ICAROB 2022

PROCEEDINGS OF THE 2022 INTERNATIONAL CONFERENCE ON ARTIFICIAL LIFE AND ROBOTICS

January 20 to 23, 2022
on line, Oita in Japan
27th AROB International Meeting Series

Editor-in-Chief
Masanori Sugisaka
Editors: Yingmin Jia, Takao Ito, Ju-Jang Lee
ISBN 978-4-9908350-7-1

ARTIFICIAL LIFE AND ROBOTICS

(ICAROB2022)

January 20-23, 2022, on line in Japan, 2022
27th AROB International Meeting Series

Editor-in-Chief
Masanori Sugisaka
Editors: Yingmin Jia, Takao Ito, Ju-Jang Lee
ISBN 978-4-9908350-7-1

Contents

1	Organization, etc.	1
2	Messages	12
3	Time Table	17
4	Opening Ceremony	20
5	Technical paper index	21
6	Abstracts	
6-1	PS abstracts	47
6-2	IS abstracts	48
6-3	OS abstracts	50
6-4	GS abstracts	120
7	Authors index	127

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

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

ADVISORY COMMITTEE

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 CHAIR

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

PROGRAM CHAIRMAN

Makoto Sakamoto (University of Miyazaki, Japan)

CO-PROGRAM CHAIR

Marion Oswald (Technische Universität Wien, Austria)

INTERNATIONAL ORGANIZING COMMITTEE

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)
Kuo-Lan Su (National Yunlin University of Science and Technology, Taiwan)
Kyungho Park (U.S. ARMY, Japan and USA)
Masao Kubo (National Defense Academy of Japan, Japan)
Masanao Obayashi (Yamaguchi University, Japan)
Maxim Sokolov (Innopolis University, Russia)
Mehta Rajiv (New Jersey Institute of Technology, USA)
Minoru Kumano (University of Miyazaki, Japan)
Peter Sapaty (Ukrainian Academy of Science, Ukraine)
Pierre Parrend (University of Strasbourg, France)
Qiang Cai (Beijing Technology and Business University, P. R. China)
Qu Yanbin (Harbin Institute of Technology, P. R. China)
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 COMMITTEE

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)
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)
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)
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)
Youji Kawamura (Kindai University, Japan)
Yu-an Zhang (Qinghai University, P. R. China)
Yue Lin (Beijing Institute of Control Engineering, P.R. China)
Yueqian Liang (China Academy of Elecronics and Information Technology, P. R. China)
Yuji Minami (National Institute of Technology, Ube College, Japan)
Yunju Chen ((Shiga University, Japan)
Yunzhong Song (Henan Polytechnic University, P. R. China)
Zacharie Mbaitiga National Institute of Technology, Okinawa College, Japan)
Zakri bin Ghazali (Universiti Malaysia Pahang, Malaysia)
Zengqiang Chen (Nankai University, P. R. China)
Zhao Peng (Huazhong University of Science and Technology, P. R. China)
Zhengquan Yang (Civil Aviation University of China, P. R. China)
Zhongxin Liu (Nankai University, P. R. China)
Zuradzman bin Mohamad Razlan (Universiti Malaysia Perlis, Malaysia)

LOCAL ARRANGEMENT COMMITTEE

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

HISTORY

The International Conference on Artificial Life and Robotics (ICAROB) resulted from the AROB-symposium (International Symposium on Artificial Life and Robotics) whose first edition was held in 1996 and the eighteenth and last edition in 2013. The AROB symposium was annually organized by Oita University and ALife Robotics Corporation Ltd., under the sponsorship of the Science and Technology Policy Bureau, the Ministry of Education, Science, Sports, and Culture (Monbusho), presently, the Ministry of Education, Culture, Sports, Science, and Technology (Monkasho), Japanese Government, Japan Society for the Promotion of Science (JSPS), the Commemorative Organization for the Japan World Exposition ('70), Air Force Office of Scientific Research, Asian Office of Aerospace Research and Development (AFOSR/AOARD), USA. I would like to express my sincere thanks to not only Monkasho (annually fund support from 1996 to 2013) but also JSPS, the Commemorative

Organization for the Japan World Exposition ('70), and various other Japanese companies for their repeated support. The old symposium (this symposium has been held every year at B-Con Plaza, Beppu, Oita, Japan except in Oita, Japan (AROB 5th ’00) and in Tokyo, Japan (AROB 6th ’01).) was organized by the International Organizing Committee of AROB and was co-operated by the Santa Fe Institute (USA), RSJ, IEEJ, ICASE (Now ICROS) (Korea), CAAI (P. R. China), ISCIE, IEICE, IEEE (Japan Council), JARA, and SICE. The old AROB-symposium expanded much by absorbing much new knowledge and technologies into it. This history and character of the former AROB symposiums are passed on the current ICAROB conference and to these journals, Journal of Robotics, Networking and Artificial Life (JRNAL) & Journal of Advances in Artificial Life Robotics (JAALR). From now on, ALife Robotics Corporation Ltd. is in charge of management of both the conference and the journals. The future of the ICAROB is brilliant from a point of view of yielding new technologies to human society in the 21st century. We also expect to establish an international research institute on Artificial Life and Robotics in the future with the help of Japanese Government and ICAROB. This conference invites you all.

AIMS AND SCOPE

The objective of this conference is the development of new technologies for artificial life and robotics which have been recently born in Japan and are expected to be applied in various fields. This conference presents original technical papers and authoritative state-of-the-art reviews on the development of new technologies concerning robotics, networking and artificial life and, especially computer-based simulation and hardware for the twenty-first century. This conference covers a broad multidisciplinary field, including areas such as:
Artificial intelligence & complexity
Artificial living
Artificial mind research
Artificial nervous systems for robots
Artificial sciences
Bipedal robot
Brain science and computing
Chaos
Cognitive science
Computational Molecular biology
Computer graphics
Data mining
Disasters robotics
DNA computing
Empirical research on network and MOT
Environment navigation and localization
Evolutionary computations
Facial expression analysis, music recommendation and augmented reality
Foundation of computation and its application

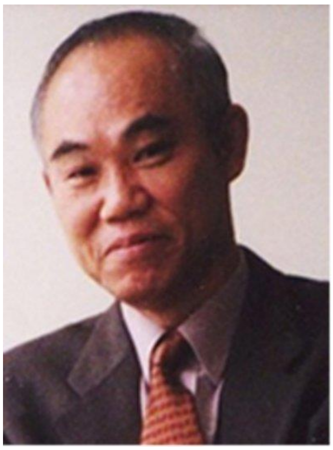
Fuzzy control
Genetic algorithms
Human-welfare robotics
Image processing
Insect-like aero vehicles
Intelligence in biological systems
Intelligent control
Management of technology
Medical surgical robot
Micro-machines
Multi-agent systems
Nano-biology
Nano-robotics
Networking
Neural circuits
Neuro-computer
Neuromorphic Systems
Neuroscience
Pattern recognition
Quantum computing
Reinforcement learning system & genetic programing
Robotics
Software development support method
System cybernetics
Unmanned underwater vehicles
Unmanned Aerial Systems Technologies
Unmanned Aerial Systems designing, controls and navigation
Unmanned Aero vehicles
Virtual reality
Visualization
Hardware-oriented submissions are particularly welcome. This conference will discuss new results in the field of artificial life and robotics

COPYRIGHTS

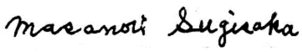
Accepted papers will be published in the proceeding of The 2022 International Conference on Artificial Life and Robotics (ICAROB2022) by ALife Robotics Corp. Ltd. Copyright belongs to ALife Robotics Corp. Ltd. Some of high-quality papers in the proceeding will be requested to re-submit their papers for the consideration of publication in Journal of Robotics, Networking and Artificial Life (JRNAL) & 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/

MESSAGES



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



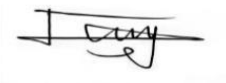
Masanori Sugisaka

General Chair of ICAROB

It is my great honor to invite you all to The 2022 International Conference on Artificial Life and Robotics (ICAROB 2022). 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 2022 International Conference on Artificial Life and Robotics (ICAROB2022) (27th 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 (ICAROB2022) will yield new merged technologies for happiness of human beings and, hence, will facilitate the establishment of an international joint research institute on Artificial Life and Robotics in future.

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



Yingmin Jia

Co-General Chair of ICAROB

It is my great pleasure to invite you to The 2022 International Conference on Artificial Life and Robotics (ICAROB 2022), Oita, Japan, January 20-23, 2022. Because of the influence of COVID-19, The ICAROB2022 will be held on-line again, and your understanding and support will be the biggest 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 27th birthday in 2022. 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 ICAROB2022.

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

Takao Ito

Takao Ito

Co General Chair of ICAROB

It is my great honor to invite you all to the 2022 International Conference on Artificial Life and Robotics (ICAROB 2022) which will be held online from January 20 to 23, 2022.

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 many 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 2022 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.

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

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 9th ICAROB 2022 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 9th ICAROB.

I'm looking forward to meeting you at the 9th ICAROB in on line and wishing you all the best.

GENERAL SESSION TOPICS

GS1 Control System Applications (6)	GS2 Learning Methods (5)
GS3 Robotics (3)	GS4 Applications (3)

ORGANIZED SESSION TOPICS

OS1 Advanced Information Processing Applications (4)	OS2 Advanced studies of interdisciplinary approach (4)
OS3 Intelligent Systems and Robotics-1 (6)	OS4 Intelligent Systems and Robotics -2- (5)
OS5 Intelligent Control (5)	OS6 Bio-inspired Artificial Vision -Systems and Applications- (5)
OS7 Media Information Processing, Music Recommendation and Artificial Intelligence (4)	OS8 Intelligent Systems and Robotics -3- (4)
OS9 Robot Control (10)	OS10 Software Development Support Method (4)
OS11 Applications of deep learning algorithms (9)	OS12 Intelligent Measurements and Control Systems (8)
OS13 Mathematical Informatics (5)	OS14 Biomedical Systems (4)
OS15 Artificial Intelligence for Embedded Systems and Robotics (5)	OS16 Robotic Manipulation (4)
OS17 Advanced Robotics (8)	OS18 Natural Computing (3)
OS19 Industrial Artificial Intelligence Robotics (7)	OS20 Application studies on sensor and RFID system (6)
OS21 Visual Signal Processing and Human-welfare Robotics I&II (6)	OS22 Advanced studies of network engineering (4)
OS23 Applications in Complex Systems (7)	OS24 Artificial Systems and Life (5)
OS25 Information Applications and Cybersecurity (6)	OS26 Intelligent Life and Data Analysis (6)
OS27 Environmental Monitoring (5)	OS28 Robot Competitions and Education (6)
OS29 Advances in Marine Robotics and Their Applications (8)	OS30 Intelligent Systems and Robotics -4- (6)
OS31 Approaches to Post-Narratology that Combines AI and Cognitive Science with Narratology (9)	OS32 Human-Machine Interface and Automation (9)
OS33 Signal Processing and Chaotic System (5)	OS34 Robotics Navigation and Control (3)

TIME TABLE (1/20)

in Japan time

1/20(Thu.)	**All Rooms meeting passcode(全部屋のミーティングのパスコードは 2022) is 2022** Group meeting for the conference (Conference Room ZOOM ID: 867 9702 9629)
1/23(Sun)	Group meeting for the next conference (Conference Room ZOOM ID: 867 9702 9629)

TIME TABLE (1/21)

1/21(Fri.)	**Meeting Room 1** (ZOOM ID: 880 7966 0454)	**Meeting Room 2** (ZOOM ID: 893 6690 0965)	**Meeting Room 3** (ZOOM ID: 816 1412 9634)	**Meeting Room 4** (ZOOM ID: 882 7162 1954)
8:40-	Registration			
9:00-10:00	OS1 Advanced Information Processing Applications (4) Chair: Toru Hiraoka	OS22 Advanced studies of network engineering (4) Chair: Wei Hong Lim	OS23 Applications in Complex Systems (7) Chair: Masao Kubo Will be end at 10:45	OS34 Robotics Navigation and Control (3) Chair: Jiwu Wang
10:00-10:20	Coffee break			
10:20-10:50	Chair : Sakamoto Makoto Opening Ceremony (Conference Room ZOOM ID: 867 9702 9629)			
11:00-12:00	Chair : Yingmin Jia (Conference Room ZOOM ID: 867 9702 9629) Plenary Speech PS2 H. Yamamoto			
12:00-13:00	Lunch			
13:00-14:00	OS2 Advanced studies of interdisciplinary approach (4) Chair: Yousin Park	OS10 Software Development Support Method (4) Chair: Tetsuro Katayama	OS14 Biomedical Systems (4) Chair: Taro Shibanoki	OS16 Robotic Manipulation (4) Chair: Kensuke Harada
14:00-14:20	Coffee break			
14:20-15:50	OS4 Intelligent Systems and Robotics -2- (5) chair: Haokang Wen OS30 Intelligent Systems and Robotics -4- (6) chair: Jichao Zhao OS33 Signal Processing and Chaotic System (5) chair: Huailin Zhao	GS3 Robotics (3) GS4 Applications (3) GS3 Chair: Hideyuki Tanaka GS4 Chair: Noritaka Sato	OS25 Information Applications and Cybersecurity (6) Chair: I-Hsien Liu	OS28 Robot Competitions and Education (6) Chair: Kazuo Ishii
15:50-16:10	coffee break			
16:10-17:10	Chair: Kazuo Ishii (Conference Room ZOOM ID: 867 9702 9629) Plenary Speech PS1 Henrik Hautop Lund			
17:10-17:30	Coffee break			

17:30-18:30	Chair: Rishav Bose & Luigi Pagliarini (Conference Room) Invited Speech IS1, IS2 Luigi Pagliarini, Rishav Bose

TIME TEBLE (1/22)

1/22(Sat.)	**Meeting Room 1** (ZOOM ID: 880 7966 0454)	**Meeting Room 2** (ZOOM ID: 893 6690 0965)	**Meeting Room 3** (ZOOM ID: 816 1412 9634)	**Meeting Room 4** (ZOOM ID: 882 7162 1954)
8:40-	Registration			
9:00-10:30	OS13 Mathematical Informatics (5) Chair: Takao Ito	OS6 Bio-inspired Artificial Vision – Systems and Applications- (5) Chair: Shinsuke Yasukawa	OS15 Artificial Intelligence for Embedded Systems and Robotics (5) Chair: Hakaru Tamukoh	GS1 Control System Applications (6) Chair: Hiroaki Wagatsuma
10:30-10:45	Coffee break			
10:45-12:00	OS5 Intelligent Control (5) Chair: Yingmin Jia	GS2 Learning Methods (5) Chair: Masato Nagayoshi	OS24 Artificial Systems and Life (5) Chair: Kuo-Hsien Hsia	OS27 Environmental Monitoring (5) Chair: Kazuo Ishii
12:00-13:00	Lunch			
13:00-14:00	Chair : Eiji Hayashi (Conference Room ZOOM ID: 867 9702 9629) Plenary Speech PS3 Takao Ito			
14:00-14:20	Coffee break			
14:20-15:20	OS21-1 Visual Signal Processing and Human-welfare Robotics I&II (4) Chair: Joo Kooi Tan	OS19-1 Industrial Artificial Intelligence Robotics (4) Chair: Eiji Hayashi	OS26-1 Intelligent Life and Data Analysis (4) Chair: I-Hsien Liu	OS29-1 Advances in Marine Robotics and Their Applications (4) chair: Keisuke Watanabe
15:20-15:30	Coffee break			
15:30-16:30	OS21-2 Visual Signal Processing and Human-welfare Robotics I&II (2) Chair: Joo Kooi Tan	OS19-2 Industrial Artificial Intelligence Robotics (3) Chair: Eiji Hayashi	OS26-2 Intelligent Life and Data Analysis (2) Chair: I-Hsien Liu	OS29-2 Advances in Marine Robotics and Their Applications (4) chair: Keisuke Watanabe
16:30-16:40	Coffee break			
16:40-17:40	OS11-1 Applications of deep learning Algorithms (4) Chair: Xiaoyan Chen	OS18 Natural Computing (3) Chair: Marion Oswald	OS17-1 Advanced Robotics (4) Chair: Evgeni Magid	
17:40-17:50	Coffee break			
17:50-19:05	OS11-2 Applications of deep learning Algorithms (5) Chair: Xiaoyan Chen		OS17-2 Advanced Robotics (4) Chair: Evgeni Magid	

TIME TABLE (1/23)

1/23(Sun.)	Meeting Room 1 (ZOOM ID: 880 7966 0454)	Meeting Room 2 (ZOOM ID: 893 6690 0965)	Meeting Room 3 (ZOOM ID: 816 1412 9634)	Meeting Room 4 (ZOOM ID: 882 7162 1954)
8:40-	Registration			
9:00-10:30	OS3 Intelligent Systems and Robotics-1 (6) chair: Peng Lu OS8 Intelligent Systems and Robotics -3- (4) chair: Tianyi Zhang		OS9-1 Robot Control (6) Chair: Yizhun Peng	OS31 Approaches to Post-Narratology that Combines AI and Cognitive Science with Narratology (6) Chair: Jumpei Ono
10:30-10:50	Coffee break			
10:50-11:50		OS7 Media Information Processing, Music Recommendation and Artificial Intelligence (4) Chair: Yasunari Yoshitomi	OS9-2 Robot Control (4) Chair: Yizhun Peng	OS31-2 Approaches to Post-Narratology (3) that Combines AI and Cognitive Science with Narratology (3) Chair: Jumpei Ono
11:50-13:00	Lunch			
13:00-14:00		OS20-1 Application studies on sensor and RFID system (4) chair: Ammar A.M. Al-Talib	OS32-1 Human-Machine Interface and Automation (4) Chair: Norrima Mokhtar	OS12-1 Intelligent Measurements and Control Systems (4) Chair: Xiaoyan Chen
14:00-14:10	Coffee break			
14:10-15:25		OS20-2 Application studies on sensor and RFID system (2) chair: Ammar A.M. Al-Talib	OS32-2 Human-Machine Interface and Automation (5) Chair: Norrima Mokhtar	OS12-2 Intelligent Measurements and Control Systems (4) Chair: Xiaoyan Chen

The 2022 International Conference on ARTIFICIAL LIFE AND ROBOTICS (ICAROB2022)

January 20 (Thursday)

17:30-19:30

Group meeting for the conference

January 21 (Friday)

10:20-10:50

Opening Ceremony

Chair: Sakamoto Makoto (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. Co-General Chairman of ICAROB	Ju-Jang Lee (KAIST, Korea)

January 23 (Sunday)

15:30-17:30

Group meeting for the next conference

TECHNICAL PAPER INDEX

January 21 (Friday)

8:40-Registration

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

11:00-12:00
Plenary Speech PS2
Chair: Yingmin Jia (Beihang University, China)

PS2 *Road To Cyber Physical Factory (Application Examples of Intelligent Factory and its Technology)*
Hidehiko Yamamoto (Gifu University, Japan)

16:10-17:10
Plenary Speech PS1
Chair: Kazuo Ishii (Kyushu Institute of Technology, Japan)

PS1 *Robotics for Growing Life*
Henrik Hautop Lund (Technical University of Denmark, Denmark)

17:30-18:00
Invited Speech IS1
Chair: Rishav Bose (Technical University of Denmark, Denmark)

IS1-1 *A physical to virtual control system implementing an art-based game*
IS1-2 *Creative Multisensory Environments*
Luigi Pagliarini (Technical University of Denmark, Denmark, Academy of Fine Arts of Macerata, Italy)

18:00-18:30
Invited Speech IS2
Chair: Luigi Pagliarini
(Technical University of Denmark, Denmark, Academy of Fine Arts of Macerata, Italy)

IS2-1 *Using GrowBots to Study Heat and Nutrient Stress in Basil*
IS2-2 *Convolutional Neural Network for Studying Plant Nutrient Deficiencies*
Rishav Bose (Technical University of Denmark, Denmark)

Meeting Room 1
9:00-10:00 OS1 Advanced Information Processing Applications (4)
Chair: Toru Hiraoka (University of Nagasaki, Japan)
Co-Chair: Masaharu Hirota (Okayama University of Science, Japan)

OS1-1 *Generation of Checkered Pattern Images Using Prewitt Filter from RGB-D Images*
Toru Hiraoka, Ryosuke Takaki (University of Nagasaki, Japan)

OS1-2 *An Approach of Analyzing Movement Patterns Using Word Embeddings from Geo-tagged Tweets*
Masaharu Hirota, Tetsuya Oda (Okayama University of Science, Japan)

OS1-3 *A Proposal of a Software Defect Predication System Using SOM*
Yoshihiro Kita[1], Kazuki Ueda[2], Kiyotaka Sakurai[2]
([1]University of Nagasaki, [2]Nihon Knowledge Co. Ltd, Japan)

OS1-4 *Mapping the Motion of Highly-inclined Triple System into a Secular Perturbation Model*
Masaya M. Saito (University of Nagasaki, Japan)
Kiyotaka Tanikawa (National Astronomical Observatory, Japan)

13:00-14:00 OS2 Advanced studies of interdisciplinary approach (4)
Chair: Yousin Park (Prefectural University of Hiroshima, Japan)
Co-Chair：Takao Ito (Hiroshima University, Japan)

OS2-1 *Measuring the entire degree centrality in Yokokai*
Tsutomu Ito[1], Matsuno Seigo[1], Sakamoto Makoto[2], Takao Ito[3]
([1]Ube National College of Technology, Japan) ([2]Univeristy of Miyazaki, Japan)
([3]Hiroshima University, Japan)

OS2-2 *A comparative study on Michinoeki's efficiency in Japan*
Tsutomu Ito[1], Matsuno Seigo[1], Sakamoto Makoto[2], Takao Ito[3]
([1]Ube National College of Technology, Japan) ([2]Univeristy of Miyazaki, Japan)
([3]Hiroshima University, Japan)

OS2-3 *The R&D Direction and Business Strategy: The case study on the cooperation of EV and battery makers*
Yousin Park[1], Iori Nakaoka[2], Yun-ju Chen[3]
([1]Prefectural University of Hiroshima, Japan) ([2]Seijoh University, Japan)
([3]Shiga University, Japan)

OS2-4 *Discovering the relationship between tourists and tourist spots in Japan*
Tsutomu Ito[1], Matsuno Seigo[1], Sakamoto Makoto[2], Takao Ito[3]
([1]Ube National College of Technology, Japan) ([2]Univeristy of Miyazaki, Japan)
([3]Hiroshima University, Japan)

14:20-15:50 OS4 Intelligent Systems and Robotics -2- & OS30 Intelligent Systems and Robotics -4- & OS33 Signal Processing and Chaotic System

OS4 Intelligent Systems and Robotics -2- (5)

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

OS4-1 *A Research on Image Dehazing Technology for Image Enhancement*
Haokang Wen, Chang Sheng (Tianjin University of Science and Technology, China,)

OS4-2 *Design of Intelligent Daylily Picking Robot*
Jiaxin Li (Tianjin University of Science and Technology, China)

OS4-3 *50KN Compression Spring Fatigue Testing Machine Design*
Peng Lu, Peng Jia (Tianjin University of Science and Technology, China)

OS4-4 *New Intelligent Unmanned Retail Shopping Container Design*
Peng Lu, Yiting Gao (Tianjin University of Science and Technology, Tianjin, China)

OS4-5 *Java-based Dream Cloud ERP System - Inventory Management Subsystem Design and Implementation*
Yiting Gao, Peng Lu (Tianjin University of Science and Technology, China)

OS30 Intelligent Systems and Robotics -4- (6)

Chair Jichao Zhao (Tianjin University of Science and Technology, China)
Co-Chair Hucheng Wang (Tianjin University of Science and Technology, China)

OS30-1 *A Design of Micromouse Control System*
Hucheng Wang (Tianjin University of Science and Technology, China)

OS30-2 *A Driver Reaction Time Detection System Design*
Yuhui Cheng[1], Mochi Li[2] ([1] Tianjin University of Science and Technology, China)
([2] Ocean University of China, China)

OS30-3 *A PID Tracking Car Design based on STM32*
Yande Xiang, Teng Zhang, Shixiang Zhao, Ling Zhou, Mingjuan Tian, Haoran Gong
(Tianjin University of Science and Technology, China)

OS30-4 *Matrix Approach to Current-state Detectability of Discrete-event Systems*
Jinliang Wang, Jiawei Wei, Xiaoguang Han
(Tianjin University of Science and Technology, China)

OS30-5 *Hardware Circuit Design Of Tracking Car Based On K60*
Peng Jia, Yande Xiang (Tianjin University of Science and Technology, China)

OS30-6 *Detachable IoT Garbage Sorting Device Based on Machine Vision*
Tao Zhu, Yang Su, Zhiqing Xiao, Fengzhi Dai
(Tianjin University of Science and Technology, China)

OS33 Signal Processing and Chaotic System (5)
Chair: Huailin Zhao (Shanghai Institute of Technology, China)
Co-Chair: Fengzhin Dai (Tianjin University of Science and technology, China)

OS33-1 *A Visual Measurement Algorithm of Approaching Vehicle Speed Based on Deep Learning*
Yurong Zhu, Huailin Zhao, Liu Junjie, Zhang Jinping, Ji Xiaojun
(Shanghai Institute of Technology, China)

OS33-2 *Target Search Based on Scene Priors*
Shengyang Lu, Lanjun Liang, Huailin Zhao, Fangbo Zhou, Feng yao
(Shanghai Institute of Technology, China)

OS33-3 *A Generalized Hamiltonian Conservative Systems with Multi-scroll Chaotic Flows*
Jingwen Liu, Zhonggao Chen (Tianjin University of Science and Technology, China)

OS33-4 *Multi-stability and FPGA Implementation of a Conservative Chaotic System*
Minghan Song (Tianjin University of Science and Technology, China)

OS33-5 *A New Hyperchaotic Financial System*
Lei Gong (Tianjin University of Science and Technology, China)

Meeting Room 2
9:00-10:00 OS22 Advanced studies of network engineering (4)
Chair Wei Hong Lim (UCSI University, Malaysia)
Co-Chair Takao Ito (Hiroshima University, Japan)

OS22-1 *New Hybridization Algorithm of Differential Evolution and Particle Swarm Optimization for Efficient Feature Selection*
Koon Meng Ang, Mohd Rizon Bin Mohamed Juhari, Wei Hong Lim, Sew Sun Tiang, Chun Kit Ang, Eryana Eiyda Hussin, Li Pan, Ting Hui Chong
(UCSI University, Malaysia)

OS22-2 *River Water Quality Monitoring Through LoRa Network*
Syarifah Nabilah Syed Taha Tahir, Mohamad Sofian Abu Talip, Mahazani Mohamad, Mohamadariff Othman, Tengku Faiz Tengku Mohmed Noor Izam, Mohd Faiz Mohd Salleh, Zati Hakim Azizul Hasan, Zeeda Fatimah Mohamad, Amir Feisal Merican Aljunid Merican (Universiti Malaya, Malaysia)

OS22-3 *Wideband Antenna with UHF Sensor Applicability for MV/HV Equipment in Smart-Grid Systems*
S. M. Kayser Azam[1], Mohamadariff Bin Othman[1], Tarik Abdul Latef[1], H. A. Illias[1], Mohd Fadzil Ain[2], Yazeed Qasaymeh[3] ([1]Universiti Malaya, Malaysia)
([2]Universiti Sains Malaysia, Malaysia) ([3]Majmaah University, Saudi Arabia)

OS22-4 *New Particle Swarm Optimization Variant with Modified Neighborhood Structure*
Koon Meng Ang[1], Mohd Rizon Bin Mohamed Juhari[1], Wy-Liang Cheng[1], Wei Hong Lim[1]*, Sew Sun Tiang[1], Chin Hong Wong[2], Hameedur Rahman[3], Li Pan[1]
([1]UCSI University, Malaysia) ([2]Fuzhou University, China)
([3]Air University, Islamabad, Pakistan)

13:00-14:00 OS10 Software Development Support Method (4)
Chair: Tetsuro Katayama (University of Miyazaki, Japan)
Co-Chair: Tomohiko Takagi (Kagawa University, Japan)

OS10-1 *N-Switch and All-Path Test Coverage Criterion for Extended Finite State Machine*
Tomohiko Takagi[1], Koichiro Sakata[1], Kouichi Akiyama[2]
([1]Kagawa University, Japan), ([2]Japan WillTech Solution Co., Ltd., Japan)

OS10-2 *Proposal of a Method to Generate Classes and Instance Variable Definitions in the VDM++ Specification from Natural Language Specification*
Kensuke Suga[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)

OS10-3 *Expansion of Application Scope and Addition of a Function for Operations into BWDM to Generate Test Cases from VDM++ Specification*
Takafumi Muto[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)

OS10-4 *Proposal of Gamma which is a spatial data sharing distributed MQTT system*
Takahiro Ueda[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)

14:20-15:50 GS3 Robotics & GS4 Applications
GS3 Robotics (3)

GS3-1 *A Three-Dimensional Design of the Multi-material Joint System to Realize a Structural Spring-Damper Compliant Mechanism with Versatility in Engineering Fields*
Pancho Dachkinov[1], Anirudha Bhattacharjee[2], Bishakh Bhattacharya[2], Hiroaki Wagatsuma[1,3]
([1]Kyushu Institute of Technology, Japan; [2]Indian Institute of Technology Kanpur, India; [3]RIKEN CBS, Japan)

GS3-2 *Haptic Device that Presents Sensation Corresponding to Palm on Back of Hand for Teleoperation of Robot Hand Report 5: Verification of development device specifications*
Kyosuke Ushimaru, Noritaka Sato (Nagoya Institute of Technology, Japan)

GS-3-3 *HBV Epidemic Control Using Time-Varying Sliding Mode Control Method*
Arsit Boonyaprapasorn[1], Suwat Kuntanapreeda[2], Parinya Sa Ngaimsunthorn[2], Thunyaseth Sethaput[3], Tinnakorn Kumsaen[4]
([1]Chulachomklao Royal Military Academy, Thailand)
([2]King Mongkut's University of Technology North Bangkok, Thailand)
([3]Thammasat University, Thailand) ([4]Khon Kaen University, Thailand)

GS4 Applications (3)

Chair: Noritaka Sato (Nagoya Institute of Technology, Japan)

GS4-1 *Blockchain Technology for Halal Supply Chain Management*
Kadeer Zulihuma, Abdul Samad Shibghatullah, Chit Su Mon
(UCSI University, Malaysia)

GS4-2 *Smart Tourism Guide Application Using Location-Based Services – Go.Travel*
Wong Yit Meng, Abdul Samad Bin Shibghatullah, Kasthuri Subaramaniam
(UCSI University, Malaysia)

GS4-3 *Gesturenomy: Touchless Restaurant Menu Using Hand Gesture Recognition*
Ian Christian Susanto, Kasthuri Subaramaniam, Abdul Samad Bin Shibghatullah
(UCSI University, Malaysia)

Meeting Room 3
9:00-10:45 OS23 Applications in Complex Systems (7)
Chair: T Masao Kubo (National Defense Academy, Japan)
Co-Chair: Hiroshi Sato (National Defense Academy, Japan)

OS23-1 *A research of infectivity rate After the Consecutive Holidays*
Saori Iwanaga (Japan Coast Guard Academy, Japan)

OS23-2 *Towards the Trusted Population-Based Optimization Systems*
Hiroshi Sato, Masao Kubo (National Defense Academy, Japan)

OS23-3 *Spatio-temporal prediction of crime occurrence spots by CNN-LSTM*
Kaede Yaji, Masao Kubo, Hiroshi Sato (National Defense Academy, Japan)

OS23-4 *Cross-View Image Geo-Localization using Multi-Scale Generalized Pooling with Attention Mechanism*
Duc Viet Bui, Masao Kubo, Hiroshi Sato (National Defense Academy, Japan)

OS23-5 *Recommendation a Emergency Patient Destinations by LightGBM*
Ryota Kawaguchi[1], Masao Kubo[1], Hiroshi Sato[1], Daizoh Saitoh[2], Takao Oya[3]
([1]National Defense Academy, [2]National Defense Medical College, [3]Saitamaseibu Fire Bureau, Japan)

OS23-6 *A Framework for Understanding the Neural Underpinnings of Symbolic and Non-Symbolic Communication Based on Global Synchronization in Human Brain Activity*
Masayuki Fujiwara, Takashi Hashimoto (Japan Advanced Institute of Science and Technology, Japan)

OS23-7 *Characterization of randomness tests by using tests results of weakly correlated chaotic sequences*
Akihiro Yamaguchi (Fukuoka Institute of Technology), Asaki Saito (Future University Hakodate, Japan)

13:00-14:00 OS14 Biomedical Systems (4)
Chair: Taro Shibanoki (Okayama University, Japan)

OS14-1 *A Mutual Control Method for a Multi-layered Non-contact Impedance Model-based mobile robots*
Masaru Sasaki[1], Taro Shibanoki[2], Hideyuki Tonooka[1], Toshio Tsuji[3]
([1]Ibaraki University, Japan), ([2]Okayama University, Japan)
([3]Hiroshima University, Japan)

OS14-2 *Relationship Between Delay Time and Sensation in Tactile Feedback for Myoelectric Prosthesis*
Taro Shibanoki (Okayama University, Japan), Kosuke Jin (Ibaraki University, Japan)

OS14-3 *Effects of Tactile Stimulation Near the Auricle on Body Sway During Foot Stamping*
Masaya Tadokoro[1], Taro Shibanoki[2], Hideyuki Tonooka[1]
([1]Ibaraki University, Japan), ([2]Okayama University, Japan)

OS14-4 *A Monitoring System of a Hamster Based on Video Image Analysis*
Yugo Yamazaki[1], Taro Shibanoki[2], Hideyuki Tonooka[1]
([1]Ibaraki University, Japan), ([2]Okayama University, Japan)

14:20-15:50 OS25 Information Applications and Cybersecurity (6)
Chair: I-Hsien Liu (National Cheng Kung University, Taiwan)
Co-Chair: Kuo-Hsien Hsia (National Yunlin University of Science and Technology, Taiwan)

OS25-1 *Extendable ICS Honeypot Design with Modbus/TCP*
I-Hsien Liu, Jun-Hao Lin, Hsin-Yu Lai, Jung-Shian Li (National Cheng Kung University, Taiwan)

OS25-2 *Industrial Control System Cybersecurity Testbed with TSN Feature*
I-Hsien Liu, Li-Yin Chang, Jung-Shian Li (National Cheng Kung University, Taiwan)
Chuan-Gang Liu (Chia Nan University of Pharmacy & Science, Taiwan)

OS25-3 *Using the Modified Delphi Method to Construct the Quality Indicators of the Counseling Service System*
Li-Min Chuang, Hsiu-Hao Liu (Chang Jung Christian University, Taiwan)

OS25-4 *The Key Success Factors of Introducing ERP System in Taiwan's Manufacturing Industry*
Li-Min Chuang, Yu-Po Lee (Chang Jung Christian University, Taiwan)

OS25-5 *The Fuzzy AHP approach for intelligent building assessment model*
Li-Min Chuang, Yu-Po Lee (Chang Jung Christian University, Taiwan),
Chien-Chih Kuo (Chien Chang Construction Co., Ltd.)

OS25-6 *Blockchain-based Verification Mechanism for Industrial Control System*
Yao-Chu Tsai, I-Hsien Liu and Jung-Shian Li (National Cheng Kung University, Taiwan)

Meeting Room 4
9:00-9:45 OS34 Robotics Navigation and Control (3)
Chair: Jiwu Wang (Beijing Jiaotong University, China)
Co-Chair: Shilong Zhen (Beijing Jiaotong University, China)

OS34-1 *Research on Path Planning Algorithms of Multiple Mobile Robots in Intelligent Warehousing*
Jiwu Wang, Shilong Zheng (Beijing Jiaotong University, China)

OS34-2 *Research on the effectiveness of improved ORB depth estimation in monocular vision slam*
Jiwu Wang, Weipeng Wan (Beijing Jiaotong University, China)

OS34-3 *Research on Research on Corner Detection Algorithm Based on Edge Contour in Automatic Loading Positioning*
Jiwu Wang, Junwei Fu (Beijing Jiaotong University, China)

13:00-14:00 OS16 Robotic Manipulation (4)
Chair: Kensuke Harada (Osaka University, Japan)
Co-Chair: Tokuo Tsuji (Kanazawa University, Japan)
Co-Chair: Akira Nakamura (Saitama Institute of Technology, Japan)

OS16-1 *Motion Planning for Retrieving an Object in a Complex Environment*
Shusei Nagato, Tomohiro Motoda, Keisuke Koyama, Weiwei Wan, Kensuke Harada
(Osaka University, Japan)

OS16-2 *Design and Control of Two-sided Gripper for Bin Picking*
He Maike, Tokuo Tsuji, Tatsuhiro Hiramitsu, Hiroaki Seki
(Kanazawa University, Japan)

OS16-3 *Training Data Augmentation for Semantic Segmentation of Food Images Using Deep Learning*
Takayuki Yamabe, Tatsuya Ishichi, Tokuo Tsuji, Tatsuhiro Hiramitsu, Hiroaki Seki
(Kanazawa University, Japan)

OS16-4 *Suitable Error Recovery Process using Combined Evaluation Standards in Robotic Manufacturing Plant*
Akira Nakamura[1], Kensuke Harada[2]
([1]Saitama Institute of Technology, [2]Osaka University, Japan)

14:20-15:50 OS28 Robot Competitions and Education (6)

Chair: Kazuo Ishii (Kyushu Institute of Technology, Japan)

Co-Chair: Yasunori Takemura (Nishinippon Institute of Technology, Japan)

OS28-1 *Underwater Acoustic Positioning Based on MEMS Microphone for a Portable Autonomous Underwater Vehicle*
Irmiya R. Inniyaka, Dominic B. Solpico, Daiki Hamada, Akihiro Sugino, Rikuto Tanaka, Yuya Nishida, Kazuo Ishii (Kyushu Institute of Technology, Japan)

OS28-2 *Autonomous Underwater Vehicle with Vision-based Navigation System for Underwater Robot Competition*
Kazuki Harada, Riku Fukuda, Yusuke Mizoguchi, Yusuke Yamamoto, Kouta Mishima, Yoshiki Tanaka, Yuya Nishida, Kazuo Ishii (Kyushu Institute of Technology, Japan)

OS28-3 *Tomato Harvesting in Greenhouse Considering the Effect of Sunlight*
Kai Shioji, Shinsuke Yasukawa, Kazuo Ishii (Kyushu Institute of Technology, Japan)

OS28-4 *Inter-University Collaboration Aimed at Integrating Different Robotic Fields: Development of Underwater Robots and Soccer Robots Through these Competitions*
Moeko Tominaga[1], Jonghyun Ahn[2], Yasunori Takemura[1], Kazuo Ishii[3]
([1]Nishinippon Institute of Technology, Japan, [2]Hiroshima Institute of Technology, Japan, [3]Kyushu Institute of Technology, Japan)

OS28-5 *Exercise on Environmental Monitoring and Control of Greenhouse by IoT Devices toward Smart Agriculture*
Yuya Nishida, Ryugo Mochizuki, Shinsuke Yasukawa, Kazuo Ishii
(Kyushu Institute of Technology, Japan)

OS28-6 *Evaluation of roller arrangement of sphere by omnidirectional integral value*
[1]Kenji Kimura, [2]Yusuke Abematsu, [3]Hirai Hiroyasu, [3]Kazuo Ishi
([1]Fukuoka Daiichi High School, Japan, [2]Kagoshima Gyokuryu High School, Japan,
[2]Kyushu Institute of Technology, Japan)

January 22 (Saturday)

8:40-Registration

Conference Room

13:00-14:00

Plenary Speech PS3

Chair: Eiji Hayashi (Kyushu Institute of Technology, Japan)

PS3 *Robot Technology, and it's Development Trend–Developing a New Networking Robot System-*
Takao Ito (Hiroshima University, Japan)

Meeting Room 1
9:00-10:15 OS13 Mathematical Informatics (5)
Chair: Takao Ito (Hiroshima University, Japan)
Co-Chair: Makoto Sakamoto (University of Miyazaki, Japan)

OS13-1 *Basic Study on Design Tool of Hula Costumes*
Taketo Kamasaka[1], Kodai Miyamoto[1], Makoto Sakamoto[1], Satoshi Ikeda[1], Amane Takei[1], Kenji Aoki[1], Tsutomu Ito[2], Takao Ito[3] ([1]University of Miyazaki, Japan), ([2]National Institute of Technology, Ube College, Japan)
([3]Hiroshima University, Japan)

OS13-2 *Basic Study on the Use of XR Technology to Support Science Education*
Kodai Miyamoto[1], Taketo Kamasaka[1], Makoto Sakamoto[1], Masahiro Yokomichi[1], Satoshi Ikeda[1], Amane Takei[1], Tsutomu Ito[2], Takao Ito[3]
([1]University of Miyazaki, Japan), ([2]National Institute of Technology, Ube College, Japan), ([3]Hiroshima University, Japan)

OS13-3 *Analysis of 5x5 board Quoridor*
Takuro Iwanaga, Makoto Sakamoto, Takao Ito, Satoshi Ikeda
(University of Miyazaki, Japan)

OS13-4 *A perfect play in 4×12 board of Othello*
Tomoyasu Toshimori, Makoto Sakamoto, Takao Ito, Satoshi Ikeda
(University of Miyazaki, Japan)

OS13-5 *Parallel full-wave electromagnetic field analysis based on domain decomposition method*
Amane Takei, Kento Ohnaka, Makoto Sakamoto (University of Miyazaki, Japan)

10:45-12:00 OS5 Intelligent Control (5)
Chair: Yingmin Jia (Beihang University, China)
Co-Chair: Weicun Zhang (University of Science and Technology Beijing, China)

OS5-1 *Geometry Structure Oriented Nonlinear Internal Model Based Manifold Consensus*
Yunzhong Song[1], Weicun Zhang[2], Fengzhi Dai[3], Huimin Xiao[4], Shumin Fei[5]
([1]Henan Polytechnic University, China), ([2]University of Science and Technology Beijing, China)
([3]Tianjin University of Science and Technology, China), ([4]Henan University of Economics and Law, China), ([5]South East University, China)

OS5-2 *Consensus Control of Linear Discrete-time Multi-agent Systems with Limited Communication Data Rate*
Jintao Hu, Yingmin Jia, Yaxin Li (Beihang University, China)

OS5-3 *Formation control for rectangular agents with communication maintenance and collision avoidance*
Yaxin Li, Yingmin Jia, Jintao Hu (Beihang University, China)

OS5-4 *Research on Sign Language Recognition Algorithm Based on Improved R(2+1)D*
Yueqin Sheng[1], Qunpo Liu[1], Naohiko Hanajima[2], Bo Su[1]
([1]Henan Polytechnic University, China), ([2]Muroran Institute of Technology, Japan)

OS5-5 *Directional Flocking of Multi-Agent system Caused by Limited Visual Field*
Yongnan Jia[1], Yong Xie[2], Weicun Zhang[1]
([1]University of Science and Technology Beijing, China)
([2]The System Design Institute of Mechanical-Electrical Engineering, China)

14:20-16:15 OS21 Visual Signal Processing and Human-welfare Robotics I&II (7)
Chair: Joo Kooi Tan (Kyushu Institute of Technology, Japan)
Co-Chair: Nobuo Sakai (Kyushu Institute of Technology, Japan)

OS21-1 *Object Acquisition Based on Human-Robot Cooperation*
Kota Ito, Masuhiro Nitta, Seiji Ishikawa, Joo Kooi Tan
(Kyushu Institute of Technology, Japan)

OS21-2 *Development of Musculoskeletal Walking Simulator for Analysis of Human Walking and Rehabilitation*
Nobuo Sakai, Yukiho Ryu, Tsubasa Ikeda, Mochimitsu Komori (Kyushu Institute of Technology), Masako Fuchi (Kyushu Nutrition Welfare University), Katsuki Hayashi (Seiai Rehabilitation Hospital)

OS21-3 *Collision Avoidance in a Human-Robot Coexistence Food Preparation Environment Using Hands Area Extraction*
Takaaki Yotsumoto, Yuta Ono, Masuhiro Nitta, Joo Kooi Tan
(Kyushu Institute of Technology, Japan)

OS21-4 *Supporting Safe Walk at a Railway Station Platform for a Visually Impaired Person Based on MY VISION*
Yuki Kawaguchi, Seiji Ishikawa, Takashi Shinomiya*, Joo Kooi Tan
(Kyushu Institute of Technology, Japan, *N&N Inc., Japan)

OS21-5 *Detecting a Pedestrian's Walk Direction Using MY VISION for Supporting Safe Walk of a Visually Impaired Person*
Shinya Iizumi, Yuta Ono, Masuhiro Nitta, Seiji Ishikawa, Joo Kooi Tan
(Kyushu Institute of Technology, Japan)

OS21-6 *Fruits and Vegetables detection using the improved YOLOv3*
Changhua Xu, Ziyue Liu, Masuhiro Nitta, Joo Kooi Tan (Kyushu Institute of Technology, Japan)

OS21-7 *Strict frequency estimation of sinusoidal signal using sampling function(withdraw)*
Masuhiro Nitta (Kyushu Institute of Technology, Japan)

16:40-19:05 OS11 Applications of deep learning algorithms (8)
Chair: Xiaoyan Chen (Tianjin University of Science and Technology, China)
Co-Chair: Shiming Wang (Tianjin University of Science and Technology, China)

OS11-1 *Human-vehicle detection based on YOLOv5*
Zhihui Chen[1], Xiaoyan Chen[1], Xiaoning Yan[2], Shuangwu Zheng[2]
([1]Tianjin University of Science and Technology, China)
([2]Shenzhen Softsz Co. Ltd., China)

OS11-2 *Low light enhancement CNN Network based on attention mechanism*
Xiwen liang[1], Xiaoyan Chen[1], Hao Feng[1], Xiaoning Yan[2], Nenghua Xu[2]
([1]Tianjin University of Science and Technology, China)
([2]Shenzhen Softsz Co. Ltd., China)

OS11-3 *Fruit Recognition Based on YOLOX*
Keying Ren[1], Xiaoyan Chen[1], Wangzi Chen[1], Xiaoning Yan[2], Dongyang Zhang[1]
([1]Tianjin University of Science and Technology, China)
([2]Shenzhen Softsz Co. Ltd., China)

OS11-4 *An improved small target detection method based on YOLOv4*
Xia Miao[1], Xiaoyan Chen[1], Keying Ren[1], Zichen Wang[1], Xiaoning Yan[2], Yue Sun[2]
([1]Tianjin University of Science and Technology, China)
([2]Shenzhen Softsz Co. Ltd., China)

OS11-5 *Yolov5-DP: A new method for detecting pedestrian aggregation*
Kunzhi Yang[1], Xiaoyan Chen[1] Xiaoning Yan[2], Dashuo Wu[2]
([1]Tianjin University of Science and Technology, China)
([2]Shenzhen Softsz Co. Ltd., China)

OS11-6 *Flame Recognition based on Yolov5 Algorithm*
Kunzhi Yang[1], Xiaoyan Chen[1] Xiaoning Yan[2], Dongyang Zhang[2]
([1]Tianjin University of Science and Technology, China)
([2]Shenzhen Softsz Co. Ltd., China)

OS11-7 *Research on face detection algorithm based on improved YOLOv5*
Zhen Mao[1], Xiaoyan Chen[1], Xiaoning Yan[2], Yuwei Zhao[2]
([1]Tianjin University of Science and Technology, China)
([2]Shenzhen Softsz Co. Ltd., China)

OS11-8 *Visibility analysis based on a novel A-VGGNet network*
Zhen Mao[1], Xiaoyan Chen[1], Xiaoning Yan[2], Shuangwu Zheng[2]
([1]Tianjin University of Science and Technology, China)
([2]Shenzhen Softsz Co. Ltd., China)

OS11-9 *Price Prediction of Diamonds*
Xiran Wen, Qiqi Xu, Zirui Su, Jiayi Fang
(The Chinese University of Hong Kong, China)

Meeting Room 2
9:00-10:15 OS6 Bio-inspired Artificial Vision -Systems and Applications- (5)
Chair: Shinsuke Yasukawa (Kyushu Institute of Technology, Japan)
Co-Chair: Yuki Hayashida (Mie University/Osaka University, Japan)

OS6-1 *The effect of preprocessing with Gabor filters on image classification using CNNs*
Akito Morita, Hirotsugu Okuno (Osaka Institute of Technology, Japan)

OS6-2 *An image coding algorithm with color constancy using the Retinex theory and the Naka-Rushton equation*
Shota Hisamitsu, Hirotsugu Okuno (Osaka Institute of Technology, Japan)

OS6-3 *A fast image sensor system with an efficient multi-scale Gaussian filtering circuit*
Yuuki Yamaji, Akito Morita, Hirotugu Okuno (Osaka Institute of Technology, Japan)

OS6-4 *A robotic vision system emulating fixational eye movements and retinal sampling*
Takanori Yotsumoto[1], Yuki Hayashida[2], Shinsuke Yasukawa[1]
([1]Kyushu Institute of Technology), ([2]Mie University, Japan)

OS6-5 *A bench-test system of the cortical prostheses utilizing retino-morphic spike as the driver signal of intracortical microstimulation*
Ryosuke Okada[1], Shinnosuke Ishikawa[2], Tetsufumi Tasaki[2], Tetsuya Yagi[2,3], Yuki Hayashida[1,2]
([1]Mie University, [2]Osaka University, [3]Fukui University of Technology, Japan)

10:45-12:00 GS2 Learning Methods (5)
Chair: Masato Nagayoshi (Niigata College of Nursing, Japan)

GS2-1 *A basic study of how to exchange work shifts using reinforcement learning on a constructive nurse scheduling system*
Masato Nagayoshi (Niigata College of Nursing, Japan)
Hisashi Tamaki (Kobe University, Japan)

GS2-2 *Developing Machine Learning and Deep Learning Models for Customer Churn Prediction in Telecomunication Industry*
Teoh Jay Shen, Abdul Samad Shibghatullah (UCSI University, Malaysia)

GS2-3 *Liver Segmentation in CT Images Using Deep-Learning and 3D CRF*
Shuntaro Nagano[1], Guangxu Li[2], Tohru Kamiya[1]
([1]Kyushu Institute of Technology, Japan), ([2]Tiangong University, Tianjin, China)

GS2-4 *Judgement on Shunt Sounds from Vascular Access using YOLO Deep Learning Model*
[1]Kyosuke Fujiwara, [1]Takayuki Yamamoto, [1]Lindsey Tate, [2]Kazuya Kibune, [1]Hiroki Tamura
([1]University of Miyazaki, Japan), ([2]Tokatsu Dialysis Hospital & Clinics, Japan)

GS2-5 *Research of Classification of Palmprint Based on Deep Learning*
Kunyu Yu, Hiroshi Matsuki (Ashikaga University, Japan)

14:20-16:15 OS19 Industrial Artificial Intelligence Robotics (7)
Chair: Eiji Hayashi (Kyushu Institute of Technology, Japan)
Co-Chair: Sakmongkon Chumkamon (Kyushu Institute of Technology, Japan)

OS19-1 *Online Deep Reinforcement Learning on Assigned Weight Spaghetti Grasping in One time using Soft Actor-Critic*
Prem Gamolped, Sakmongkon Chumkamon, Tomofumi Tsuji, Nattapat Kloomklang, Chanapol Piyavichayanon, Ranatchai Laosiripong, Eiji Hayashi (Kyushu Institute of Technology, Japan)

OS19-2 *The research about editing system of performance information for player piano. -Inference in the same phrase including ostinato-*
Haruna Yamasaki, Sakmongkon Chumkamon, Eiji Hayashi (Kyushu Institute of Technology, Japan)

OS19-3 *Weight estimation for noodle products in food layout of a home replacement meal*
Tomofumi Tsuji[1], Sakmongkon Chumkamon[1], Chanapol Piyavichayanon[1], Prem Gamolped[1], Ranatchai Laosiripong[1], Ayumu Tominaga[2], Ryusuke Fujisawa[1], Eiji Hayashi[1]
([1]Kyushu Institute of Technology, Japan)
([2] National Institute of Technology, Kitakyushu College, Japan)

OS19-4 *Cognition of surrounding conditions for a field robot Slope detection using a multilayer perceptron classifier with point cloud as input –*
Takumi Tomokawa[1], Sakmongkon Chumkamon[1], Ayumu Tominaga[2], Sylvain Geiser[1], Ryusuke Fujisawa[1], Eiji Hayashi[1]
([1]Kyushu Institute of Technology, Japan)
([2]National Institute of Technology, Kitakyushu College, Japan)

OS19-5 *Particle Filter Based SLAM 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)

OS19-6 *Anomaly Detection using Autoencoder with Gramian Angular Summation Field in Time Series Data*
Umaporn Yokkampon[1], Sakmongkon Chumkamon[1], Abbe Mowshowitz[2], Eiji Hayashi[1]
([1]Kyushu Institute of Technology, Japan), ([2]The City College of New York, USA)

OS19-7 *Autonomous Robot Packaging Ready Meal in Conveyor Production Line*
Sakmongkon Chumkamon, Tomofumi Tsuji, Prem Gamolped, Chanapol Piyavichayanon, Umaporn Yokkampon, Eiji Hayashi (Kyushu Institute of Technology)

16:40-17:25 OS18 Natural Computing (3)
Chair: Marion Oswald (Technische Universität Wien, Austria)
Co-Chair: Yasuhiro Suzuki (Nagoya University, Japan)

OS18-1 *An Acoustic Artificial Life System Using the Game of Life and its Application for Performing Arts*
Yasuhiro Suzuki (Nagoya University Japan)

OS18-2 *The Effect of Non-audible Low Frequency, Deep Micro Vibrotactile, DMV Sounds on Music*
Yasuhiro Suzuki (Nagoya University, Japan)

OS18-3 *Response of Yeast to Low Frequency Sound Exposure*
Yasuhiro Suzuki (Nagoya University, Japan)

Meeting Room 3
9:00-10:15 OS15 Artificial Intelligence for Embedded Systems and Robotics (5)
Chair: Hakaru Tamukoh (Kyushu Institute of Technology, Japan)
Co-Chair: Takuya Nanami (University of Tokyo, Japan)

OS15-1 *INT8 Activation Ternary or Binary Weights Networks*
Ninnart Fuengfusin, Hakaru Tamukoh (Kyushu Institute of Technology, Japan)

OS15-2 *A parameter tuning method for PQN model*
Daimon Sakai, Takuya Nanami, Takashi Kohno (University of Tokyo, Japan)

OS15-3 *Hardware Development of Edge-Preserving Bubble Image Conversion in High-level Synthesis*
Jiang Qin, Akira Yamawaki (Kyushu Institute of Technology, Japan)

OS15-4 *Development of Haze Removing Hardware Using High-Level Synthesis*
Daiki Shirai and Akira Yamawaki (Kyushu Institute of Technology, Japan)

OS15-5 *Automatic approximation of primitive shapes using point clouds*
Yuma Yoshimoto, Hakaru Tamukoh (Kyushu Institute of Technology, Japan)

10:45-12:00 OS24 Artificial Systems and Life (5)
Chair: Kuo-Hsien Hsia (National Yunlin University of Science and Technology, Taiwan)
Co-Chair: Chung-Wen Hung (National Yunlin-University of Science and Technology, Taiwan)

OS24-1 *An EtherCAT Based Delta Robot Synchronous Control Application*
Chung-Wen Hung[1], Yu-Hsuan Tseng[1], Chau-Chung Song[2], Guan-Yu Jiang[1]
([1]National Yunlin-University of Science and Technology, Taiwan, [2]National Formosa University, Taiwan)

OS24-2 *Web-based SCADA using MQTT Protocol and AES*
Jr-Hung Guo, Tzu-Yuan Lin, Kuo-Hsien Hsia
(National Yunlin-University of Science and Technology, Taiwan)

OS24-3 *Smart Identification System of Teaching-type Autonomous Vehicles*
Chun-Chieh Wang (Chienkuo Technology University, Taiwan)

OS24-4 Automatic Anti-Lock Brake System for Anti-Rollover Control of Autonomous Heavy-Duty Truck
Chian C. Ho, Riki Umami Sanaz Ulfitria
(National Yunlin-University of Science and Technology, Taiwan)

OS24-5 *Development of Intelligent Beehive and Network Monitoring System for Bee Ecology*
Chau-Chung Song[1], Geng-Yi Lin[1], Chi-Chung Peng[2] and Chung-Wen Hung[3]
([1, 2]National Formosa University, Taiwan, [3]National Yunlin-University of Science and Technology, Taiwan)

14:20-16:00 OS26 Intelligent Life and Data Analysis (6)
Chair: I-Hsien Liu (National Cheng Kung University, Taiwan)
Co-Chair: Chu-Fen Li (National Formosa University, Taiwan)

OS26-1 *Data Balanced Algorithm Based on Generative Adversarial Network*
I-Hsien Liu, Cheng-En Hsieh, Wei-Min Lin, Jung-Shian Li (National Cheng Kung University, Taiwan), Chu-Fen Li (National Formosa University, Taiwan)

OS26-2 *Fault-Tolerant Control System Design for Nonlinear System with Actuator Faults*
Ho-Nien Shou (Air Force Institute of Technology, Taiwan)
Hsin-Yu Lai (National Cheng Kung University, Taiwan)

OS26-3 *Key Indicators for Successful E-Oriented Operation and Management of the Nutrition Consulting Service System*
Ling-Mei Hsu (Chang Jung Christian University, Taiwan)

OS26-4 *The key factors for the application of blockchain into ocean Freight Forwarders: An Industry Perspective*
Chu-Ting Hsu, Ming-Tao Chou, Ji-Feng Ding (Chang Jung Christian University, Taiwan)

OS26-5 *Key Success Factors for Implementation Quality Assurance of Information Technology in Tourism Industry*
Shuen-Huei Yao, (Chang Jung Christian University, Taiwan), Cheng Chung Yeh
(National University of Tainan, Taiwan), Wen Jung Tsai (Chang Jung Christian University, Taiwan)

OS26-6 *AI Big data analysis and application: Patient Safety Culture of Nursing Staff in an Operation Room*
Su-Chiu Hsiao (Chang Jung Christian University, Taiwan)

16:40-18:50 OS17 Advanced Robotics (8)
Chair: Evgeni Magid (Kazan Federal University, Russia)
Co-Chair: Kuo-Hsien Hsia (National Yunlin University of Science and Technology, Taiwan)

OS17-1 *Experience in efficient real office environment modelling in Gazebo: a tutorial*
Bulat Abbyasov[1], Kirill Kononov[1], Tatyana Tsoy[1], Martínez-García Edgar A.[2], Evgeni Magid[1]
([1]Kazan Federal University, Russia), ([2]The Autonomous University of Ciudad Juarez, Mexico)

OS17-2 *Graphical user interface design for a UAV teleoperation*
Roman Lavrenov[1], Ramil Safin[1], Bai Yang[2], Martínez-García Edgar A.[3], Roman Meshcheryakov[4]
([1]Kazan Federal University, Russia), ([2]Ritsumeikan University, Japan)
([3]The Autonomous University of Ciudad Juarez, Mexico)
([4]V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Russia)

OS17-3 *Numerical solution approach for the ROBOTIS OP2 humanoid hand inverse kinematics*
Zagidullin Linar[1], Tatyana Tsoy[1], Roman Meshcheryakov[2], Kuo-Hsien Hsia[3], Evgeni Magid[1]
([1]Kazan Federal University, Russia)
([2]V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Russia)
([3]National Yunlin University of Science and Technology, Taiwan)

OS17-4 *Alvus modelling in Gazebo*
Liaisan Safarova[1], Bulat Abbyasov[1], Tatyana Tsoy[1], Hongbing Li[2], Evgeni Magid[1]
([1]Kazan Federal University, Russia), ([2]Shanghai Jiao Tong University, China)

OS17-5 *Testing procedures architecture for establishing a fiducial marker recognition quality in UAV-based visual marker tracking task in Gazebo simulator*
Mikhail Kilin[1], Roman Lavrenov[1], Bai Yang[2], Mikhail Svinin[2], Evgeni Magid[1]
([1]Kazan Federal University, Russia), ([2]Ritsumeikan University, Japan)

OS17-6 *Feature importance evaluation method for multi-agent deep reinforcement learning in advanced robotics task allocation*
Sergey Ryabtsev[1], Mikhail Gurchinsky[1], Igor Struchkov[1], Vyacheslav Petrenko[1], Fariza Tebueva[1], Sergey Makarenko[2]
([1]North-Caucasus Federal University, Russia),
([2]Saint Petersburg Federal Research Center of the Russian Academy of Sciences, Russia)

OS17-7 *Iterative method of labor division for multi-robotic systems*
Sergey Ryabtsev[1], Artur Sakolchik[2], Vladimir Antonov[1], Vyacheslav Petrenko[1], Fariza Tebueva[1], Sergey Makarenko[3]
([1]North-Caucasus Federal University, Russia),
([2]Belarusian State University, Belarus)
([3]Saint Petersburg Federal Research Center of the Russian Academy of Sciences, Russia)

OS17-8 *Development of Bowling Machine Using VEX IQ*
Kuo-Hsien Hsia[1], Ya-Chun Chen[1], Evgeni Magid[2], Xin-Ying Zeng[1]
([1]National Yunlin University of Science and Technology, Taiwan)
([2]Federal Kazan University, Russia)

Meeting Room 4
9:00-10:30 GS1 Control System Applications (6)
Chair: Hiroaki Wagatsuma (Kyushu Institute of Technology, Japan)

GS1-1 *Design of local linear models using Self tuning Control System for PID Tuning According to error*
Shinichi Imai (Tokyo gakugei university, Japan)

GS1-2 *A Systematic Analysis of the Knee Support Exoskeleton Based on Multibody Dynamics Toward Personalization with 3D Printed Spring-Damper Components*
Shintaro Kasai[1], Pancho Dachikinov[1], Kohei Tanaka[1], Hiroaki Wagatsuma[1,2]
([1]Kyushu Institute of Technology, Japan; [2]RIKEN CBS, Japan)

GS1-3 *A Drone-Based Concrete Crack Inspection System by Using Morphological Component Analysis and Sub-Pixel Width Estimation*
Ankur Dixit[1], Wataru Oshiumi[1], Hiroaki Wagatsuma[1,2]
([1]Kyushu Institute of Technology, Japan; [2]RIKEN CBS, Japan)

GS1-4 *A Systematic Geometric Design Method of Flexible Bars Available for Personalized Knee Orthoses with Spring-Damper Functions*
Pancho Dachikinov[1], Shintaro Kasai[1], Kohei Tanaka[1], Hiroaki Wagatsuma[1,2]
([1]Kyushu Institute of Technology, Japan; [2]RIKEN CBS, Japan)

GS1-5 *State-space modeling of fingers motion measured by the Leap Motion Controller*
Ryuichi Usami, Hideyuki Tanaka (Hiroshima University, Japan)

GS1-6 *Applicability Verification of iWakka Game to Children with Developmental Coordination Disorder*
Masakazu Nomura[1], Moe Nishiya[1], Yoshifumi Morita[1], Hideo Yamagiwa[2]
([1]Nagoya Institute of Technology, Japan) ([2]Tokyo Metropolitan Tobu Medical Center, Japan)

10:45-12:00 OS27 Environmental Monitoring (5)
Chair: Kazuo Ishii (Kyushu Institute of Technology)
Co-Chair: Keisuke Watanabe (Tokai University)

OS27-1 *Biofouling Monitoring Experiments of Underwater Concrete Samples for Offshore Platform Cleaning Robot Development*
Keisuke Watanabe[1], Hiroki Goda[2], Koji Harada[3]
([1]Tokai University, Japan, [2]Kyusyu Institute of Technology, Japan, [3]Nishimatsu Construction Co., Ltd., Japan)

OS27-2 *Fall risk notification system using LiDAR sensor for the visually impaired people*
Daigo Katayama[1], Kazuo Ishii[1], Shinsuke Yasukawa[1], Satoshi Nakadomari[2], Koichi Wada[2], Akane Befu[2], Chikako Yamada[2] ([1]Kyushu Institute of Technology, Japan, [2]NEXT VISION, Japan)

OS27-3 *Reflection Coefficient Estimation through the Modelling of Ultrasonic Transmission*
Ryuugo Mochizuki, Yuya Nishida, Kazuo Ishii (Kyushu Institute of Technology, Japan)

OS27-4 *Evaluation of Maps Constructed by Crawler-type Agricultural Robot in Different Farms*
Takuya Fujinaga, Tsuneo Nakanishi (Fukuoka University, Japan)

OS27-5 *An Estimation Method of Coastal Ocean Debris Using Aerial Drone*
Kazuo Ishii[1], Kanako Shirahashi[1], Yuya Nishida[1], Moeko Tominaga[2], Yoshiki Tanaka[1], Dominic B. Solpico[1] ([1]Kyushu Institute of Technology, Japan, [2]Nishinippon Institute of Technology, Japan)

14:20-16:30 OS29 Advances in Marine Robotics and Their Applications (8)
Chair: Keisuke Watanabe (Tokai University)
Co-Chair Kazuo Ishii (Kyushu Institute of Technology)

OS29-1 *A Sensor Network to Estimate Fish Activity and Assist Feeding Decisions in Marine Aquaculture*
Dominic B. Solpico, Yuya Nishida, Kazuo Ishii (Kyushu Institute of Technology, Japan)

OS29-2 *Proposal of a Swarm Intelligent Underwater Glider System for a Long-term Three-dimensional Wide-area Ocean Observation*
Kanako Kobatake, Masakazu Arima (Osaka Prefecture University, Japan)

OS29-3 *Underwater Acoustic Communication using QPSK Modulation Method*
Yuya Nishida, Yuichiro Uemura, Rikuto Tanaka, Kazuo Ishii (Kyushu Institute of Technology, Japan)

OS29-4 *Ultrasonic Cleaner using Two Transducers for Ship Hull Cleaning Robot*
Yuya Nishida, Toshihiro Matsumura, Kazuo Ishii (Kyushu Institute of Technology, Japan)

OS29-5 *Motion Control of a Ship Hull Cleaning Robot*
Hyoga Yamamoto, Yuya Nishida, Takayuki Matsuo, Kazuo Ishii (Kyushu Institute of Technology, Japan)

OS29-6 *Development of a USV Testbed and Its System Check Experiments at Sea*
Keisuke Watanabe, Masatoshi Shimpo (Tokai University, Japan)

OS29-7 *Development of a Seabed Walking Platform for Ore Sample Drilling in Deep Sea Mining*
Keisuke Watanabe[1], Hideyuki Suzuki[2], Yoshiyasu Watanabe[1]
([1]Tokai University, Japan), ([2]The University of Tokyo, Japan)

OS29-8 *Development of Remotely Operated Vehicle for Small-size Jellyfish Extermination and its Evaluation of Extermination Motion Control*
Hiroyuki Yokota[1], Shinsuke Yasukawa[2], Jonghyun Ahn[1]
([1]Hiroshima Institute of Technology, Japan, [2]Kyushu Institute of Technology, Japan)

January 23 (Sunday)

8:40-Registration

Meeting Room 1

9:00-10:30 OS3 Intelligent Systems and Robotics -1- & OS8 Intelligent Systems and Robotics -3-

OS3 Intelligent Systems and Robotics -1- (6)
Chair: Peng Lu (Tianjin University of Science and Technology, China)
Co-Chair: Fengzhi Dai (Tianjin University of Science and Technology, China)

OS3-1 *A Distributed Optimal Formation Control for Multi-Agent System Based on UAVs*
Jichao Zhao[1], Fengzhi Dai[1], Yunzhong Song[2]
([1]Tianjin University of Science and Technology, China), ([2]Henan Polytechnic University, China)

OS3-2 *A Design of Multi-Agent System Simulation Platform Based on Unmanned Ground Vehicles and A Research on Formation Control Protocol*
Chuang Zhang, Jichao Zhao, Fengzhi Dai (Tianjin University of Science and Technology, China)

OS3-3 *A Study of Weighted Average Method for Multi-sensor Data Fusion*
Peng Lu[1], Fengzhi Dai[1,2]
([1] Tianjin University of Science and Technology, China),
([2] Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

OS3-4 *A Design of Wide-angle Open and Close Multifunctional Smart Windows*
Zilong Liu, Fengzhi Dai*, Luqi Shen, Jiarun Xu, Longyu Gao
(Tianjin University of Science and Technology, China)

OS3-5 *A Low-intensity Laser Control System Design*
Yuhui Cheng, Fengzhi Dai (Tianjin University of Science and Technology, China)

OS3-6 *A Review of Machine Vision Based Fruit Recognition Applications*
Tianyi Zhang, Fengzhi Dai (Tianjin University of Science and Technology, China)

OS8 Intelligent Systems and Robotics -3- (4)
Chair: Tianyi Zhang (Tianjin University of Science and Technology, China)
Co-Chair: Yiting Gao (Tianjin University of Science and Technology, China)

OS8-1 *A Portable Electrocardiograph System Design based on STM32 Chip*
Tianyi Zhang, Huating Liu (Tianjin University of Science and Technology, China)

OS8-2 *Design and Development of The Parking Space Autonomous Management System*
Yiting Gao, Tianyi Zhang (Tianjin University of Science and Technology, China)

OS8-3 *An Overview of Obstacle Avoidance Methods for Unmanned Vehicles*
Peng Lu, Haokang Wen (Tianjin University of Science and Technology, China)

OS8-4 *Synchronization of Novel 5D Hyperchaotic Systems*
Hong Niu (Tianjin University of Science & Technology, China)

Meeting Room 2
10:50-11:50 OS7 Media Information Processing, Music Recommendation and Artificial Intelligence (4)
Chair: Yasunari Yoshitomi (Kyoto Prefectural University, Japan)
Co-Chair: Masayoshi Tabuse (Kyoto Prefectural University, Japan)

OS7-1 *Data expansion method by combining unnecessary sentence deletion and most important sentence addition*
Tomohito Ouchi, Masayoshi Tabuse (Kyoto Prefectural University, Japan)

OS7-2 *Evaluation of a system that the reading of sentences by a voice synthesizer and the highlighting sentences*
Sota Kobayashi, Masayoshi Tabuse (Kyoto Prefectural University, Japan)

OS7-3 *Mouse Cursor Control System using Eye Gaze and detection of eye opening and closing*
Masayoshi Tabuse, Mana Fukumoto, Yasunari Yoshitomi, Taro Asada
(Kyoto Prefectural University, Japan)

OS7-4 *Music Recommendation System Driven by Facial Expression Recognition*
Taro Asada[1], Motoki Kawamura[2], Yasunari Yoshitomi[1], Masayoshi Tabuse[1]
([1]Kyoto Prefectural University, Japan, [2]Fujitsu Limited, Japan)

13:00-14:40 OS20 Application studies on sensor and RFID system (6)
Chair: Ammar A.M. Al-Talib (UCSI university, Malaysia)
Co-Chair: Takao Ito (Hiroshima University, Japan)

OS20-1 *A Pedal Powered Water Purifier*
Ammar A.M. Al-Talib, Ting Kee Yuan, Sarah 'Atifah Saruchi (UCSI University, Malaysia)

OS20-2 *A Smart Node (Maintenance & Lifespan Prediction System)*
Kam Heng Chaw[1], Ammar A.M. Al-Talib[2], Tarek Fawzi[2], Jonathan Yong Chung Ee[2]
([1]MODU System (S) Pte Ltd, Malaysia), ([2]UCSI University, Malaysia)

OS20-3 *A Healthcare Laundry Management System using RFID System*
Eryana Hussin, Wong Jie Jian (UCSI University, Malaysia)

OS20-4 *A Monitoring System with Humidity and Growth Level Detection for Horticulture*
Eryana Hussin, Ng Joon Wen (UCSI University, Malaysia)

OS20-5 *A Double Identification Attendance System Using High Frequency RFID System*
Eryana Hussin, Wong Chee Ming (UCSI University, Malaysia)

OS20-6 *Levitating Frictionless-Vertical Windmill-*
Ammar A.M. Al-Talib, Muhammad R. Md. Redzuan, A. R. Abd Hamid (UCSI University, Malaysia)

Meeting Room 3
9:00-11:50 OS9 Robot Control (10)
Chair: Yizhun Peng (Tianjin University of Science and Technology, China)
Co-Chair: Junhui Yin (Tianjin University of Science and Technology, China)

OS9-1 *Design of Analog Electromagnetic Gun Based on Arduino*
Lintao Hu, Yizhun Peng (Tianjin University of Science and Technology, China)

OS9-2 *A Design of Intelligent House Inspection Robot*
Sicong Wang, Yizhun Peng (Tianjin University of Science and Technology, China)

OS9-3 *A Design of Embedded Plate & Ball Control System Based on Machine Vision*
Wendinig Luo[1], FuChen Zhao[2]
([1]Beijing Huadian Jingneng Energy Technology Co., China)
([2]Beijing University of Posts and Telecommunications, China)

OS9-4 *Design of Blood Circulation System of Medical Simulation Robot*
Songyun Shi, Yang Ge, Chomgxu Guo, Yizhun Peng
(Tianjin University of Science and Technology, China)

OS9-5 *A Design of Dynamic Exoskeleton for Self-learning Human Movements*
Qingliang Liu, Yucheng, Pengyu Yao, Dechao Wang, Yizhun Peng
(Tianjin University of Science and Technology, China)

OS9-6 *Design of Autonomous navigation building Climbing and Handling Robot Based on SLAM*
Linhui Chen, Junjie Tai, Yizhun Peng (Tianjin University of Science and Technology, China)

OS9-7 *Simulation Research on Automatic Navigation of Indoor Wheelchair*
Peng Shi, Yizhun Peng (Tianjin University of Science and Technology, China)

OS9-8 *Design of a Humanoid Dance Robot for Dancing Baduanjin*
Zhuofan Xu, Ruitao Li, Binfu Zhong, Yizhun Peng
(Tianjin University of Science and Technology, China)

OS9-9 *Design of Intelligent Shading System Suitable for Parenting Products*
Yihan Yan, Songyun Shi, Yaxin Shen, Yizhun Peng
(Tianjin University of Science and Technology, China)

OS9-10 *Design of Intelligent Personalized Nutrition Supplement Machine*
Xinpeng Yang, Sidan Liu, Rui Li, Junhui Yin, Yizhun Peng
(Tianjin University of Science and Technology, China)

13:00-15:25 OS32 Human-Machine Interface and Automation (9)
Chair: Norrima Mokhtar (University of Malaya, Malaysia)
Co-Chair: Fakhrul Hazman Yusoff (UITM, Malaysia)

OS32-1 *A Derivative Oriented Thresholding Approach for Feature Extraction of Mold Defects on Fine Arts Painting*
Hilman Nordin, Bushroa Abdul Razak, Norrima Mokhtar, Mohd Fadzil Jamaludin
(Universiti Malaya, Malaysia)

OS32-2 *Imaginary Finger Control Detection Algorithm Using Deep Learning with Brain Computer Interface (BCI)*
Suresh Gobee[1], Norrima Mokhtar[1], Hamzah Arof[1], Noraisyah Md Shah[1], Wan Khairunizam[2]
([1]Universiti Malaya, Malaysia), ([2]University Malaysia Perlis, Malaysia)

OS32-3 *Investigating the Effect of Individuality Factors in Measuring Aggresion Induced by Human Brain*
Wan Khairunizam[1], Kai Xu tung[1], Lugieswaran[1], Wan Azani Mustafa[1], Hashimah Ali[1],
Zuradzman M. Razlan[2], Shahriman AB[2], Norrima Mokhtar[3]

([1]University Malaysia Perlis, Malaysia), ([2]Universiti Malaya, Malaysia)

OS32-4 *Voice User Interface(VuI) Smart Office Door Application in the Context of Covid-19 Pandemic*
Muhammad Zharif Aiman Alias, Wan Norsyafizan W. Muhamad, Suzi Seroja Sarnin, Darmawaty Mohd Ali (University Technology MARA, Malaysia)

OS32-5 *Endometrial Cell Images Segmentation: A Comparative Study*
Wan AzaniMustafa[1], Nurul Umaira Salim[1], Wan Khairunizam[1], Shahrina Ismail[1], Hiam Alquran[2]
([1]Universiti Malaysia Perlis, Malaysia; [2]Yarmouk University, Jordan)

OS32-6 *Temperature Control Using Fuzzy Controller for Variable Speed Vapor Compression Refrigerator System*
Siti Qurrata Ain Suhaimi, M. Saifizi, S.M. Othman, Azri A. Aziz, Wan Azani Mustafa,
Wan Khairunizam (Universiti Malaysia Perlis, Malaysia)

OS32-7 *Automatic Dry Waste Classification for Recycling Purpose*
Muhammad Nuzul Naim Baharuddin[1], Hasan Mehmood Khan[1], Norrima Mokhtar[1], Heshalini Rajagopal[2], Tarmizi Adam[3], Wan Amirul Wan Mahiyiddin[1],
Jafferi Jamaluddin[4]
([1]Universiti Malaya, Malaysia). ([2]Manipal International University, Malaysia),
([3]Universiti Teknologi Malaysia, Malaysia), ([4]UMPEDAC Universiti Malaya, Malaysia)

OS32-8 *A Low Cost Smart Parcel Box System with Enhanced Security*
Ahmad Luqmanulhakim, Wan Norsyafizan W. Muhamad, Suzi Seroja Sarnin, Meor Mohd Azreen Meor Hamzah (Universiti Teknologi MARA, Malaysia)

OS32-9 *Classification of Body Mass Index Based Face Images Using Facial Landmarks Approach and PCA plus LDA*
Hasimah Ali[1], Ho Yong Kang[1], Wan Khairunizam Wan Ahmad[1], Mohamed Elshaikh[1],
Norrima Mokhtar[2]
([1]University Malaysia Perlis, Malaysia), ([2] University of Malaya, Malaysia)

Meeting Room 4
9:00-11:35 OS31 Approaches to Post-Narratology that Combines AI and Cognitive Science with Narratology (9)
Chair: Jumpei Ono (Aomori University, Japan)
Co-Chair: Hiroki Fxyma (Tainan University of Technology, Taiwan)
Co-Chair: Takashi Ogata (Iwate Prefectural University, Japan)

OS31-1 *Story Units of the Types of Japanese Folktales and the Combination with a Noun Conceptual Dictionary*
[1]Jumpei Ono, [2]Motoki Kumagai, [2]Takashi Ogata
([1]Aomori University, Japan), ([2]Iwate Prefectural University, Japan)

OS31-2 *Visualization of the Unconscious in Quality Inspection in Manufacturing*
Jun Nakamura (Chuo University, Japan)

OS31-3 *The Study on the Relationship Between the Comic Artists' Styles and the Visual Languages: From the Stylistic Changes in the Work of Japanese Comic Artists*
Kaori Otsuru (Tainan University of Technology, Taiwan (ROC))

OS31-4 *Theoretical Backgrounds toward Text Mining for a Phenomenological Model of Taste Perception*
Hiroki Fxyma (Tainan University of Technology, Taiwan)

OS31-5 *Why is the Early Detection of Dementia Failed?*
Yuki Hayashi (Chiba University/National Institute for Japanese Language and Linguistics, Japan)

OS31-6 *Relationship Between World-view and Advertising Techniques*
Yoji Kawamura (Kindai University, Japan)

OS31-7 *Adjective and Adjective Verb Conceptual Dictionaries in an Integrated Narrative Generation System*
Jumpei Ono (Aomori University, Japan), Takashi Ogata (Iwate Prefectural University, Japan)

OS31-8 *Prototyping Animation System that Combines a Kabuki Work and its Background Story:* Kyōganoko Musume Dōjōji *and the Legend of* Dōjōji
[1]Miku Kawai, [2]Shunta Kudo, [3]Jumpei Ono, [2]Takashi Ogata
([1]The Open University of Japan, Japan), ([2]Iwate Prefectural University, Japan)
([3]Aomori University, Japan)

OS31-9 *How Will Art Appreciations Change According to Information Change?*
Akinori Abe (Chiba University, Japan)

13:00-15:10 OS12 Intelligent Measurements and Control Systems (8)

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

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

OS12-1 *Research on an AGV path planning method*
Zhihui Chen, Xiaoyan Chen, Shiming Wang (Tianjin University of Science and Technology, China)

OS12-2 *Boiler level measurement and control system*
Xiwen Liang, Shengli Sun, Xiaoyan Chen (Tianjin University of Science and Technology, China)

OS12-3 *Recurrence quantification and time-frequency analysis of two-phase flow patterns*
Keying Ren, Xiaoyan Chen , Meng Du (Tianjin University of Science and Technology, China)

OS12-4 *Application of deep learning in automatic driving*
Wei Su, Xiaoyan Chen (Tianjin University of Science and Technology, China)

OS12-5 *An intelligent home security system based on STM32*
Wei Su[1], Xiaoyan Chen[1], Guangyong Xi[2]
([1]Tianjin University of Science and Technology, China)
([2]Zhengzhou University of Light Industry, China)

OS12-6 *Design of smart bracelet based on STM32 microcontroller*
Xia Miao, Xiaoyan Chen, Jianliang Li (Tianjin University of Science and Technology, China)

OS12-7 *Design of material conveying and automatic sorting control system based on PLC*
Qian Wang[1], XiaoYan Chen[1], Shengmin Cao[2]
([1]Tianjin University of Science and Technology, China), ([2]TangShan College, China)

OS12-8 *Image reconstruction based on ResV-Net for electrical impedance tomography*
Qian Wang, ZiChen Wang, Di Wang, XiaoYan Chen
(Tianjin University of Science and Technology, China)

Group Meeting for the next Conference

N-1 *Compact Ultra-Wideband Slotted Microstrip Patch Antenna for 5G, IoT and RFID Applications*
Gurney Nga, Mastaneh Moakyef, Shahid Manzoor, Manickam Ramasamy (UCSI University, Malaysia)

N-2 *Drone Performance Analysis Based on SNR Factor*
Gershom Phiri, Mastaneh Moakyef, Sew Sun Tiang, Wong Chin Hong (UCSI University, Malaysia)

Abstracts

PS abstracts (3)

PS1 Robotics for Growing Life

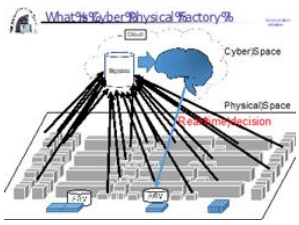
Henrik Hautop Lund (Technical University of Denmark, Denmark)

We present a novel direction of artificial life robotics in which we use robotics to control the growth of real, natural life. The concept of using robotics to grow life present itself as a potential sustainable solution for food production, allowing an optimization of food quality and outcome. We illustrate this concept with our development of the Growbot, which is a tabletop size robotic green house for growing edible food plant. The GrowBots use sensors such as humidity, CO2, temperature, water level and camera sensors, and actuators such as full spectrum LEDs, IR LEDs, UV LEDs, fertilizer and water pumps, air change and air fan. The software acts as recipes for the plant growth in the robotic greenhouse adjusting the environmental condition for the growth of the living plants such as salad, parsley and basil.

PS2 Road To Cyber Physical Factory(Application Examples of Intelligent Factory and its Technology)

Hidehiko Yamamoto (Gifu University, Japan)

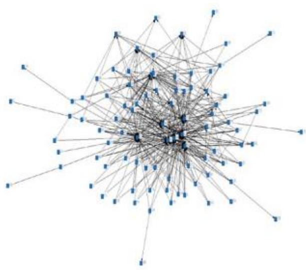
Since 1980, Computerized machine tools and robots have been developed. I talk about the history of the Intelligent manufacturing systems and their technology including Artificial Intelligence and GA. The technology are as follows. The future factory by using Autonomous System, we call Autonomous Decentralized Flexible Manufacturing Systems, is presented. The virtual factory and several kinds of simulations for production systems and scheduling problems is presented. The simulations application examples for automobile production lines are presented. IoT production and Cyber Physical Factory which is the near future manufacturing model are presented. Its application example for automobile parts production is also presented.



PS3 Robot Technology, and it’s Development Trend–Developing a New Networking Robot System-

Takao Ito (Hiroshima University, Japan)

Robot technology has been changed dramatically with massive development of internet environment. The author reviewed a plethora of literature and investigated advanced robot technologies. Today, most of typical robot technologies are used in single-cause-oriented products, such as robot vacuum cleaner and Asimo, a humanoidrobot invented by Honda. These advanced products played important role in our modern society. For further development, a networking robot system with advanced technologies of internet and artificial intelligence is required in order to copy with theuncertainty in the future. Different technology should be combined and linked together for multiple-goal-oriented approach in the networking robot system. For evaluating the validity of our new system, a centrality index is introduced in this research.



IS abstracts (3)

IS1-1 A physical to virtual control system implementing an art-based game.

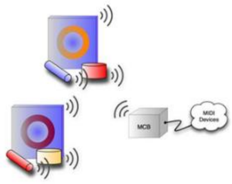
Luigi Pagliarini[1,2], Henrik Hautop Lund[1]
([1]Technical University of Denmark, Denmark) ([2] Academy of Fine Arts of Macerata, Italy)

We hereby present Tiler, an art-based game where a virtual world made of tiles and controlled through a set of electronic cubes where players can/should gather the decoration of a floor based on aesthetical criteria. Such a tool allows projecting and designing bi-dimensional shapes by physically manipulating tri-dimensional objects. Besides that, Tiler, is an application that brings to life a clear example on how we can achieve new ways of interfacing the physical world with virtual ones. In the following article we introduce the logical and technical aspects of this real-to-virtual interface and show its potential applications in different fields

IS1-2 Creative Multisensory Environments

Henrik Hautop Lund[1], Luigi Pagliarini[1,2]
([1]Technical University of Denmark, Denmark), (Academy of Fine Arts of Macerata, Italy)

We outline the concept of playful creative multisensory environments. Multisensory environments are characterized by facilitating multimodal interaction by users through a composition of different objects. These objects allow manipulation of the environment. Multisensory rooms are often used to enhance users' sensations and emotions. In order to explore the development of such multisensory environments, which are playful and provide possibilities to become creative, we suggest the use of modular playware as part of the Playware ABC concept. This allows for the development of flexible, portable environments to be used by anybody, anywhere, anytime, and facilitates that the users can construct, combine and create. In this way, the users can become creative in seamless interaction with aesthetically pleasing environments. The paper outlines the lessons learned from the development and testing of a number of such playful creative multisensory rooms. Main findings are that such multisensory environments need to provide explicit immediate feedback, be simple in design, build on high quality aesthetics, provide variation, and be playful to result in intrinsic motivation.



IS2-1 Using GrowBots to Study Heat and Nutrient Stress in Basil

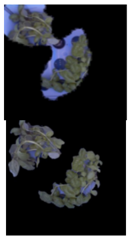
Rishav Bose, Henrik Hautop Lund (Technical University of Denmark, Denmark)

We present how a novel type of robot called a food computer can be used to simulate abiotic stresses and study their impact on hydroponically grown Italian basil, Ocimum basilicum. The food computer called the GrowBot is a tabletop sized robotic greenhouse for growing edible food plants. The GrowBot's actuators were used to alter the environmental conditions in the growth chamber to study different aspects of plant growth and food production in varying climate scenarios. The experiments show how we can used the LED lights to control the temperature to a certain, desired range (29ºC – 35ºC) for the heat stress experiments, while measurements show that we can simultaneously obtain the photosythetically active radition (PAR) values to be in the ideal range for growth of the basil plants.

IS2-2 Convolutional Neural Network for Studying Plant Nutrient Deficiencies

Rishav Bose, Henrik Hautop Lund (Technical University of Denmark, Denmark)

We discuss the development of a vision-based plant phenotyping system based on a novel type of robotic system called a food computer. The food computer used in this project is called the GrowBot. It has a host of sensors to help analyse the growth chamber including a Raspberry Pi camera. The project revolved around developing a system to segment the plant canopy from its background and analyse nutrient deficiencies from the images taken by the camera. The pilot project investigated how a segmentation model called U-Net could be used to study the images. One of the drawbacks of many existing vision-based plant phenotyping systems is that their convolutional neural networks (CNNs) were trained to analyse very ideal images of individual leaves. This pilot project tried to address that issue, while at the same time explored how to train the neural networks to learn segmentation from a small image dataset.



OS abstracts

OS1 Advanced Information Processing Applications (4)

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

OS1-1 Generation of Checkered Pattern Images Using Prewitt Filter from RGB-D Images

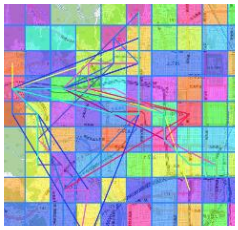
Toru Hiraoka, Ryosuke Takaki (University of Nagasaki, Japan)

A non-photorealistic rendering method for generating checkered pattern images from gray-scale photographic images using Prewitt filter with an expanded window size has been proposed. In this paper, we propose an extension of the conventional method to apply to RGB-D images. Our method can change the size of the checkered patterns depending on the depth. To verify the effectiveness of our method, we conducted experiments that are visually confirmed the checkered patterns by changing the parameters in our method.

OS1-2 An Approach of Analyzing Movement Patterns Using Word Embeddings from Geo-tagged Tweets

Masahru Hirota, Tetsuya Oda (Okayama University of Science, Japan)

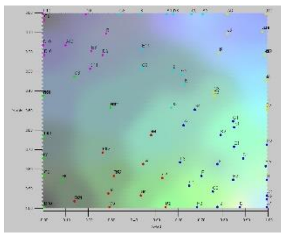
Many tourists upload geo-tagged content about tourist attractions to social media sites. The trajectories composed of those contents contribute to many applications, such as analyzing tourist behavior patterns, traffic flow analysis, and next location prediction. In this paper, we analyze the movement patterns of users by clustering their movement trajectories in Twitter data. We use the word embedding model to learn movements between two places. Also, we cluster the trajectories based on the similarity calculated by the generated embedding vectors. We visualized the clustered trajectories on a map and discussed the result.



OS1-3 A Proposal of a Software Defect Predication System Using SOM

Yoshihiro Kita, Kazuki Ueda, Kiyotaka Sakurai (University of Nagasaki, Nihon Knowledge Co. Ltd, Japan)

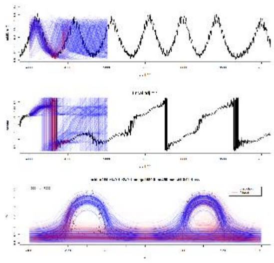
The goal of software testing is to detect all potential defects. However, it is difficult to know how many latent defects remain and where they are hidden. In this research, we propose a system that analyzes the characteristics and tendencies of already detected defects and predicts where the potential defects are likely to be. Specifically, the system inputs the data of detected defects into a Self-Organization Map (SOM) and predicts the locations that contain many defects from this map. In order to confirm the validity of this proposal, we input past defect data into the SOM, analyze the trend of defects, and evaluate the predictability of the potential defects.



OS1-4 Mapping the Motion of Highly-inclined Triple System into a Secular Perturbation Model

Masaya M. Saito (University of Nagasaki, Japan)
Kiyotaka Tanikawa (National Astronomical Observatory, Japan)

The motions of three bodies like Sun-Asteroid-Jupiter system or triple star system are formalized as hierarchical three body problem. When the third body orbits around the rest in a highly inclined elliptic orbit, the system undergoes the oscillation, called the Kozai oscillation, where the eccentricity may increase with decrease of the inclination of the orbital plane. While the Kozai oscillation seems to be a key process in orbital evolution, including disruption of triple system, its reflection into actual trajectories is quite complicated to analyze. For this reason, we try to map these trajectories into a secular perturbation model with data assimilation and demonstrate the extraction of state and its transition (libration to circulation and vice versa) as the Kozai oscillation.



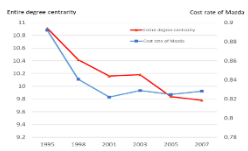
OS2 Advanced studies of interdisciplinary approach (4)

Chair Yousin Park (Prefectural University of Hiroshima, Japan))
Co-Chair Takao Ito (Hiroshima University, Japan)

OS2-1 Measuring the entire degree centrality in Yokokai

Tsutomu Ito[1], Matsuno Seigo[1], Sakamoto Makoto[2], Takao Ito[3]
([1]Ube National College of Technology, Japan), ([2]Univeristy of Miyazaki, Japan),
([3]Hiroshima University, Japan)

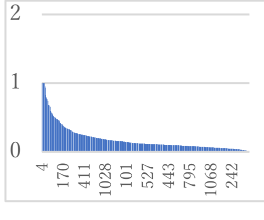
Centrality is one of the most important indexes in network calculation. Based on the definition, more than 400 different centrality such as degree, betweenness have been developed. All centrality indexes are calculated using the number of connection line, and its position in a given network. In automotive industry, keiretsu is considered as one of typical networks. It is critical to measure the centrality of transaction network in the keiretsu. It is widely well-known that different parts play different roles in a car. Thus, the importance of each connect line in a transaction network should be measured based upon the importance of the parts. A new parts-importance weighted centrality model is proposed in this paper.



OS2-2 A comparative study on Michinoeki's efficiency in Japan

Tsutomu Ito[1], Matsuno Seigo[1], Sakamoto Makoto[2], Takao Ito[3]
([1]Ube National College of Technology, Japan), ([2]Univeristy of Miyazaki, Japan),
([3]Hiroshima University, Japan)

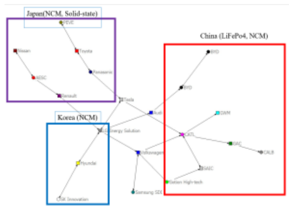
Michinoeki is considered as one of the most successful social experiments in Japan now. A plethora of literature of current situation analysis of Michinoeki's have been published over past decades. The authors reviewed many important and typical literatures and found that the quantitative studies are still sparse. Obviously, it is a critical issue to measure the efficiency of Michinoeki's for its further development and revitalization of the local economy. Thus, the data of Michinoeki are gathered from all Michinoeki in Japan, and the relative efficiency are calculated using DEA model for comparison. Furthermore, the managerial implication of the results is discussed in this paper.



OS2-3 The R&D Direction and Business Strategy: The case study on the cooperation of EV and battery makers

Yousin Park[1], Iori Nakaoka[2], Yun-ju Chen[3]
([1]Prefectural University of Hiroshima, Japan), ([2]Seijoh University, Japan), ([3]Shiga University, Japan)

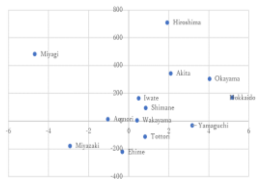
This paper focuses on the R&D direction and the business strategy of EV (Electric Vehicle) and battery makers. M. E. Porter (1996) claimed that the productivity frontier represents the maximum value that the organization can deliver at any a given cost, using technologies, skills and purchased inputs. He argued that strategic decisions are ones that are aimed at differentiating an organization from its competitors in a way that is sustainable in the future. We use the patent information of EV companies (Toyota, Tesla, VW) and battery makers (Panasonic, CATL, LG Chem) as the cases. And we examine our propositions by social network analysis and text mining. The analysis in this paper include: 1) To try to distinguish between differentiation and low-cost strategy, and visualize each firm's R&D direction, 2) To make discussion on business strategy and the R&D direction of EV companies and battery makers. In this paper we suppose that patterns of cooperation EV and battery makers.



OS2-4 Discovering the relationship between tourists and tourist spots in Japan

Tsutomu Ito[1], Matsuno Seigo[1], Sakamoto Makoto[2], Takao Ito[3]
([1]Ube National College of Technology, Japan), ([2]Univeristy of Miyazaki, Japan),
([3]Hiroshima University, Japan)

It is widely well-known that Japan currently has serious problems with the declining birthrate and aging society. As actions to counteract our aging and shrinking population, revitalization of local economy is becoming the most effective economic development strategy. Obviously, tourism is considered as one of the effective policies to develop local economy today. To provide more practical evidence of tourism development, this paper is focusing on discovering the relationship between tourists and tourist spots using conventional regression model and 4-cell model developed on tourists and tourist spots. Moreover, as one of the successful factors, the importance of the development of public traffic among different tourist spots has been confirmed.

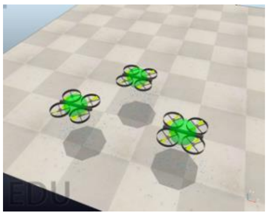


OS3 Intelligent Systems and Robotics-1 (6)
Chair Peng Lu (Tianjin University of Science and Technology, China)
Co-Chair Fengzhi Dai (Tianjin University of Science and Technology, China)

OS3-1 A Distributed Optimal Formation Control for Multi-Agent System Based on UAVs

Jichao Zhao[1], Fengzhi Dai[1], Yunzhong Song[2]
([1]Tianjin University of Science and Technology, China), ([2]Henan Polytechnic University, China)

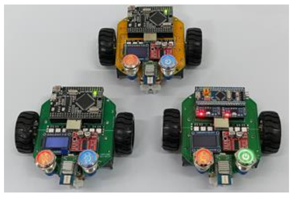
In this paper, the distributed optimization problem of multi-agent system (MAS) formation control composed of unmanned aerial vehicles (UAVs) is solved. Aiming at the situation that the internal state of a single UAV can be fully understood, the internal optimal control law of a UAV is designed by using the optimal control theory. To solve the problem that each agent in the system can only communicate with some agents, the distributed formation control law of the system is designed based on the communication topology of the system, and the stability of the system is analyzed by the knowledge of graph theory. The validity of the formation protocol is verified by numerical simulation and UAV platform.



OS3-2 A Design of Multi-Agent System Simulation Platform Based on Unmanned Ground Vehicles and A Research on Formation Control Protocol

Chuang Zhang, Jichao Zhao, Fengzhi Dai (Tianjin University of Science and Technology, China)

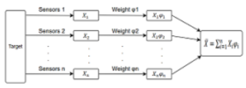
Inspired by ants and fishes in nature, the multi-agent system (MAS) plays a huge role in production and manufacturing in modern society. This paper takes unmanned ground vehicles (UGVs) as the object and designs a multi-vehicle test platform, which is composed of UGVs, ultra-wideband (UWB), and Bluetooth Mesh. At the same time, the UGV can achieve different control objectives by changing the main controller. The formation control protocol of UGVs is designed and its stability is analyzed. The effectiveness of the protocol is verified by numerical simulation. Finally, the designed control protocol is applied to the developed hardware test platform to verify the effectiveness of the experimental platform.



OS3-3 A Study of Weighted Average Method for Multi-sensor Data Fusion

Peng Lu[1], Fengzhi Dai[1,2]
([1] Tianjin University of Science and Technology, China),
([2] Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

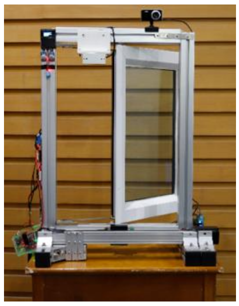
With the development of sensor technology, multi-sensor data fusion has become an important research direction in the field of sensors. And parametric classification algorithms have become intensive in the field of multi-sensor data fusion. The weighted average method is the most important one among the parametric classification algorithms. This paper describes the composition and development of parameter classification algorithms, focusing on the process, steps and recent developments of the weighted average method, and uses the algorithm to fuse data from ultrasonic and infrared sensors. The simulation results prove that the weighted average method has a better fusion effect.



OS3-4 A Design of Wide-angle Open and Close Multifunctional Smart Windows

Zilong Liu, Fengzhi Dai*, Luqi Shen, Jiarun Xu, Longyu Gao
(Tianjin University of Science and Technology, China)

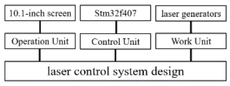
Current smart windows always have low intelligence and a few functions. And almost all kinds of traditional windows have the problem of inconvenient cleaning outside. The paper proposed a new mechanical structure design of the window, so that the window can be turned by nearly 180 degrees, which is convenient for cleaning the inside and outside of the window. In terms of intelligent systems, a variety of sensors combined with Internet modules are used to realize the remote intelligent control (turning with the wind, and automatically closing windows when there is no one at home when bad weather). The added child mode uses visual inspection to identify children. When a child is detected, the window will be automatically closed to prevent the child from climbing the window and falling from the building.



OS3-5 A Low-intensity Laser Control System Design

Yuhui Cheng, Fengzhi Dai (Tianjin University of Science and Technology, China)

Laser-assisted therapy has a good therapeutic effect on specific symptoms in clinical and experimental. This paper proposes a low-intensity laser control system design scheme, which is divided into the operation unit, control unit and work unit. The main control unit is based on the STM32F407ZGT6 chip design, the operating unit uses a DWIN 10.1-inch serial screen, and the working unit consists of 8 low-intensity laser generators. The system can control the power, frequency and working time of 8 laser generators at the same time, and has the function of automatically saving working parameters when the power is turned off.



OS3-6 A Review of Machine Vision Based Fruit Recognition Applications

Tianyi Zhang, Fengzhi Dai (Tianjin University of Science and Technology, China)

Machine vision is widely used in various fields, The main applications in the field of fruit picking are fruit identification, fruit quality detection, fruit ripeness detection and grading, etc. And fruit ripeness detection technology is important to improve the quality and market competitiveness of fresh and stored fruits. This paper focuses on the application of machine vision in fruit identification, fruit ripeness detection and grading in the past three years, and the application is more mature in many fruits such as citrus, blueberry, cherry, etc. It uses a number of algorithms to enable accurate identification of fruits and processing of their images to control the robotic arm for a variety of operations such as picking.

OS4 Intelligent Systems and Robotics -2-(5)
Chair Haokang Wen (Tianjin University of Science and Technology, China)
Co-Chair Jiaxin Li (Tianjin University of Science and Technology, China)

OS4-1 A Research on Image Dehazing Technology for Image Enhancement
Haokang Wen, Chang Sheng (Tianjin University of Science and Technology, China,)

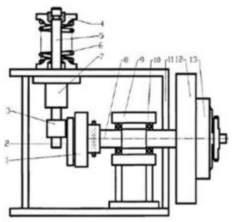
Image defogging is to study the method of image enhancement in foggy weather with low definition and lighter color. Image defogging technology aims to improve image contrast and scene clarity, and has broad application prospects in the fields of target recognition, traffic navigation, and remote sensing. Three defogging algorithms in image defogging technology are introduced based on the current research status: global histogram equalization, local histogram equalization, and Retinex algorithm. This article introduces the main steps of the three algorithms and discusses the advantages and disadvantages of each algorithm. Finally, the processing results of these three algorithms are compared and comprehensively evaluated.

OS4-2 Design of Intelligent Daylily Picking Robot
Jiaxin Li (Tianjin University of Science and Technology, China)

The picking conditions of daylily are bad, and long-term picking is likely to cause extremely serious harm to human body, especially to hands. Through the research and summary of the biological characteristics of daylilies, this paper aims to develop a kind of intelligent picking mechanism with strong applicability. The main work includes: designing binocular recognition system for rapid segmentation and recognition of daylily; Workflow design of daylily intelligent picking robot; Software and hardware design of daylily picking control system.

OS4-3 50KN Compression Spring Fatigue Testing Machine Design
Peng Lu, Peng Jia (Tianjin University of Science and Technology, China)

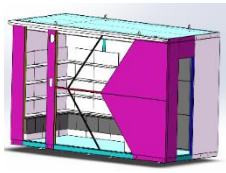
Spring fatigue test is the key process of spring performance testing, especially for automotive suspension springs, train damping springs, engine valves and other critical parts of the spring, must do the reliability assessment of spring fatigue performance. Because the requirements of different springs are different, the frequency and amplitude used in the test also have different requirements. In this case, a compression spring fatigue tester was developed and designed to test the maximum number of cycles of a spring under a given failure condition by applying a cyclic variable load to the spring. Through the design of the compression spring fatigue testing machine, the test prototype is finally developed.



OS4-4 New Intelligent Unmanned Retail Shopping Container Design

Peng Lu, Yiting Gao (Tianjin University of Science and Technology, Tianjin, China)

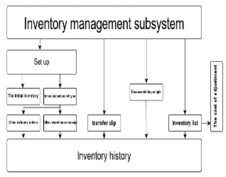
During the COVID-19 epidemic, unmanned retail has seen new opportunities for growth. The common unmanned retail methods are vending machines and unmanned supermarkets, which have the disadvantages of low product variety and high store costs, respectively. In this paper, a new intelligent unmanned retail shopping container is designed, integrating unattended technology with the container. Through the two-dimensional code, electronic anti-theft, RFID and other technologies, designed to achieve the shopping process of the goods that are taken away, the whole data. At the same time, the designed shopping containers can be deployed closer to consumers, providing a safer and more convenient shopping experience for people.



OS4-5 Java-based Dream Cloud ERP System - Inventory Management Subsystem Design and Implementation

Yiting Gao, Peng Lu (Tianjin University of Science and Technology, China)

With the continuous development of economy, large and medium-sized enterprises of heavy production tasks, more and more high to the requirement of inventory management. The traditional manual management will cause low efficiency of goods supply. The problems such as material management and data collection have become increasingly unable to meet the requirements of enterprise inventory management. The traditional inventory after making the decision need human convey, but the inventory management system as new way of management, inventory management system can provide managers with a convenient platform directly communicate need to complete the operation, task allocation directly down, people can be faster to implement.



OS5 Intelligent Control (5)

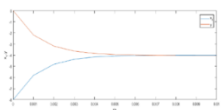
Chair Yingmin Jia (Beihang University, China)

Co-Chair Weicun Zhang (University of Science and Technology Beijing, China)

OS5-1 Geometry Structure Oriented Nonlinear Internal Model Based Manifold Consensus

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

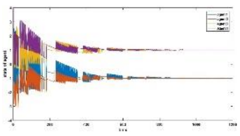
This note comes with manifold consensus based on nonlinear internal model. To be special, scheme demonstrated here is not necessary to inject the nonlinear internal model with additional extraneous augmented system. And this amazing result is made possible empowered by geometry structure, to be specific, Riemannian metric is employed to modeling the internal model of the nonlinear manifold. In case of completeness, the consensus of a first order linear agent and another one second order oscillator is provided to verify the suggested program.



OS5-2 Consensus Control of Linear Discrete-time Multi-agent Systems with Limited Communication Data Rate

Jintao Hu, Yingmin Jia, Yaxin Li (Beihang University, China)

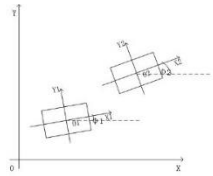
This paper investigates the consensus problem of linear discrete-time multi-agent systems with limited communication data rate and the cooperative-antagonistic interactions. A consensus control protocol is designed based upon a dynamic encoding-decoding mechanism. By means of the proposed control protocol, it is guaranteed that the agents can attain bipartite consensus if the signed graph is connected and structurally balanced, and besides the states of agents can reach zeros if the signed graph is connected but structurally unbalanced. Moreover, the clear form of the convergence rate is given. Finally, the numerous simulations are presented to illustrate the feasibility of the proposed control protocol.



OS5-3 Formation control for rectangular agents with communication maintenance and collision avoidance

Yaxin Li, Yingmin Jia, Jintao Hu (Beihang University, China)

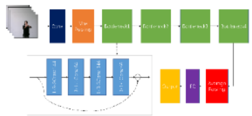
Formation control has many potential applications in various fields. This paper mainly discusses the rectangular agents, which is not often considered in other papers. Coordinate transformation is used to describe the location relationship between two agents. Obstacle function between agents, obstacle function between agents and obstacles are designed to ensure no collisions. Graph Laplacian is used to formation rotation. The distributed control protocol is designed to achieve desired formation with no collisions and communication maintenance between agents. Stability analysis proves the effectiveness of the algorithm.



OS5-4 Research on Sign Language Recognition Algorithm Based on Improved R(2+1)D

Yueqin Sheng[1], Qunpo Liu[1], Naohiko Hanajima[2], Bo Su[1]
([1]Henan Polytechnic University, China), ([2]Muroran Institute of Technology, Japan)

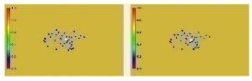
Sign language is a significant communication tool for deaf or hearing-impaired people. The study of sign language recognition can promote the communication between deaf-mute people and normal people, which is of great importance to improve the quality of life of deaf-mutes. Meanwhile, as an important part of multi-mode human-computer interaction technology, the study of sign language recognition is also pushing the progress of intelligent human-computer interaction. Traditional sign language recognition methods can only solve the problem of sign language recognition with a certain data scale. These algorithms are complex and have low generalization. So this paper uses deep learning technology for sign language recognition. In this paper, an improved R(2+1)D model is proposed for isolated word recognition. This model separated the spatial convolution and temporal convolution, so that the optimization can be decomposed and the model can be optimized better. The validity of the proposed algorithm is verified on CSL data set.



OS5-5 Directional Flocking of Multi-Agent system Caused by Limited Visual Field

Yongnan Jia[1], Yong Xie[2], Weicun Zhang[1]
([1]University of Science and Technology Beijing, China)
([2]The System Design Institute of Mechanical-Electrical Engineering, China)

Experiment evidence has proved that the visual field of each individual in biological swarms is usually non-omnidirectional. Therefore, we introduce limited visual field to the egalitarian flocking model. The directional flocking problem refers to the flocking problem that all the individuals are expected to move in a specified direction, which is decided by the leader. This paper mainly compared the limited-visual-field flocking model with the classic flocking model (that is the egalitarian one) from the point view of rate of convergence. Experimental results indicated that limited-visual-field flocking model is more efficient than the omnidirectional one for the directional flocking problem.



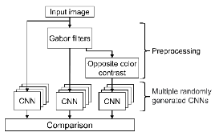
OS6 Bio-inspired Artificial Vision -Systems and Applications-(5)

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

OS6-1 The effect of preprocessing with Gabor filters on image classification using CNNs

Akito Morita, Hirotsugu Okuno (Osaka Institute of Technology, Japan)

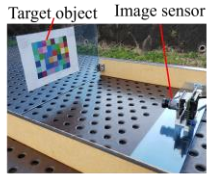
Preprocessing of input images is one of the promising approaches for improving the accuracy in image classification using artificial neural networks. In this study, we investigated the effect of neuro-inspired preprocessing, such as Gabor filtering and opposite color contrast, on the classification accuracy of convolutional neural networks (CNNs). We compared the classification accuracy of CNNs with the following four types preprocessing: no preprocessing, Gabor filtering, opposite color contrast, and the combination of Gabor filtering and opposite color contrast. To eliminate the influence of the topology of the CNN, the average accuracy of multiple CNNs whose parameters (the number of layers and neurons and the size of kernels for convolution) were chosen randomly was evaluated. The STL10 dataset was used for training and evaluation. The results showed that Gabor filtering increased the classification accuracy, whereas the opposite color contrast did not increase the accuracy significantly.



OS6-2 An image coding algorithm with color constancy using the Retinex theory and the Naka-Rushton equation

Shota Hisamitsu, Hirotsugu Okuno (Osaka Institute of Technology, Japan)

Illumination-independent coding of color information is required for a wide range of image recognition tasks because the lighting conditions strongly affect the apparent color of the object. In this study, we proposed an image-coding algorithm with color constancy based on the center / surround (C/S) retinex model and the Naka-Rushton equation; this equation was used in place of the logarithmic function used in the original C/S retinex model in order to encode the intensity information efficiently. Using images acquired under various lighting conditions, we compared the output of our algorithm with the gray world algorithm and the original C/S retinex algorithm. Using the same image set, we also investigated how the parameters of the Naka-Rushton equation affect the change in the color information output. The results suggested that setting the parameters depending on the image statistics around the region of interest could improve the performance of color discrimination.

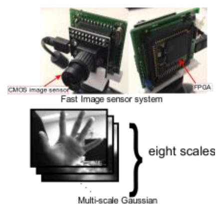


Images were acquired outdoors at the daytime and the evening

OS6-3 A fast image sensor system with an efficient multi-scale Gaussian filtering circuit

Yuuki Yamaji, Akito Morita, Hirotugu Okuno (Osaka Institute of Technology, Japan)

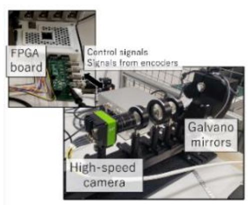
Real-time extraction of multi-scale visual features is one of the most important functions in robotic vision systems. In particular, many bio-inspired visual processing algorithms rely on multi-scale Gaussian and/or Gabor filters because the early stages of visual nervous system are modeled as such filters. Fast multi-scale Gaussian filtering is an essential component for implementing such functions. In the present study, we designed a multi-scale Gaussian filtering circuit whose coefficients of the standard deviation are selectable from any multiple of the square root of two. We also developed an image sensor system composed of a CMOS image sensor and a field-programmable gate array (FPGA) that contains the proposed filtering circuit. The system provided eight images whose resolution is 160 x 120 filtered by different scales of Gaussian at 156 frames / second.



OS6-4 A robotic vision system emulating fixational eye movements and retinal sampling

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

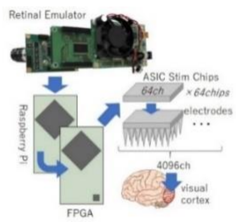
Recent studies on visual physiology have suggested that fixational eye movements contribute to the information processing function on the retina. In this study, we built a robotic vision system that emulates the characteristics of fixational eye movements and retinal sampling as an assistant tool for clarifying information processing through the collaboration of eye movements and neural circuits. The proposed robot vision system consists of two galvanometers, A field-programmable gate array (FPGA) board, and a high-speed, high-resolution camera. We measured the frequency response of this robot vision system and the results showed that it satisfied the requirements for emulating the frequency characteristics of biological eye movements. We also generated command signals in the FPGA to emulate biological eye movements, and confirmed the satisfactory operation of the system.



OS6-5 A bench-test system of the cortical prostheses utilizing retino-morphic spike as the driver signal of intracortical macrostimulation

Ryosuke Okada[1], Shinnosuke Ishikawa[2,] Tetsufumi Tasaki[2], Tetsuya Yagi[2,3], Yuki Hayashida[1,2]
([1]Mie University, [2]Osaka University, [3]Fukui University of Technology, Japan)

Neuromorphic retinal emulators, which mimic computations and/or functional architectures in biological retinas, have been developed as hardware systems so that the information processing is executed in biological time scale, thereby making them available for the visual prosthetics for acquired blindness suffering from retinal diseases, like glaucoma or diabetic retinopathy. In this study, we developed a prototype hardware system of such a prosthetic, in which point-process spike signals from our previously developed retinal emulator were utilized for the electrical microstimulation applied to the primary visual cortex of the cerebrum. The system was consisted of the retinal emulator, a single board computer, a FPGA board, and our previously developed microstimulator ASIC chip connecting to multiple stimulating electrodes. By substituting the electrodes with a micro-LED array, the system operations were verified through bench tests, in which the spatial pattern of stimuli via 4096 output channels was controlled correspondingly to the visual scene.

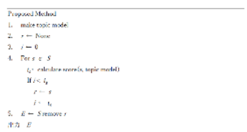


OS7 Media Information Processing, Music Recommendation and Artificial Intelligence (4)
Chair Yasunari Yoshitomi (Kyoto Prefectural University, Japan)
Co-Chair Masayoshi Tabuse (Kyoto Prefectural University, Japan)

OS7-1 Data expansion method by combining unnecessary sentence deletion and most important sentence addition

Tomohito Ouchi, Masayoshi Tabuse (Kyoto Prefectural University, Japan)

We are studying data expansion methods in automatic summarization systems. What has been found in the research to date is that the method of expanding the input article with unnecessary sentences deleted is the most effective of the extended methods. In the existing research, we have tried a method of adding most important sentences. In this research, we propose a method that combines the deletion of unnecessary sentences and the addition of most important sentences. We propose a hybrid method with two methods, one is to add important sentences first and the other is to add important sentences last.



OS7-2 Evaluation of a system that the reading of sentences by a voice synthesizer and the highlighting sentences

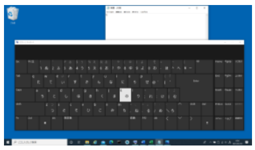
Sota Kobayashi, Masayoshi Tabuse (Kyoto Prefectural University, Japan)

It is generally known that there are individual differences in human cognitive function. For example, some people have an advantage over reading the text, while others have an advantage over hearing the words. We made a hypothesis; giving both visual and audio stimuli may make it easier for people to catch information. To test this hypothesis, we conducted an experiment. For the experiment, three indicators are set consisting of memory, understanding and concentration. Memory and understanding scores are measured by using a problem set given in a previous study. The problem set allows us to estimate the scores of verbatim memory and understanding. Concentration scores are estimated by the heart rate: previous study suggested a relationship between concentration and the heart rate. In this study, we made a system that realizes the reading of sentences by a voice synthesizer and the highlighting sentences. In the experiment using the system, we measured the difference in learning effect due to the difference in reading.

OS7-3 Mouse Cursor Control System using Eye Gaze and detection of eye opening and closing

Masayoshi Tabuse, Mana Fukumoto, Yasunari Yoshitomi, Taro Asada (Kyoto Prefectural University, Japan)

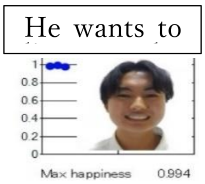
It is necessary to support of computer operation for a physically disabled person. One of the possible physical movements of the physically disabled person is Eye movement. Eye gaze and detection of eye opening and closing of a person makes it possible to operate a computer. In our system, a web camera, Gaze pointer, dlib C++ library and OpenCV library are used to recognize eye gaze and detect eye opening and closing. Changing the eye gaze, we can move a mouse cursor. Recognizing closed eye, we can carry out an operation of mouse click. In this paper, we evaluated the effect on operability due to the eye gaze and detection of eye opening and closing.



OS7-4 Music Recommendation System Driven by Facial Expression Recognition

Taro Asada[1], Motoki Kawamura[2], Yasunari Yoshitomi[1], Masayoshi Tabuse[1]
([1]Kyoto Prefectural University, Japan, [2]Fujitsu Limited, Japan)

We propose a system for music recommendation through an interaction between a user and a personified agent using facial expression recognition, synthesized voice and facial expression. Though expressing one of happy and neutral facial expressions, the user tells the system his or her evaluation on a song reproduced by the system. A happy facial expression means that he or she wants to listen to the song again, while a neutral facial expression means vice versa. The system exploits a series of his or her evaluations on previously recommended songs for deciding the next recommendation of song. The facial expression of the user is recognized using Face API. The effectiveness of the proposed system is verified.



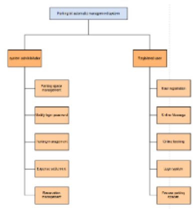
OS8 Intelligent Systems and Robotics -3 (4)-
Chair Tianyi Zhang (Tianjin University of Science and Technology, China)
Co-Chair Yiting Gao (Tianjin University of Science and Technology, China)

OS8-1 A Portable Electrocardiograph System Design based on STM32 Chip
Tianyi Zhang, Huating Liu (Tianjin University of Science and Technology, China)

ECG monitoring instruments occupy an important position in the medical field of application.It has a great reference value for the testing of basic cardiac functions and its pathological studies. For the shortcomings of conventional ECG monitoring instruments and equipment, which are large and not easy to carry, this paper designs a portable ECG monitor. In this paper, we designed a portable infrared induction heart rate measurement device, using STM32C8T6 as the core chip, to complete the extraction and processing analysis of ECG signals. The complete set thus composed can make basic diagnosis of the collected data and is successfully applied to daily life.

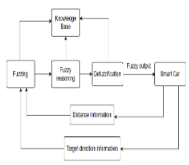
OS8-2 Design and Development of The Parking Space Autonomous Management System
Yiting Gao, Tianyi Zhang (Tianjin University of Science and Technology, China)

The article purpose is to change the manual or semi-manual management model that parking Spaces now have, so as to make the parking space management more standardized and more effective. This system used the Browser/Server architecture, JSP(Java Server Page) as the front r & d tool, SQL Server as the background data warehouse for r & d. The system key functions realized in the last system include: user registration and login, message block, parking reservation, change of login password, fee settlement block and parking information query. The design manufacture of this parking space autonomous management system can increase the effect of parking space management to a large extent, all reflecting the information period of the special good.



OS8-3 An Overview of Obstacle Avoidance Methods for Unmanned Vehicles
Peng Lu, Haokang Wen (Tianjin University of Science and Technology, China)

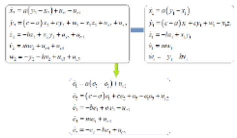
Autonomous obstacle avoidance is one of the popular research elements in the field of intelligent unmanned vehicles, and it is a key technology to realize the automatic travel of intelligent unmanned vehicles. This paper introduces the traditional algorithms and intelligent algorithms related to autonomous obstacle avoidance, and analyze the advantages and drawbacks of the corresponding algorithms. The development of the corresponding control strategies is also summarized and summarized in order to provide some reference for the research on obstacle avoidance of intelligent unmanned vehicles.



OS8-4 Synchronization of Novel 5D Hyperchaotic Systems

Hong Niu (Tianjin University of Science & Technology, China)

In this paper, synchronization of novel five-dimensional (5D) autonomous hyperchaotic systems is studied. The synchronization control law is proposed based on the center translation method. A structure compensator is formulated to make the mathematical model of the error system the same as that of the response system, and a linear feedback controller is designed via the Lyapunov stability theory to make the error system globally asymptotically stable at the origin. Thus, the two 5D hyperchaotic systems are synchronized. Some relevant numerical simulation results, such as the curves of the corresponding synchronization state variables and the errors, are given to illustrate the feasibility and effectiveness of the synchronization control law.



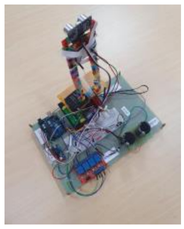
OS9 Robot Control (10)

Chair Yizhun Peng (Tianjin University of Science and Technology, China)
Co-Chair Junhui Yin (Tianjin University of Science and Technology, China)

OS9-1 Design of Analog Electromagnetic Gun Based on Arduino

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

The design of analog electromagnetic gun based on Arduino simulates the whole automated process of the electromagnetic gun discovering the target and launching the shell. The OpenMV camera on the simulated electromagnetic gun recognizes and determines the position of the target-shaped guide mark with a red radius of 20cm, then transmits the data to the Arduino to adjust the position of the simulated turret. The next step is that ultrasonic module measures the distance between the target-shaped guide and the turret before controlling the steering gear to rotate the barrel to the corresponding position. The shell in the barrel which is under the action of the magnetic field is accurately shot onto the target with a radius of 5cm on the ground.



OS9-2 A Design of Intelligent House Inspection Robot

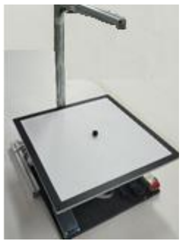
Sicong Wang, Yizhun Peng (Tianjin University of Science and Technology, China)

Aiming at the social pain point that the tedious procedure of housing quality inspection leads to the consumption of a lot of manpower, an intelligent house quality inspection robot is designed. The robot can independently detect the quality of the walls and other structures of the indoor house. It is an intelligent detection robot based on Internet of things technology. The robot is independently developed and designed by our ROS operating system, combined with embedded devices and other devices. It can independently navigate and control by remote control. It can adapt to terrain in a variety of complex environments and obtain various environmental information for mapping. It can measure the defects of the house more quickly and accurately. It does not need to carry other large detection tools, which makes the detection process more convenient, saves a lot of manpower, and realizes the unmanned and intelligent detection.

OS9-3 A Design of Embedded Plate & Ball Control System Based on Machine Vision

Wendinig Luo[1], FuChen Zhao[2]
([1]Beijing Huadian Jingneng Energy Technology Co., China)
([2]Beijing University of Posts and Telecommunications, China)

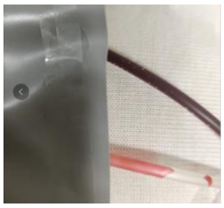
The control system is an experimental device to track the position of the target object and control the trajectory of the sphere by the actuator-rudder-driven platform motion. The research content of this design includes image processing, machine vision, motion control and many other fields. In this paper, the mechanical structure of the plate & ball control system and the printed circuit board of the main controller are designed, and the PID-based control algorithm is designed on the basis of the ball control ability analysis and the real-time detection analysis of machine vision. On the hardware platform with good self-designed performance, the precise sphere coordinates are obtained by the Hough circle detection algorithm, and then the position and speed of the sphere are controlled by the PID controller.



OS9-4 Design of Blood Circulation System of Medical Simulation Robot

Songyun Shi, Yang Ge, Chomgxu Guo, Yizhun Peng (Tianjin University of Science and Technology, China)

Medical robot is the most active direction in the research field of medical industry in recent years. Its development potential is very huge. It has greatly promoted the progress of modern medical equipment. The blood circulation system of our medical simulation robot can be used in medical universities, medical clinical training and even military fields. The robot adopts ROS robot operating system and FreeRTOS real-time operating system, and uses multi-sensor fusion data analysis to obtain the information in the blood sample and realize the simulation of blood circulation. Iit can simulate the touch of syringe in human injection. The whole medical process is optimized.



OS9-5 A Design of Dynamic Exoskeleton for Self-learning Human Movements

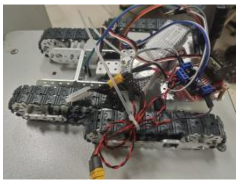
Qingliang Liu, Yucheng, Pengyu Yao, Dechao Wang,Yizhun Peng
(Tianjin University of Science and Technology, China)

In order to make the athletes more flexible and simple in the movement training, this project designed a kind of intelligent exoskeleton for the athletes to learn and complete a set of fixed movements with a higher efficiency in the training. In order to optimize the user's sports experience and adapt to more sports scenes, the product is divided into three parts: embedded device, mobile phone client and background server. The smart exoskeleton can be mechanically worn only by connecting the leg bandage and hand bandage to the human body. A single person can be completed, more convenient. It is a multi-degree-of-freedom humanoid robot system that can follow the movement of human limbs in real time. The equipment sets a series of standard movements, which leads human to repeat exercises and slowly forms muscle memory. The main target of the design is the people who want to quickly and accurately achieve the purpose of action learning.

OS9-6 Design of Autonomous navigation building Climbing and Handling Robot Based on SLAM

Linhui Chen, Junjie Tai, Yizhun Peng (Tianjin University of Science and Technology, China)

When there is no elevator at the station and airport or the elevator is crowded, the ramp of baggage checking is steep, the weight of luggage package is too heavy, it takes time and effort, and it is very easy to knock against and cause damage to the contents of the package. Our climbing robot carries out autonomous navigation based on visual SLAM. Equipped with ROS robot operating system, it adopts laser SLAM mapping navigation technology and tracks as running parts, which is suitable for various terrains. The adoption of binocular vision, can be more accurate analysis of the surrounding environment. The robot is equipped with MPU9250 attitude sensor, which is convenient for solving robot attitude. Jetson Nano is used as the upper computer and STM32F429 as the lower computer.



OS9-7 Simulation Research on Automatic Navigation of Indoor Wheelchair

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

A new auto-navigation wheelchair based on ROS system is proposed to deal with the global aging and the low behavioral ability of the elderly. Lidar is used to locate and map the active area using gmapping algorithm. Real-time map information is transmitted to the processor by the camera and lidar working together. Automatic navigation is completed by A* algorithm calculation. Important information points are marked by QR code and precisely positioned by camera recognition, which enables wheelchair to have automatic navigation function. It can help older people move safer and more freely at home; It can also be applied to nursing homes to reduce the pressure of nurses and centralize management of the elder.

OS9-8 Design of a Humanoid Dance Robot for Dancing Baduanjin

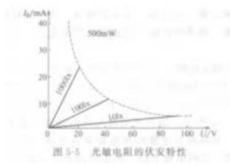
Zhuofan Xu, Ruitao Li, Binfu Zhong, Yizhun Peng (Tianjin University of Science and Technology, China)

This paper introduces a humanoid robot which consists of 16 servo motors driving joint and adopts STM32 as the main control chip and dual mode Bluetooth3.0/4.0BLE+EDR wireless control.By programming, the motor of humanoid robot can reach the specified space position and attitude at the specified time.At the same time, digital closed-loop control PID algorithm and synchronous compensation algorithm are used to modify and optimize the Angle of the servo motor in real time, and the designed action is vividly displayed.

OS9-9 Design of Intelligent Shading System Suitable for Parenting Products

Yihan Yan, Songyun Shi, Yaxin Shen, Yizhun Peng (Tianjin University of Science and Technology, China)

With the development of science and technology, if you want to provide a more safe and healthy growth environment for babies, simple human protection is far from enough. And as smart home is gradually accepted by the public, smart home does bring people a more convenient life In daily life, proper exposure to the sun can help nutrients be quickly absorbed and used, so that babies can better explore this new world, but direct sunlight will bring great harm to babies' eyes This product is an intelligent shading system, which can be connected with the smart home system and the Internet. Based on specific seasons, time periods and different geographical locations, the position of the sun and light intensity can be determined, so as to adjust the shading Angle in time, so that babies can bask more scientifically.



OS9-10 Design of Intelligent Personalized Nutrition Supplement Machine

Xinpeng Yang, Sidan Liu, Rui Li, Junhui Yin, Yizhun Peng
(Tianjin University of Science and Technology, China)

Reasonable nutritional supplement is the necessity for growth and development of living organisms. Residents need various nutrients daily to maintain bodies. However, in the post-epidemic era, people's physical and mental pressure has increased due to the fast life pace, and most Chinese residents are sub-healthy. To improve the current incomplete and untimely citizen nutrition supplement mode, this work uses big data, artificial intelligence and Internet of things to design an unmanned intelligent personalized nutrition supplement machine. Combining individual body data automatically, the machine can prepare exclusive nutritional supplement drinks timely, avoid crowd contact and support the citizen health quality improvement with effective software and hardware .

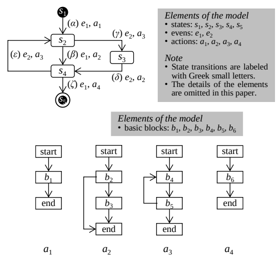
OS10 Software Development Support Method (4)

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

OS10-1 N-Switch and All-Path Test Coverage Criterion for Extended Finite State Machine

Tomohiko Takagi[1], Koichiro Sakata[1], Kouichi Akiyama[2]
([1]Kagawa University, Japan), ([2]Japan WillTech Solution Co., Ltd., Japan)

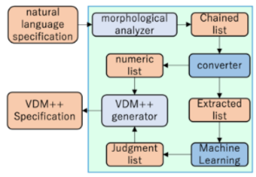
EFSM (Extended Finite State Machine) enables engineers to define the expected behavior of software from the aspect of not only state transitions but also actions on the state transitions. Test cases are usually created from EFSM models so as to satisfy a test coverage criterion called N-switch. However, it is originally designed for FSM, and actions are not taken into account in it. To address this problem, we propose a new test coverage criterion for EFSM. Our criterion requires that (i) test cases cover all the successive state transition sequences of specified length, and also (ii) the test cases cover all the paths on control flow graphs of actions that accompany each of the successive state transition sequences. (i) and (ii) are the characteristics of N-switch and AP (All-Path test coverage criterion), respectively.



OS10-2 Proposal of a Method to Generate Classes and Instance Variable Definitions in the VDM++ Specification from Natural Language Specification

Kensuke Suga[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)

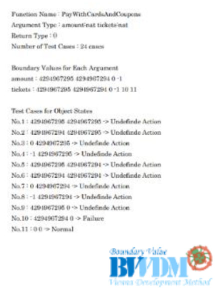
VDM is one of methodology on the formal methods to write the specification without ambiguity. Writing VDM++ specification is difficult. Our laboratory proposed a method for automatically generating VDM++ specifications from natural language specifications using machine learning. However, the existing method is not useful because it only supports type definitions and constant definitions in the VDM++ specification. This paper proposes a method to generate classes and instance variable definitions in the VDM++ specification from natural language specification. The superordinate and subordinate relationships between words are quantified, and then they are used as new parameters for machine learning. It is confirmed that the proposed method gave more useful results than the existing method.



OS10-3 Expansion of Application Scope and Addition of a Function for Operations into BWDM to Generate Test Cases from VDM++ Specification

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

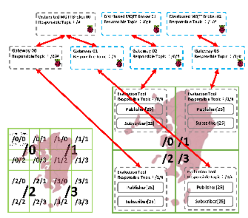
Generating test cases from formal specification descriptions VDM++, which is a method for the ambiguity of specifications, is time-consuming and labor-intensive. Therefore, we developed BWDM, which is an automatic test case generation tool for VDM++ specifications, in our laboratory. However, the existing BWDM does not support type definition blocks and conditional expressions for invariant conditions, pre-conditions, and post-conditions. Moreover, it cannot generate test cases for operation definitions that manipulate a state of objects. Therefore, to improve the usefulness of BWDM, this research extends BWDM to solve the above three problems. Consequently, it is confirmed that the extended BWDM can save about 17 minutes in generating test cases compared to test cases generation by hand.



OS10-4 Proposal of Gamma which is a spatial data sharing distributed MQTT system

Takahiro Ueda[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)

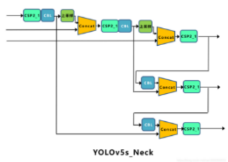
This paper proposes Gamma, which is a new distributed MQTT system, to improve the usefulness of distributed MQTT systems for sharing spatial data. Gamma that adds a Manager to the previous study is implemented with five functions. In the experiment, it was found that Gamma can process more messages than a single MQTT broker. It is confirmed that Gamma achieves scalability by increasing the number of Gateways and distributed MQTT brokers. Furthermore, it is confirmed that the efficiency of the distributed MQTT system can be improved by setting the responsibility area of Gateways so that its load is not unevenly distributed to a specific Gateway.



OS11 Applications of deep learning algorithms (9)
Chair Xiaoyan Chen (Tianjin University of Science and Technology, China)
Co-Chair Shiming Wang (Tianjin University of Science and Technology, China)

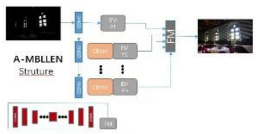
OS11-1 Human-vehicle detection based on YOLOv5
Zhihui Chen[1], Xiaoyan Chen[1], Xiaoning Yan[2], Shuangwu Zheng[2]
([1]Tianjin University of Science and Technology, China), ([2]Shenzhen Softsz Co. Ltd., China)

With the continuous improvement of social development level, traffic has become complicated. Therefore, the detection of people and vehicles becomes important. There are many application scenarios for human-vehicle detection, such as autonomous driving and transportation. This paper mainly introduces the research status of human-vehicle detection, analyzes the advantages and disadvantages of various current target detection algorithms, and focuses on YOLOv5 algorithm. Because the YOLOv5 model is much smaller than YOLOv4, and also has strong detection ability. Finally, YOLOv5 is used to carry out human-vehicle detection experiments. The results the detection accuracy is improved slightly.



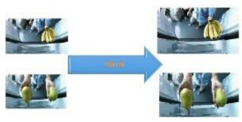
OS11-2 Low light enhancement CNN Network based on attention mechanism
Xiwen liang[1], Xiaoyan Chen[1], Hao Feng[1], Xiaoning Yan[2], Nenghua Xu[2]
([1]Tianjin University of Science and Technology, China), ([2]Shenzhen Softsz Co. Ltd., China)

Low-light enhancement is a challenging task. With the image brightness increasing, the noises are amplified, and with the contrast and detail increasing, the false information is generated. In order to solve this problem, this paper proposes a novel end-to-end attention-guided method (A-MBLLEN) based on multi-branch convolutional neural network. The proposed network is composed with enhancement module (EM) and Convolutional Block Attention Module (CBAM). The attention module can make the CNN network structure gradually focus on the weak light area in the image, and the enhancement module can fully highlight the multi-branch feature graph under the guidance of attention. In this manner, image quality is improved from different aspects. Extensive experiments demonstrate that our method can produce high fidelity enhancement results for low-light images quantitatively and visually.



OS11-3 Fruit Recognition Based on YOLOX
Keying Ren[1], Xiaoyan Chen[1], Wangzi Chen[1], Xiaoning Yan[2], Dongyang Zhang[1]
([1]Tianjin University of Science and Technology, China), ([2]Shenzhen Softsz Co. Ltd., China)

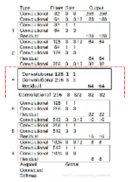
Pattern recognition is an urgent problem to be solved in the field of computer vision. In this paper, the network of fruit recognition based on YOLOX is studied. Due to the problem of slow training speed and low accuracy in the classical algorithms, the de-coupling detection head is optimized in YOLOX to overcome the above shortcomings. In terms of data enhancement, a new method combining Mosaic and MixUp is proposed. Through experimental verification, the method proposed in this paper has a great improvement over related algorithms such as YOLOV5, the accuracy is 98.6%, which is increased 5.2%.



OS11-4 An improved small target detection method based on YOLOv4

Xia Miao[1], Xiaoyan Chen[1], Keying Ren[1], Zichen Wang[1], Xiaoning Yan[2], Yue Sun[2]
([1]Tianjin University of Science and Technology, China), ([2]Shenzhen Softsz Co. Ltd., China)

In order to improve the efficiency and accuracy of small target detection in current traffic flow, this research proposes an improved YOLOv4 framework and applies it to small target detection task. A new small target-friendly 4-fold down-sampling residual is added between the second and third residual blocks of CSP Darknet-53 block to improve the detection accuracy of small target. The novel YOLOv4 model is optimized by above strategy. Compared with the original network, the modified framework can significantly improve the recall rate and average detection accuracy of small target.



OS11-5 Yolov5-DP: A new method for detecting pedestrian aggregation

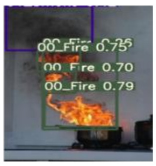
Kunzhi Yang[1], Xiaoyan Chen[1] Xiaoning Yan[2], Dashuo Wu[2]
([1]Tianjin University of Science and Technology, China), ([2]Shenzhen Softsz Co. Ltd., China)

In this paper, a novel network Yolov5-DP (Yolov5-DBSCAN-P) is proposed. Deep separable convolution and ACON-C activation function are added into Yolov5 network to improve the detection accuracy of pedestrians. Firstly, DBSCAN-P is used as the clustering detector to detect pedestrians in the area. Secondly, the depth-separable convolution is used to replace the common convolution in Yolov5. Finally, the loss function Swish is improved to ACON to increase the model speed and reduce the model size. The Yolov5-DP network is tested on the public dataset MOT20Det. The experimental results show that good detection results and accurate aggregation detection results are obtained.

OS11-6 Flame Recognition based on Yolov5 Algorithm

Kunzhi Yang[1], Xiaoyan Chen[1] Xiaoning Yan[2], Dongyang Zhang[2]
([1]Tianjin University of Science and Technology, China), ([2]Shenzhen Softsz Co. Ltd., China)

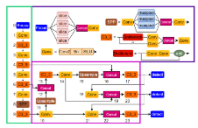
To address the problem of low accuracy and speed of flame detection, this paper proposes an improved YOLOv5 for flame detection. The new network is based on YOLOv5 by changing the loss function to DIoU(Distance Intersection over Union). Through introducing a large number of training data sets, it is hoped to improve the object detection accuracy. The experimental results show that the proposed YOLOv5 algorithm is effective with higher accuracy and faster detection for different flames.



OS11-7 Research on face detection algorithm based on improved YOLOv5

Zhen Mao[1], Xiaoyan Chen[1], Xiaoning Yan[2], Yuwei Zhao[2]
([1]Tianjin University of Science and Technology, China), ([2]Shenzhen Softsz Co. Ltd., China)

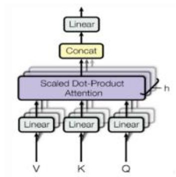
Face detection technology is one of the research hotspots in the field of deep learning in recent years. Aiming at the problems of slow detection speed and low accuracy of various target detection algorithms, this paper proposes an improved target detection algorithm based on YOLOv5. By introducing lightweight network, changing the depth and width of YOLOv5 network structure and reducing the number of model parameters, the network reasoning speed can be greatly accelerated. At the same time, the method uses Acon adaptive activation function to further improve the accuracy of face detection. Experimental results show that the improved algorithm has faster detection speed and higher detection accuracy than the traditional algorithms.



OS11-8 Visibility analysis based on a novel A-VGGNet network

Zhen Mao[1], Xiaoyan Chen[1], Xiaoning Yan[2], Shuangwu Zheng[2]
([1]Tianjin University of Science and Technology, China). ([2]Shenzhen Softsz Co. Ltd., China)

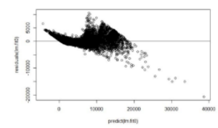
In recent years, visibility analysis through deep learning processing and analysis of video images for different places has become a hot research topic that attracts people's attention. A new deep learning model (A-VGGNet) is proposed to evaluate the visibility of real scenes. The model is constructed on the basis of the VGG classification model, and the classification accuracy of the deep learning model is improved by adding an attention mechanism. The experimental results show that the training success rate is 97.62%, the verification set test accuracy rate is 75.05%, and the test set classification accuracy rate is 85.05%. The proposed model has a good effect on the accuracy evaluation and classification of visibility.



OS11-9 Price Prediction of Diamonds

Xiran Wen, Qiqi Xu, Zirui Su, Jiayi Fang (The Chinese University of Hong Kong, China)

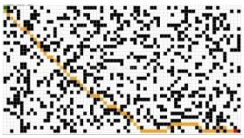
The experiment aimed at price prediction based on diamond dataset which contains 53940 rows of information. The model is constructed based on linear regression model with the lowest estimated test error among all methods including tree and nonlinear models. The experimental results show that the mean square error for the training dataset and validation dataset are 592182.6 and 603833.2 respectively, and the R^2 reached 98%. The test MSE is 631947. The proposed model can well predict diamond prices.



OS12 Intelligent Measurements and Control Systems (8)
Chair Xiaoyan Chen (Tianjin University of Science and Technology, China)
Co-Chair Shiming Wang (Tianjin University of Science and Technology, China)

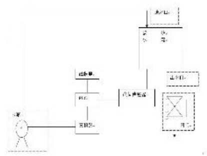
OS12-1 Research on an AGV path planning method
Zhihui Chen, Xiaoyan Chen, Shiming Wang (Tianjin University of Science and Technology, China)

AGV is an acronym for Automatic Guided Transport Vehicle. At present, the key technologies of AGV mainly include navigation and positioning technology, path planning technology, multi-AGV coordinated control technology and multi-sensor information fusion technology, etc. This paper studies A* algorithm in path planning. The A* algorithm can find the shortest path between two points. It mainly studies the principle of A* algorithm and simulates it in Matlab. By comparing the path length of ant colony algorithm and A* algorithm in Matlab grid graph, the excellent performance of A* algorithm is proved.



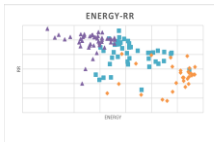
OS12-2 Boiler level measurement and control system
Xiwen Liang,Shengli Sun, Xiaoyan Chen (Tianjin University of Science and Technology, China)

This paper develops a boiler level control system to measure and control boiler level easily. This level control system applies PID control algorithm in SIEMENS S7-300PLC, uses ladder diagram and statement table programming, and completes variable connection and screen editing in MCGS monitoring system. The PLC boiler level control system is completed by establishing the system mathematical model, designing the level control algorithm, and verifying the simulation. Through debugging, modifying parts of the parameters, the control effect of the control system achieve satisfied expectation.



OS12-3 Recurrence quantification and time-frequency analysis of two-phase flow patterns
Keying Ren, Xiaoyan Chen , Meng Du (Tianjin University of Science and Technology, China)

Two phase flow often occurs in industrial production. If it is not detected in real time, it will do great harm to industrial production. In this paper, a method combining recurrence quantitative analysis and time-frequency representation is proposed to identify the flow patterns of the Gas-liquid flow. From the construction of the experimental device to the collection of conductance fluctuation signal of two-phase flow, the recurrence plot and WVD distribution map are drawn by MATLAB which are used in the final flow pattern analysis of two-phase flow. Experimental results show that our method can accurately identify the flow pattern of two-phase flow.



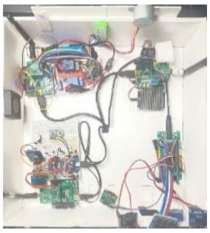
OS12-4 Application of Deep learning in automatic driving
Wei Su, Xiaoyan Chen (Tianjin University of Science and Technology, China)

With the continuous development of science and technology, the field of artificial intelligence has become a research hotspot, especially deep learning, which has attracted much attention from all walks of life. Starting with the automatic driving solution, this paper mainly expounds the important role and technical route of the deep learning applications, and finally looks forward to the development direction and application of automatic driving technology based on deep learning.

OS12-5 An intelligent home security system based on STM32

Wei Su[1], Xiaoyan Chen[1], Guangyong Xi[2]
([1]Tianjin University of Science and Technology, China), ([2]Zhengzhou University of Light Industry, China)

Based on STM32F103C8T6 host control chip, this paper completes the design and development of an intelligent home security system combined with various hardware modules, cloud server and android app. The system integrates OLED display screen, buzzer, human infrared, Wifi and other modules. The serial port is connected with ZigBee coordinator. The terminal equipment with temperature, humidity and smoke sensors is adopted to monitor the environment, and the relays and stepping motors are designed to control the equipment at home by simulation. The access control system takes STC89C52 as the core, and integrates the functions of photographing module and steering gear simulation door opening. Android app uses socket technology to complete remote data communication through Alibaba cloud server and indoor security system. After testing, the system runs stably and has good performance.



OS12-6 Design of smart bracelet based on STM32 microcontroller

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

With the development of electronic information technology, the demand for high-precision and convenient electronic detection equipment for biomedical signals goes straightly high. Heart rate and steps counting are two important indicators of the human body. Based on this, this design studies an intelligent bracelet with health monitoring system, which can detect human movement state and steps counting, measure and analyze the heart rate, and connect wirelessly through Bluetooth module. A STM32 microcontroller is used to collect and analyze the information of motion state, heart rate and step quantity, and then send them to the APP, which is developed based on Android and displayed through Bluetooth module. The users can change their lifestyle by above parameters. Smart bracelets can play a role in reducing the risk of disease.

OS12-7 Design of material conveying and automatic sorting control system based on PLC

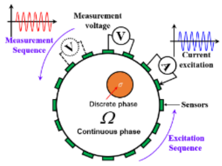
Qian Wang[1], XiaoYan Chen[1], Shengmin Cao[2]
([1]Tianjin University of Science and Technology, China), ([2]TangShan College, China)

The automatic sorting technology is widely used in all walks of life, such as logistics distribution center, post office, mining, port, etc. It can replace human to do a lot of monotonous reciprocating or high-precision work, in order to liberate people's hands to do more challenging and innovative work, and greatly improve work efficiency. This design uses programmable controller PLC and configuration software to design an automatic sorting control system. The PLC and MCGS present a friendly man-machine interface and powerful data management functions. The simulation results show that the system has high sorting efficiency and stable performances.

OS12-8 Image reconstruction based on ResV-Net for electrical impedance tomography

Qian Wang, ZiChen Wang, Di Wang, XiaoYan Chen (Tianjin University of Science and Technology, China)

Electrical impedance tomography (EIT) is a nonlinear and ill-posed inverse mathematical problem. Due to the above problem, the reconstruction image suffers from the serious artifacts. To overcome shortcomings, we proposed a residual V-shaped deep convolutional neural network (ResV-Net). It consists of feature extraction module and image reconstruction module which are optimized by ResBlock. The residual connection method can effectively increase the number of forward information flow and reverse gradient flow in deep CNN, and alleviate the problem of non-convergence caused by gradient vanishing. The simulation and experimental results show that the ResV-Net has better visualization effect than the related imaging method.



OS13 Mathematical Informatics (5)

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

OS13-1 Basic Study on Design Tool of Hula Costumes

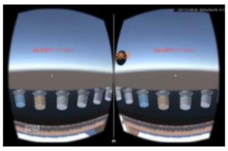
Taketo Kamasaka[1], Kodai Miyamoto[1], Makoto Sakamoto[1], Satoshi Ikeda[1], Amane Takei[1], Kenji Aoki[1], Tsutomu Ito[2], Takao Ito[3]
([1]University of Miyazaki, Japan), ([2]National Institute of Technology, Ube College, Japan)
([3]Hiroshima University, Japan)

Miyazaki Prefecture has a large hula population, probably due to its similarity to Hawaii in mythology and climate. On the other hand, many hula costumes are handmade, and it costs tens of thousands of yen to produce an original design. Therefore, we are developing a 3D computer-aided design (CAD) system for hula costumes, based on the idea that it is possible to reduce the number of failures by checking the behavior of the fabric when dancing in the designed costume before making it. There are some research examples on ordinary apparel CAD, but there is no precedent specific to hula costumes.

OS13-2 Basic Study on the Use of XR Technology to Support Science Education

Kodai Miyamoto[1], Taketo Kamasaka[1], Makoto Sakamoto[1], Masahiro Yokomichi[1], Satoshi Ikeda[1], Amane Takei[1], Tsutomu Ito[2], Takao Ito[3]
([1]University of Miyazaki, Japan), ([2]National Institute of Technology, Ube College, Japan), ([3]Hiroshima University, Japan)

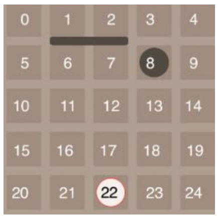
According to the results of a survey on science teaching in 2016, the percentage of students who answered that they like science is lower than other subjects. However, more than 80% of students said that they like experiments and observations. In addition, a 2019 survey on smartphone penetration showed that about 90% of students are familiar with the technology. Also, XR technology has made remarkable progress in recent years. Based on the above, I conducted this research because I thought that creating a simulation application using XR technology with smartphones would change the way we think about science classes. In this paper, we have developed a simulation application for science experiments. The subjects were asked to experience the created application and answer a questionnaire. As a result, the average score was 4 out of 5, which was not a bad result. At the same time, however, we found a problem. The problem was that since this was a simulation application, the user experience was not very good. So we wanted to make it a little easier to use. While improving the problems, I would like to create apps for other fields as well.



OS13-3 Analysis of 5x5 board Quoridor

Takuro Iwanaga, Makoto Sakamoto, Takao Ito, Satoshi Ikeda (University of Miyazaki, Japan)

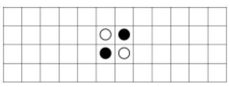
In this paper, we analyze Quoridor using retrograde analysis. *Quoridor* is a 2 or 4-player intuitive strategy game designed by Mirko Marchesi and published by Gigamic Games. In this study, we will analyze the game for two players. In this case, Quoridor is classified as finite two-person zero-sum games, so it is always classified as either a must-win game, a must-win game, or a tie game. In this case, Quoridor is finite two-person zero-sum games, so it can be classified as either a must-win game, a must-lose game, or a tie. First, we analyze the game with 25 squares (5x5) and one board each (originally 81 squares and 10 boards each). After that, we will increase the number of boards and squares to achieve the regular version of the analysis.



OS13-4 A perfect play in 4×12 board of Othello

Tomoyasu Toshimori, Makoto Sakamoto, Takao Ito, Satoshi Ikeda (University of Miyazaki, Japan)

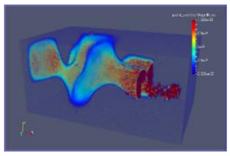
In 1993, mathematician Joel Feinstein discovered that White would win in a 6x6 Othello if both players did their best. In 2015, Takeshita worked on the analysis of the reduced Othello and succeeded in the complete analysis of 4x4, 4x6, 4x8, and 4x10 boards. A complete analysis of the 4x12 board predicted that the search space was about 10 ^ 4 times larger than the 4x10 board. Therefore, Takeshita tried to reduce the search space by referring to the perfect play of 4x10 board. As a result of searching after the 7th move when the procedure up to the 6th move on the 4x12 board was the same as that on the 4x10 board, it was confirmed that Black would win. In this paper, we will perform a complete analysis of the 4x12 board Othello before the 7th move.



OS13-5 Parallel full-wave electromagnetic field analysis based on domain decomposition method

Amane Takei, 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.



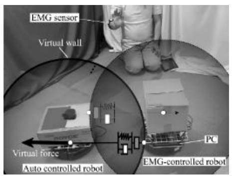
OS14 Biomedical Systems (4)

Chair Taro Shibanoki (Okayama University, Japan)
Co-Chair Hideyuki Tonooka (Ibaraki University, Japan)

OS14-1 A Mutual Control Method for a Multi-layered Non-contact Impedance Model-based mobile robots

Masaru Sasaki[1], Taro Shibanoki[2], Hideyuki Tonooka[1], Toshio Tsuji[3]
([1]Ibaraki University, Japan), ([2]Okayama University, Japan), ([3]Hiroshima University, Japan)

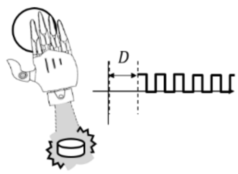
This paper proposes a mutual control method for multi-layered non-contact impedance model-based mobile robots. In the proposed system, the motion priority is set to the robot, and the stiffness, viscosity, and inertia parameters of the non-contact impedance model are changed appropriately according to the priority value, so that the robots can avoid collision with each other and obstacles at the same time. In the experiment, two mobile robots, including one controlled by EMG signals, were prepared and operated to intersect. The other mobile robot automatically stopped and resumed its movement in response to EMG-controlled robot with the high priority, indicating that the proposed method can be used to control multiple robots.



OS14-2 Relationship Between Delay Time and Sensation in Tactile Feedback for Myoelectric Prosthesis

Taro Shibanoki (Okayama University, Japan), Kosuke Jin (Ibaraki University, Japan)

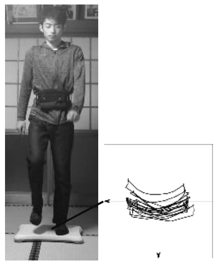
In this paper, we aim to develop a new tactile feedback method for myoelectric prosthetic hands and model the relationship between delay time and sensation in vibration stimulation. For myoelectric prosthetic hands, tactile sensation can be expressed by vibrating an oscillator attached to the socket based on information obtained from tactile sensors attached to the prosthesis's fingertips. In this case, if there is a time gap between the sensory input and the stimulus, there is a possibility of causing discomfort. Therefore, in the experiment performed, a delay time, D [s], is set between the start of contact with an object and the start of vibration using a tactile sensor and conducted NRS evaluation. The results showed that the discomfort was generated up to D = 0.8 [s] and then decreased according to the delay time. The results showed that the discomfort was induced by controlling the timing of the vibration stimulus.



OS14-3 Effects of Tactile Stimulation Near the Auricle on Body Sway During Foot Stamping

Masaya Tadokoro[1], Taro Shibanoki[2], Hideyuki Tonooka[1]
([1]Ibaraki University, Japan), ([2]Okayama University, Japan)

This paper describes the effects of tactile stimulation near auricles on body sway in during foot stamping. The system measures center of pressure, acceleration of upper body, and whole-body movements by skeleton tracking with a depth camera while the subject performs foot stamping on the stabilometer, and extracts evaluation indices based on four perspectives: 1. the amplitude, 2. the variation, and 3. rhythm of body sway, and 4. the correlation of each limb. In the prototype experiment conducted with one healthy mail, the body sway during foot stamping for ten trials in non-stimulus condition is compared with that for two trials in stimulus condition that the constant tactile stimulus is applied. As a result, the variation of the movement of lower limb and upper body sway were significantly reduced. This result implies that applying the constant tactile stimulation near both of auricles may be effective in stabilizing posture during foot stamping.



OS14-4 A Monitoring System of a Hamster Based on Video Image Analysis

Yugo Yamazaki[1], Taro Shibanoki[2], Hideyuki Tonooka[1]
([1]Ibaraki University, Japan), ([2]Okayama University, Japan)

This paper proposes a monitoring system of a hamster using a video camera. The proposed system first processed the video image taken from the top of the cage, and then extract features related to posture information and internal state. These features are used to discriminate between daily activities and other activities using machine learning technique. This allows the system to alert when the hamster behaves differently from its daily routine. In the experiments, we analyzed the daily behaviors of hamsters using the proposed system, and showed that the behaviors of hamsters change when stimuli are given from outside the cage, and that the system may be able to discriminate them appropriately.

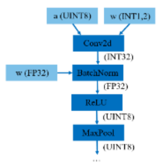
OS15 Artificial Intelligence for Embedded Systems and Robotics (5)

Chair Hakaru Tamukoh (Kyushu Institute of Technology, Japan)
Co-Chair Takuya Nanami (University of Tokyo, Japan)

OS15-1 INT8 Activation Ternary or Binary Weights Networks

Ninnart Fuengfusin, Hakaru Tamukoh (Kyushu Institute of Technology, Japan)

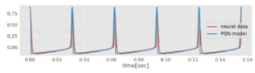
In this paper, we propose binary or ternary weights 8-bit integer activation convolutional neural network. This model is designed to fit as a middle ground between the 8-bit integer and low-bit (1-bit or 2-bit) quantized models. We discover that conventional low-bit quantization techniques (i.e., BinaryConnect and Ternary Weight Network) can be utilized with 8-bit integer quantization without any fractions. Based on these methods, we evaluate our model with the VGG16-like model and CIFAR10 dataset. Our model provides competitive results to the general floating-point model.



OS15-2 A parameter tuning method for PQN model

Daimon Sakai, Takuya Nanami, Takashi Kohno (University of Tokyo, Japan)

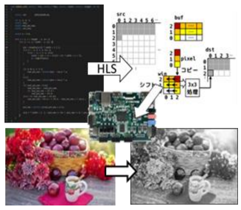
A piecewise quadratic neuron (PQN) model is a spiking neuron model that can be efficiently implemented on digital arithmetic circuits. In addition, this model can reproduce a variety of neuronal activities precisely with optimized parameter sets. In previous studies, we have optimized the parameters using meta-heuristic methods, which required a lot of computational time. In this paper, we propose a parameter estimation method that takes into account the mathematical structure of the model and reproduces the electrical activities of a targeted neuron with less computational time. We expect that this method can be used to construct silicon neuronal networks that faithfully replicate the nervous system.



OS15-3 Hardware Development of Edge-Preserving Bubble Image Conversion in High-level Synthesis

Jiang Qin, Akira Yamawaki (Kyushu Institute of Technology, Japan)

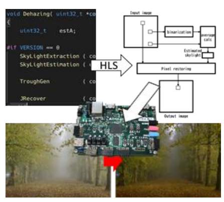
The non-photorealistic rendering, NPR, is widely used in social networking service on the mobile device. To realize a real time NPR with low power making battery life of mobile device longer, we attempt to develop hardware module by using high-level synthesis, HLS, converting software to hardware automatically. This research focuses on edge-preserving bubble image, EPB, converting photos into image like filled with bubbles. We proposed a software description method for EPB algorithm so that HLS can generate a high-performance and low-power hardware module. Through the practical experiments, we show that our proposed description method can make HLS generate good hardware module improving the performance and power efficiency compared with the conventional method.



OS15-4 Development of Haze Removing Hardware Using High-Level Synthesis

Daiki Shirai and Akira Yamawaki (Kyushu Institute of Technology, Japan)

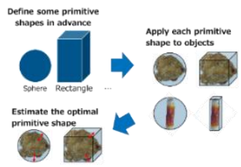
We develop a haze removing hardware using high-level synthesis (HLS). Haze removing removes a haze in the picture taken with a camera. It is frequently used as a pre-process for advanced image processing. However, haze removing generally requires a large amount of computational power due to its expensive and complex processing flow in software. Therefore, we decided to develop the haze removing as hardware to save power and improve processing speed at runtime. When developing hardware, we have used HLS converting software to hardware design automatically to flexibly respond to shortened development periods and specification changes. For the software used in high-level synthesis, the algorithm was changed and described so that the performance of the development hardware would be improved. Finally, the developed hardware was able to achieve an 1.43 times performance improvement and a run-time power improvement of about 41.4 times compared to the software execution.



OS15-5 Automatic approximation of primitive shapes using point clouds

Yuma Yoshimoto, Hakaru Tamukoh (Kyushu Institute of Technology, Japan)

This paper proposes a method to estimate appropriate primitive shapes by automatically using a point cloud of objects. The process is as shown in the flow. First, the method estimates a rotation angle of the object. Next, the method places the primitive shape in the center of the object. The primitive shape stretched or compressed to fit the object. The distance between all the points of the object and the primitive shape is measured. The method applies these processes using various primitive shapes. Finally, the most appropriate primitive shape is determined. The experiment confirmed that the method estimates objects like apples and chikuwa to primitive objects, such as spheres and cylinders.



OS16 Robotic Manipulation (4)

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

OS16-1 Motion Planning for Retrieving an Object in a Complex Environment

Shusei Nagato, Tomohiro Motoda, Keisuke Koyama, Weiwei Wan, Kensuke Harada (Osaka University, Japan)

This paper proposes a method for retrieving a target object by a robot in a complex environment. In such an environment, it becomes difficult to observe the state of the target object owing to occlusion, and to grasp it. In this work, we propose a method for selecting the viewpoint to observe the occlusion part of the target object by using RGB-D images and selecting the motion from grasping/dragging to retrieve the target object according to the situation. Finally, we confirmed that a robot can retrieve a certain target from the complex environments.

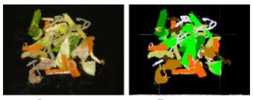
OS16-2 Design and Control of Two-sided Gripper for Bin Picking

He,Maike, Tokuo Tsuji, Tatsuhiro Hiramitsu, Hiroaki Seki (Kanazawa University, Japan)

Bin picking is a core problem in industry automation. The goal is to have a robot with sensors and cameras attached to it pick-up known objects with random poses out of a bin using a gripper. Bin picking technology has the good effect in solving the current shortage of labor. In this research we have completed The designed gripper grasp object from both sides. Proposed an idea of automation, completed the image processing and robot calibration.

OS16-3 Training Data Augmentation for Semantic Segmentation of Food Images Using Deep Learning

Takayuki Yamabe, Tatsuya Ishichi, Tokuo Tsuji, Tatsuhiro Hiramitsu, Hiroaki Seki
(Kanazawa University, Japan)



Input Output

With the advent of deep learning, the technology for category recognition of meal images has reached a practical stage. However, for more advanced food management, it is desirable to develop a technology that can recognize the distribution of food ingredients. We have focused on semantic segmentation, which is a technique for recognizing the distribution of food ingredients in pixels of an image. In order to achieve highly accurate image recognition, it is necessary for humans to carefully paint the boundaries of food items and assign correct labels, which requires a great deal of effort. In this study, we propose a method to automatically amplify appropriate training data by cropping, flipping/rotating, image composition, and color manipulation based on images of only one type of food.

OS16-4 Suitable Error Recovery Process using Combined Evaluation Standards in Robotic Manufacturing Plant

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

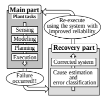


Fig. 1 Robot task system with an error recovery function

The number of manufacturing plants where industrial robots work is increasing. Therefore, errors during work are likely to occur. For big errors, it is often necessary to go back to the previous step and resume work. There are two issues: which step to return to and what kind of work to do from the point of return. In this paper, we will show that it is good to use a combination of multiple evaluation standards to decide the planning.

OS17 Advanced Robotics (8)

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

OS17-1 Experience in efficient real office environment modelling in Gazebo: a tutorial

Bulat Abbyasov[1], Kirill Kononov[1], Tatyana Tsoy[1], Martínez-García Edgar A.[2], Evgeni Magid[1]
([1]Kazan Federal University, Russia), ([2]The Autonomous University of Ciudad Juarez, Mexico)

New robotic solutions should be carefully verified before executing with real robots in real environments. Simulation provides a significant support in testing, but requires test sites with a high level of realism. In this case, 3D modeling can be used to produce the necessary 3D digital representation of real objects with varying difficulty. This article presents a step-by-step tutorial on modeling a realistic office environment. The environment contains a building skeleton frame, windows, building tiles, and furniture. Blender modeling toolset was used to create high-quality 3D models in the Gazebo simulator. The constructed virtual environment was validated with a lidar-based SLAM task for a UGV.

OS17-2 Graphical user interface design for a UAV teleoperation

Roman Lavrenov[1], Ramil Safin[1], Bai Yang[2], Martínez-García Edgar A.[3], Roman Meshcheryakov[4]
([1]Kazan Federal University, Russia), ([2]Ritsumeikan University, Japan)
([3]The Autonomous University of Ciudad Juarez, Mexico)
([4]V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Russia)

The number of drones being used around the world grows at a high speed. New drones' manufacturers are emerging, new drone designs are being developed. Most drones are controlled by a remote control, while UAV joysticks and communication protocols are different. However, the Robotic Operating System (ROS) allows unifying control process for drones. In this article, we present a universal graphical interface for controlling drones using ROS. The program is written in C++ and Qt and allows to control UAV and receive and visualize data from drones. Due to the use of ROS topics, this program can be applied to any drone with ROS.

OS17-3 Numerical solution approach for the ROBOTIS OP2 humanoid hand inverse kinematics

Zagidullin Linar[1], Tatyana Tsoy[1], Roman Meshcheryakov[2], Kuo-Hsien Hsia[3], Evgeni Magid[1]
([1]Kazan Federal University, Russia)
([2]V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Russia)
([3]National Yunlin University of Science and Technology, Taiwan)

Small-size humanoids are widely used in human-robot interaction (HRI) projects and activities. To operate robot limbs in HRI and pick-and-place tasks it is required to solve an inverse kinematics problem. Classical approaches are closed-form solutions with algebraic or geometric approaches or a numerical solution. While a typical numerical solution is supposed to search for joint variables using an iterative optimization, in this paper we suggest an off-line solution for a ROBOTIS OP2 humanoid upper limb via a forward kinematics approach that allows to calculate in advance all possible solutions for an end effector pose within a robot workspace with several levels of the workspace discretization. The solution was obtained in a simulation and successfully validated with a real ROBOTIS OP2 humanoid.

OS17-4 Alvus modelling in Gazebo

Liaisan Safarova[1], Bulat Abbyasov[1], Tatyana Tsoy[1], Hongbing Li[2], Evgeni Magid[1]
([1]Kazan Federal University, Russia), ([2]Shanghai Jiao Tong University, China)

Insufficient testing of medical robots can lead to accidents during a surgery and damage an expensive equipment. A simulated 3D patient permits a preliminary checking of robotics-based medical scenarios without threatening a real patient's health. This article presents a 3D model of a human abdomen, which contains vital organs: intestine, liver, stomach, and kidneys. There are 3 layers of an abdominal wall: skin, adipose, and muscle. Several types of pathologies were modeled: cysts, tumors, and gallstones. Blender modeling software was used to create realistic 3D models of organs with their distinctive features for Gazebo simulator. The model is presented as a ROS package with necessary configuration files and can be used by other researchers to simulate medical operations in Gazebo environment.

OS17-5 Testing procedures architecture for establishing a fiducial marker recognition quality in UAV-based visual marker tracking task in Gazebo simulator

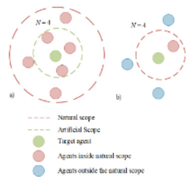
Mikhail Kilin[1], Roman Lavrenov[1], Bai Yang[2], Mikhail Svinin[2], Evgeni Magid[1]
([1]Kazan Federal University, Russia), ([2]Ritsumeikan University, Japan)

Fiducial markers could be used in different tasks, including UAV and UGV marker-based localization. In most cases developers do not consider features of fiducial markers' systems (FMS) while selecting a particular FMS for a project. However, this selection might significantly influence results of experiments and thus the quality of a resulting product, an algorithm or a software. In this work, we define an architecture of experimental framework that allows finding an optimal marker for a UAV in a mobile ground object following task. The proposed framework estimates an average deviation of a detected Aruco marker position and an accuracy of the UAV landing on the marker. The framework uses Robot Operating System and employs UAV PX4 LIRS model in the Gazebo simulator.

OS17-6 Feature importance evaluation method for multi-agent deep reinforcement learning in advanced robotics task allocation

Sergey Ryabtsev[1], Mikhail Gurchinsky[1], Igor Struchkov[1], Vyacheslav Petrenko[1], Fariza Tebueva[1], Sergey Makarenko[2]
([1]North-Caucasus Federal University, Russia),
([2]Saint Petersburg Federal Research Center of the Russian Academy of Sciences, Russia)

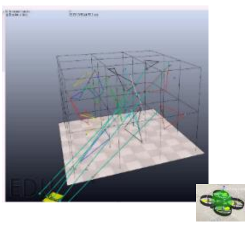
The need to tackle intelligent tasks using advanced robotics multi-agent systems (MAS) actualize the use of artificial neural networks (ANNs) and multi-agent deep reinforcement learning technology. The article aims to solve the problem of exponential growth of ANN complexity with an increase in the number of agents in the MAS. To solve this problem, we propose an evaluation method for input data features importance. This method allows to optimize the input data feature set to reduce the computational complexity of the ANN inference while providing the same level of performance.



OS17-7 Iterative method of labor division for multi-robotic systems

Sergey Ryabtsev[1], Artur Sakolchik[2], Vladimir Antonov[1], Vyacheslav Petrenko[1], Fariza Tebueva[1], Sergey Makarenko[3]
([1]North-Caucasus Federal University, Russia),
([2]Belarusian State University, Belarus)
([3]Saint Petersburg Federal Research Center of the Russian Academy of Sciences, Russia)

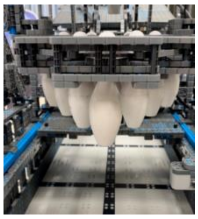
Labor division in multi-robotic systems allows to distribute tasks between agents in order to increase the efficiency of performing the global task. Collective decision-making methods allow agents to form the "agent-task" pairs. In this paper, we consider the case when the number of tasks significantly exceeds the number of agents. We propose an iterative method of labor division in multi-robotic systems. It uses collective decision-making to assign a cluster of subtasks to an agent. The paper examines different ratios between cluster size, number of clusters, and number of agents in order to find ratios that provide minimal average global task execution time and minimal average energy consumption.



OS17-8 Development of Bowling Machine Using VEX IQ

Kuo-Hsien Hsia[1], Ya-Chun Chen[1], Evgeni Magid[2], Xin-Ying Zeng[1]
([1]National Yunlin University of Science and Technology, Taiwan)
([2]Federal Kazan University, Russia)

VEX IQ is a kind of educational robotics platform focuses on semi-automatic and semi-remote. In this paper, a fully automatic bowling machine based on VEX IQ educational robotics platform has been developed. Since the brain of VEX IQ cannot communicate to other VEX IQ brains and the components of VEX IQ are made of plastic, it is necessary to overcome these problems to create mechanism similar to steel-made constructions and form an intelligent large system. We use sensors as communication interface for the brains of VEX IQ. Totally 3 brains and 10291 VEX IQ plastic components are used for the construction of the bowling machine. The overall size is about 252x93x90 in centimeters.



OS18 Natural Computing (3)

Chair Marion Oswald (Technische Universität Wien, Austria)
Co-Chair Yasuhiro Suzuki (Nagoya University, Japan)

OS18-1 An Acoustic Artificial Life System Using the Game of Life and its Application for Performing Arts

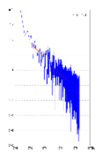
Yasuhiro Suzuki (Nagoya University., Japan)

Sound and vibration are forms of energy propagation. We have constructed an artificial life form that takes sound as energy and produces energy. The system converts sound into a two-dimensional pattern and uses it as input to the Game of Life. After n steps, the Game of Life is re-transformed into sound, and the sound is output. The sound allows the artificial life in the PC to interact with the outside world. We used this system to create an artwork that interacts with a dancer interactively through sound.

OS18-2 The Effect of Non-audible Low Frequency, Deep Micro Vibrotactile, DMV Sounds on Music

Yasuhiro Suzuki (Nagoya University, Japan)

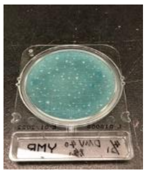
The 1/f fluctuation is a fluctuation in which the power spectrum is inversely proportional to the frequency f. 1/f is found in natural and manufactured phenomena and is widespread in music. Although sounds below 20Hz are inaudible to humans, the lowest notes of a pipe organ are 8Hz or 16Hz, and orchestral music contains inaudible low-frequency sounds. We have shown that a not 1/f can fluctuate at 1/f by adding shallow frequency sounds.



OS18-3 Response of Yeast to Low Frequency Sound Exposure

Yasuhiro Suzuki (Nagoya University, Japan)

The primordial Earth, where life began, has been exposed to solid infrasound due to active mountain building and intense weather. Yeast, Saccharomyces cerevisiae, was the first eukaryotic organism to emerge 2.1 billion years ago. We investigated the response of yeast to low-frequency sound. It has known that the yeast is cultured by shaking a culture flask, the intensity of shaking changes the growth rate. We have confirmed that the power of the low-frequency exposed sound may affect the growth rate.



OS19 Industrial Artificial Intelligence Robotics (7)

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

OS19-1 Online Deep Reinforcement Learning on Assigned Weight Spaghetti Grasping in One time using Soft Actor-Critic

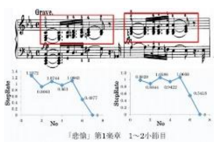
Prem Gamolped, Sakmongkon Chumkamon, Tomofumi Tsuji, Nattapat Kloomklang, Chanapol Piyavichayanon, Ranatchai Laosiripong, Eiji Hayashi (Kyushu Institute of Technology, Japan)

With the growth of the food industry recently, especially in pandemics, Artificial Intelligence and robotics have become essential and widely used to package food. Human error in food production can cause accuracy and performance issues since the food is soft, non-rigid, and nonpattern, mainly noodle-like. This paper proposes a challenging novel self-learning robotics grasping for spaghetti using deep reinforcement learning (DRL) based on the soft actor-critic. The manipulator is trained for one-shot grasping spaghetti where the robot could get the state from the RBGD camera. We also implement spaghetti detection and segmentation for the input state of the DRL. Finally, we present the evaluation and discussion of the spaghetti grasping.

OS19-2 The research about editing system of performance information for player piano. -Inference in the same phrase including ostinato-

Haruna Yamasaki, Sakmongkon Chumkamon, Eiji Hayashi (Kyushu Institute of Technology, Japan)

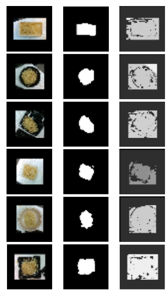
Playing the piano expressive by player piano, it is necessary to adjust the volume, length, and timing of music. In the case of piano music, there are often 1000 or more notes in the score of even a short piece of music. So, to edit music data manually requires not only knowledge but also a huge amount of time and effort. Therefore, we aimed to develop a system that, like a skilled pianist, can perform even the first musical score based on information related to previous skills and experience. So, we developed a system that automatically estimates the performance expression of unedited music using edited performance data and score data. In this paper, I studied the changes in performance expression when similar phrases such as ostinato are repeated in a piece by F. Chopin, and described a new method of inference.



OS19-3 Weight estimation for noodle products in food layout of a home replacement meal

Tomofumi Tsuji[1], Sakmongkon Chumkamon[1], Chanapol Piyavichayanon[1], Prem Gamolped[1], Ranatchai Laosiripong[1], Ayumu Tominaga[2], Ryusuke Fujisawa[1], Eiji Hayashi[1]
([1]Kyushu Institute of Technology, Japan)
([2] National Institute of Technology, Kitakyushu College, Japan)

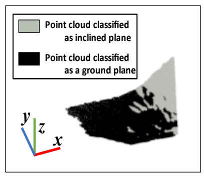
In recent years, there has been an increasing demand for robot automation to improve productivity in the food layout of lunchboxes and prepared foods in the Japanese home replacement meal industry. In this research, we are developing autonomous work robots that can perform midday meal serving tasks and developing the technology for industrial food automation using Artificial Intelligence (AI) to improve productivity, security, and safety. In this paper, we perform weight estimation of the served object to identify the amount of spaghetti grasped. We created our dataset of spaghetti used for weight estimation. The spaghetti is in different types of containers, the various weight of spaghetti, and the random position of spaghetti in the robot workspace. We also propose a deep learning method using RGB-D cameras for weight estimation and describe its validation and evaluation.



OS19-4 Cognition of surrounding conditions for a field robot Slope detection using a multilayer perceptron classifier with point cloud as input -

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

In the Japanese forestry industry, automation of work to supplement labor is desired to achieve sustainable forest management. In this study, the field robot for the automation of forestry is developed. In the field robot, recognition of the surrounding situation is an important function for safe movement. In this paper, we focus on the recognition of terrain. The terrain in a mountainous area has various conditions such as slope, presence of weeds and trees, and unevenness. In this study, the classifier for ground and sloped surfaces using Multi Layered Perceptron (MLP) is developed. This classifier classifies each point of the 3D point cloud acquired from the RGB-D camera into the ground plane of the robot and the slope plane where the robot cannot climb. The accuracy of the classification was verified by training the classifier on a dataset acquired in a real environment.



OS19-5 Particle Filter Based SLAM For Forestry Robot

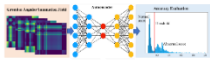
Sylvain Geiser[1], Sakmongkon Chumkamon[1], Ayumu Tominaga[2], Takumi Tomokawa[1], Eiji Hayashi[1]
([1]Kyushu Institute of Technology, Japan)
([2] National Institute of Technology, Kitakyushu College, Japan)

In Japan, the forestry workforce is dramatically declining. Therefore, field robots are investigated to replace humans for dangerous actions. Task execution with such mobile robots requires localization and mapping. This research focuses on online SLAM implemented on SOMA forestry robot developed at Hayashi Laboratory. In this approach, the core algorithm is a Rao-Blackwellized particle filter, and the environment is represented by a map of features which are trees. Furthermore, motion is captured by odometry through rotary encoders and observation is described by a range-bearing model. The raw pointcloud of the mounted lidar is processed in order to get distance and azimuth for each detected landmark. A realistic simulation has been build using Gazebo and the results of first experiments speak for real-time capability.

OS19-6 Anomaly Detection using Autoencoder with Gramian Angular Summation Field in Time Series Data

Umaporn Yokkampon[1], Sakmongkon Chumkamon[1], Abbe Mowshowitz[2], Eiji Hayashi[1]
([1]Kyushu Institute of Technology, Japan), ([2]The City College of New York, USA)

Uncertainty is ubiquitous in data and also represents a real-life challenge. To deal with this challenge, we propose a novel method for detecting anomalies in time series data based on the Autoencoder method, encoding the time series as images by means of Gramian Angular Summation Field (GASF). Time series data is represented as 2D image data to enhance anomaly detection. The proposed method is validated with four time-series data sets. Experimental results show that our proposed method can improve validity and accuracy on all criteria. Therefore, analysis of time series data by combining Autoencoder and Gramian Angular Summation Field methods can effectively detect anomalies.



OS19-7 Autonomous Robot Packaging Ready Meal in Conveyor Production Line

Sakmongkon Chumkamon, Tomofumi Tsuji, Prem Gamolped, Chanapol Piyavichayanon, Umaporn Yokkampon, Eiji Hayashi (Kyushu Institute of Technology)

Food automation technology becomes increasingly important in industrial and scientific research especially in a social distance of pandemics. In this paper, we investigate the robotic motion planning and grasping for the food, which is nonrigid, nonpattern, and soft since the food is difficult to grasp. Moreover, in automation, the process has to be organized on time since the automation uses the conveyor production line. This paper contributes three points. The first is motion planning online while the production line conveyor operating. The second is real-time Non-Pattern Food Segmentation. The third is grasping non-rigid, nonpattern and soft objects of food. Additionally, the robotics grasping also proposes the finishing decorating spaghetti packaging by rolling the spaghetti while the conveyor moving which is our novel proposed.

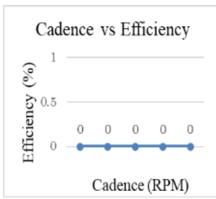
OS20 Application studies on sensor and RFID system (6)

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

OS20-1 A Pedal Powered Water Purifier

Ammar A.M. Al-Talib, Ting Kee Yuan, Sarah 'Atifah Saruchi (UCSI University, Malaysia)

Water is a necessity for every living organism on Earth. However, in certain parts of the world especially in rural areas, clean drinking water is a luxury to the residents. This is because they are not supplied with clean water and thus must source for water from rivers, rainwater, and wells which are often contaminated and unsafe for drinking. The objective of this study is to design and fabricate a functional prototype of a pedal powered water purifier, and to ensure that purified water is safe for consumption based on chemical water analysis. The quality of the purified water will be compared to the Drinking Water Guidelines provided by the World Health Organization (WHO). The system utilizes the distillation method to produce clean drinking water. The charging efficiency of the generator has reached 48.74% at 60rpm cadence. Based on the chemical analysis of purified water, it has been proven that it meets the standards set by WHO and is safe for consumption.



OS20-2 A Smart Node (Maintenance & Lifespan Prediction System)

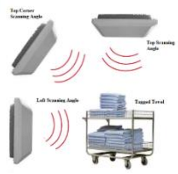
Kam Heng Chaw[1], Ammar A.M. Al-Talib[2], Tarek Fawzi[2], Jonathan Yong Chung Ee[2]
([1]MODU System (S) Pte Ltd, Malaysia), ([2]UCSI University, Malaysia)

Since the present business environment is suffering uncertainty due to the economic fluctuations, not only the failure anticipation of production lines is fundamental, but also it is critical to preserve company's capital and production flowing. Failure doesn't occur overnight, as the warning signal from many different sources emerge, and even the production quality/quantity evolve prior to the failure. Hence, surveillance of these signals 'sources could be used as an input for a system that depends on smart factory principles to predict failure and parts' lifespan in advance. This idea was echoing for few years but now it became achievable due to the development of the artificial intelligence (AI), machine learning, data mining and data reservoir technologies. The project aimed to develop novel maintenance mechanism based on industry 4.0 principles, implementing of the IoT monitoring system on manufacturing line and verifying and validating the performance of the maintenance system.

OS20-3 A Healthcare Laundry Management System using RFID System

Eryana Hussin, Wong Jie Jian (UCSI University, Malaysia)

The paper presented the use of Ultra-High Frequency (UHF) Radio Frequency Identification System (RFID) for laundry management system in healthcare. The laundry management system consists of the waterproof RFID laundry tags, RFID reader and the data recording system which was build using Visual Basic and Microsoft Access software. In this project, the users will be able to monitor the process of the hospital's linen which includes the pillowcases, bedsheets, towels, and other hospital linens. The advantage of using this laundry management can reduce the human error of missed count the items, and to ensure smooth laundry item cleaning process to be executed. Moreover, the user will be able to monitor the status of certain linen items. The database recoding system presented in this paper can reduce the human error, effective and systematic compared to the traditional laundry counting and recording system. This laundry management system also can be used for hotels laundry management for the linen that can be tagged to prevent the loss of items such as towels, blanket, and curtains.



OS20-4 A Monitoring System with Humidity and Growth Level Detection for Horticulture

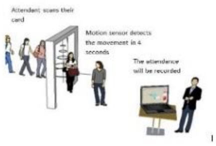
Eryana Hussin, Ng Joon Wen (UCSI University, Malaysia)

Smart Farming System with humidity and growth level detection presented in this paper is designed to monitor and control the production of the fresh produce. This smart farming system is designed to monitor the changes due to environment parameters which includes temperature, soil moisture level of the crops, and growth level of the crops. A sensor is placed at certain distance from the plant and the growth rate of the plant is monitored from time to time. In this prototype, the 12V fan is connected to the relay and will be turn on when the temperature drops below the minimum accepted range for the produce. The Smart Farming System proposed in this paper is built with a combination management system including environment parameters monitoring system and automatic technology control system. Arduino UNO is the central unit that plays an important role to receive and control all the function of the project. This combined system is design for modern smart horticultural or large-scale horticultural to produce more sustainable production.

OS20-5 A Double Identification Attendance System Using High Frequency RFID System

Eryana Hussin, Wong Chee Ming (UCSI University, Malaysia)

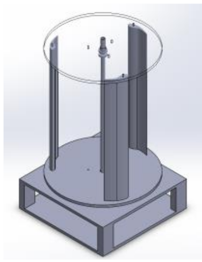
In this paper, the double identification attendance system integrated with High Frequency Radio Frequency Identification (RFID) and sensors is presented. This double identification system requires users to be detected by two different mechanisms. This double identification system is proposed to overcome the problem of false attendance by the students who just stopped by to tag their student identification card without attending the class. The attendance system presented in this paper can capture the time in and time out of each student which will be automatically captured if the students successfully pass through both RFID reader and motion sensor. The proposed attendance system is designed to replace the manual attendance system which still implemented in this modern era. The unique features of the RFID system which able to capture and store the student attendance in the database system effectively and can reduce the time for the teachers to take the attendance daily. This double identification attendance system also will be useful for any event or training provider to record the attendance of the attendant effectively.



OS20-6 Levitating Frictionless-Vertical Windmill-

Ammar A.M. Al-Talib, Muhammad R. Md. Redzuan, A. R. Abd Hamid (UCSI University, Malaysia)

To harness wind energy in low wind speed areas, a vertical axis wind turbine (VAWT) is generally more suitable than a horizontal axis wind turbine. To improve the feasibility of using a VAWT, a magnetic levitation bearing concept, utilizing neodymium magnets, is used to make the bearing frictionless. A new experimental type of airfoil called EN0005 is used that boasts better self-starting ability than conventional airfoils. The design parameters of the wind turbine are explored and reviewed to create a final design that is modelled in SolidWorks. This final design is then put through simulations and analysis in ANSYS FLUENT before being fabricated as a prototype. This final prototype produces similar results to other findings and validates the report. The EN0005 blade profile appears to improve the self-starting capability of the wind turbine and improve viability of Darrieus turbines in lower wind speeds. These findings verify the benefits of using a magnetic frictionless bearings for wind turbines.



OS21 Visual Signal Processing and Human-welfare Robotics I&II (7)

Chair Joo Kooi Tan (Kyushu Institute of Technology, Japan)
Co-Chair Nobuo Sakai (Kyushu Institute of Technology, Japan)

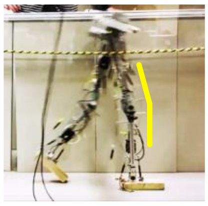
OS21-1 Object Acquisition Based on Human-Robot Cooperation

Kota Ito, Masuhiro Nitta, Seiji Ishikawa, Joo Kooi Tan (Kyushu Institute of Technology, Japan)

In recent years, the population of elderly persons who are unable to go for shopping to buy their daily necessities is increasing. Putting an intelligent shopping robot as a future goal, in this paper, we propose a human-robot cooperative system in which a developed robot goes to a spot where a user specifies, acquires goods which the user wants and brings the goods back to the user. To realize this, real-time user-robot communication, processing of the two kinds of images provided from an RGB-D camera, and robot manipulation and travel control are done. The effectiveness of the proposed method was verified by experiments. The proposed robot system will provide a support system for shopping refugee, such as an elderly, a bedridden or a physically-disadvantage person, in future based on online real-time goods selection.

OS21-2 Development of Musculoskeletal Walking Simulator for Analysis of Human Walking and Rehabilitation

Nobuo Sakai, Yukiho Ryu, Tsubasa Ikeda, Mochimitsu Komori (Kyushu Institute of Technology), Masako Fuchi (Kyushu Nutrition Welfare University), Katsuki Hayashi (Seiai Rehabilitation Hospital)



In the field of rehabilitation, the explanation of walking motion, called ‘Rocker function’, is one of the practical indexes for the facilitation of human walking in clinical settings. If this explanation is true, we can reconstruct it by artificial materials. In this study, we tried to develop the musculoskeletal walking simulator, which can actually reproduce bipedal walking according to the Rocker function. Muscles and tendons including biarticular arrangements were represented by springs and cables. In this report, 3 muscles were actuated by servomotor. The simulator reproduced the human musculoskeletal walking motion generated from its intellectual structure in nature. The results would support the insight of the explanation of Rocker function in the rehabilitative treatments.

OS21-3 Collision Avoidance in a Human-Robot Coexistence Food Preparation Environment Using Hands Area Extraction

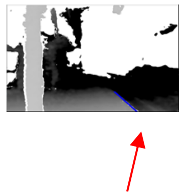
Takaaki Yotsumoto, Yuta Ono, Masuhiro Nitta, Joo Kooi Tan (Kyushu Institute of Technology, Japan)

In Japan, the population of the working-age between 15 and 64 years old has been declining. In order to solve the labor shortage, the introduction of industrial robots that can perform tasks equivalent to humans especially in a food preparation industry is strongly requested. In a human-robot cooperative environment, it is indispensable to avoid collision between workers and robot. In this paper, we propose a method of automatic hand area extraction of a robot and a worker in a food preparation work line based on an ego-camera and the use of color distribution and GrabCut. on the images provided from the camera. A warning is issued, if the hands of the side workers/robots approach close to the central work area. Experimental results show satisfactory results. The method is under refinement so that it will adapt to various hand colors and illumination change.

OS21-4 Supporting Safe Walk at a Railway Station Platform for a Visually Impaired Person Based on MY VISION

Yuki Kawaguchi, Seiji Ishikawa, Takashi Shinomiya*, Joo Kooi Tan
(Kyushu Institute of Technology, Japan, *N&N Inc.)

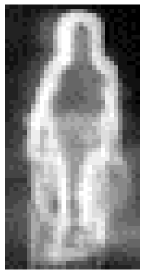
The platforms of a railway station is a dangerous place for visually impaired person. However, not all stations, except those in large cities, are equipped with platform doors. Therefore, a supporting system is requested for safe walk of a visually impaired at a station platform. In this paper, we propose a system that prevent a visually impaired person from falling on the track from the edge of a platform using MY VISION. The proposed system employs a depth image captured by MY VISION. It uses Line Segment Detector and Graph-Based Image Segmentation to detect the edge of the station platform. If the users of MY VISION approach the edge too close, the system gives a warning to the user. The effectiveness of the proposed method was verified by experiments.



OS21-5 Detecting a Pedestrian's Walk Direction Using MY VISION for Supporting Safe Walk of a Visually Impaired Person

Shinya Iizumi, Yuta Ono, Masuhiro Nitta, Seiji Ishikawa, Joo Kooi Tan
(Kyushu Institute of Technology, Japan)

As a means of assisting a visually impaired person to walk safely, a system has been proposed that recognizes surrounding pedestrians and their approaching directions using the images obtained by MY VISION. However, since the conventional method used a pedestrian's model to generate MSC-HOG (Multiple-Scale-Cell Histograms of Oriented Gradients) features, it is specialized only for recognizing pedestrians. Thus it is difficult to recognize other passers-by such as cyclists. To solve this problem, we propose a method to recognize not only pedestrians, but also cyclists using the average edge images of pedestrians and cyclists based on MSC-HOG. We also use different discriminators for robustly detecting passers-by who approach a MY VISION user. The effectiveness of the proposed method was verified by experiments.



OS21-6 Fruits and Vegetables detection using the improved YOLOv3

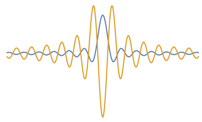
Changhua Xu, Ziyue Liu, Masuhiro Nitta, Joo Kooi Tan (Kyushu Institute of Technology, Japan)

As the global aging intensifies, it is more convenient for a robot go for shopping to buy things like vegetables or fruits instead of elderly persons themselves. It is then important that a robot is more human-like to select items according to a user's personal preferences such as maturity of fruits, sweetness, freshness and so on. However, fruits or vegetables are generally displayed in a disorderly manner. Therefore, thorough detection and recognition of fruits and vegetables is a difficult task for a robot. This paper proposes an improved YOLOv3 and also pre-training the networks to detect fruits and vegetables and to recognize their maturity. The effectiveness of the proposed method is shown by experiments.

OS21-7 Strict frequency estimation of sinusoidal signal using sampling function(withdraw)

Masuhiro Nitta (Kyushu Institute of Technology, Japan)

This paper considers an accurate frequency estimation problem of a sinusoidal signal acquired by an analog-to-digital converter. By making use of an A/D converter, pure sinusoidal signal is digitized and quantized error occurs. Thus the estimated frequency slightly differs from the original one. The aim of this study is to identify the original frequency based on the theory of distributions. As the continuous test function with compact support, the paper utilizes a sampling function. Although the proposed estimation method requires the derivative of the sinusoidal signal, sampled sinusoidal signal becomes a nondifferentiable function. The distributions overcome this difficulty by using integration by parts of the sinusoidal signal and the sampling function. The effectiveness of the proposed method is demonstrated by some numerical simulations.



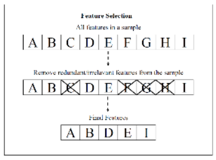
OS22 Advanced studies of network engineering (4)

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

OS22-1 New Hybridization Algorithm of Differential Evolution and Particle Swarm Optimization for Efficient Feature Selection

Koon Meng Ang, Mohd Rizon Bin Mohamed Juhari, Wei Hong Lim, Sew Sun Tiang, Chun Kit Ang, Eryana Eiyda Hussin, Li Pan, Ting Hui Chong (UCSI University, Malaysia)

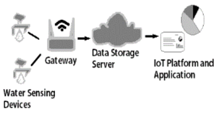
Feature selection is a popular pre-processing technique applied to enhance the learning performances of machine learning models by removing irrelevant features without compromising their accuracies. The rapid growth of input features in big data era has increased the complexities of feature selection problems tremendously. Given their excellent global search ability, differential evolution (DE) and particle swarm optimization (PSO) are considered as the promising techniques used to solve feature selection problems. In this paper, a new hybrid algorithm is proposed to solve feature selection problems more effectively by leveraging the strengths of both DE and PSO. The proposed feature selection algorithm is reported to achieve an average accuracy of 90.54% when solving 13 datasets obtained from UCI Machine Learning Repository.



OS22-2 Implementation of LoRa in River Water Quality Monitoring

Syarifah Nabilah Syed Taha Tahir, Mohamad Sofian Abu Talip, Mahazani Mohamad, Mohamadariff Othman, Tengku Faiz Tengku Mohmed Noor Izam, Mohd Faiz Mohd Salleh, Zati Hakim Azizul Hasan, Zeeda Fatimah Mohamad, Amir Feisal Merican Aljunid Merican (Universiti Malaya, Malaysia)

Emergence of Long Range (LoRa) in network technologies become game changer for Internet of Things (IoT) application. Deployment of LoRa enable IoT application of environment monitoring to cover wide area while maintain at low energy and low cost. Water quality monitoring program was developed to maintain and protect quality of water resources for daily purpose. Also, to prevent pollution and disease epidemic peculiarly during Covid19. This research aimed to build autonomous water quality monitoring prototype implemented with LoRa network for support decision system. The Wireless Sensor Nodes (WSN) that embedded with five type of water quality sensors of pH, turbidity, total dissolved solid (TDS), dissolved oxygen (DO) and temperature linked to single gateway. Water environmentalist able to view the result of timely water quality from mobile application dashboard. Though the performance not severely affected, acquired results revealed non-line of sight condition, transmission power and Spread Factor (SF) value influenced LoRa performance in urban environment. In conclusion, a few improvements on the system grant LoRa high capabilities to be integrated with IoT environment application in urban environment.



OS22-3 Wideband Antenna with UHF Sensor Applicability for MV/HV Equipment in Smart-Grid Systems

S. M. Kayser Azam[1], Mohamadariff Bin Othman[1], Tarik Abdul Latef[1], H. A. Illias[1], Mohd Fadzil Ain[2], Yazeed Qasaymeh[3] ([1]Universiti Malaya, Malaysia) ([2]Universiti Sains Malaysia, Malaysia) ([3]Majmaah University, Saudi Arabia)

Abstract—In this paper, a wideband antenna is proposed to be used as an ultra-high frequency (UHF) sensor for medium to high voltage equipment for the next generation smart-grid systems. The antenna is designed on a heat-protected substrate so that it can withstand the extreme environment of a power substation. The proposed antenna operates within the UHF range with wideband characteristics. The peak and average values of the realized gain indicate that the antenna has the ability to largely improve the received signal before delivering it for data analysis and processing. The antenna radiates with an omni-directional pattern in the three-dimensional view. Efficiency of the proposed antenna is quite decent, and the physical dimension of the antenna is compact enough to be used as an UHF sensor. The proposed wideband antenna as an UHF sensor is designed to be aimed especially at the faulty-insulation detection in medium to high voltage equipment like power cables, power transformers, gas-insulated switchgears etc. while adopting the smart-grid technology.

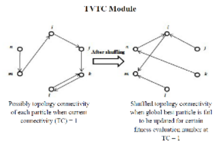
$$??_{??} = \frac{7.48 \times h}{??^{(??_0 \frac{\sqrt{??_{??}+1.41}}{87})}} - 1.25 \times ??$$

$??_{??}$ = feedline width
$??_{??}$ = feedline length
$??_0$ = characteristic impedance
$??_{??}$ = relative dielectric permitivity
h = substrate height
?? = conductive trace thickness

OS22-4 New Particle Swarm Optimization Variant with Modified Neighborhood Structure

Koon Meng Ang[1], Mohd Rizon Bin Mohamed Juhari[1], Wy-Liang Cheng[1], Wei Hong Lim[1*], Sew Sun Tiang[1], Chin Hong Wong[2], Hameedur Rahman[3], Li Pan[1]
([1]UCSI University, Malaysia) ([2]Fuzhou University, China) ([3]Air University, Islamabad, Pakistan)

Numerous particle swarm optimization (PSO) variants were proposed in past decades to tackle different types optimization problems more robustly. Nevertheless, the imbalance of explorative and exploitative search behaviors remains as an on-going research challenge that can restrict the performance of PSO. In this paper, a new variant known as PSO with time-varying topology connectivity (PSO-TVTC) is proposed. A time-varying topology connectivity (TVTC) module is designed to achieve the proper regulation on explorative and exploitive behaviors of PSO via dynamic modification of particle's topology connectivity throughout the optimization process. Experimental results reveal that the proposed PSO-TVTC has exhibited prominent performance among its competitors by producing 7 best mean fitness out of 8 benchmark functions.



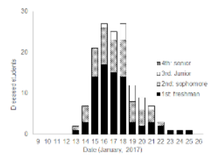
OS23 Applications in Complex Systems (7)

Chair Masao Kubo (National Defense Academy, Japan)
Co-Chair Hiroshi Sato (National Defense Academy, Japan)

OS23-1 A research of infectivity rate After the Consecutive Holidays

Saori Iwanaga (Japan Coast Guard Academy, Japan)

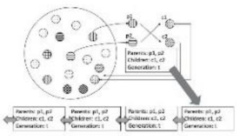
I found there are super-spreaders in the Japan Coast Guard Academy case, but super-spreading depends on the timing of bringing in. After three consecutive holidays from January 7 to January 9, 2017, students returned dormitory in JCGA and started to take classes. Then, 25 % of students were infected in two weeks. We had introduced a refinement to the SEIR model to previously infectious "P" state in the incubation period and proposed a discrete mathematical SEPIR model for influenza. By examining infection from the students of previously infectious "P" state, I found that there are super-spreaders who directly infected over 10 students in this case. But, super-spreading doesn't depend on the features of a person, it depends on the timing of bringing in. I found that students can super-spread seasonal influenza until the first infected students are found. After that, because most students take measures of seasonal influenza, the number of patients decreases.



OS23-2 Towards the Trusted Population-Based Optimization Systems

Hiroshi Sato, Masao Kubo (National Defense Academy, Japan),

Following the development of evolutionary computation, various population-based optimization methods have been proposed. In these systems, optimization is achieved through the interactions of many individuals/particles/agents. However, when the system is implemented in a distributed environment, reliability becomes an issue. In such an environment, it may not be possible to trust others. There are numerous cases why we cannot guarantee trust, such as malfunction of distributed parts or failure to synchronize. Therefore, we have to make trust between distributed individuals/particles/agents. The record of past actions is usually a good tool for generating trust. This paper introduces the blockchain mechanism into the population-based optimization system to make a trust management system. By using blockchain, we can implement it without a central authority. In the system, all interactions are reviewed and get feedback, and the feedback is used to calculate the trust score. We consider several scoring methods for this type of system.



OS23-3 Spatio-temporal prediction of crime occurrence spots by CNN-LSTM

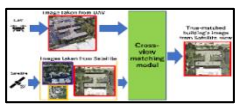
Kaede Yaji, Masao Kubo, Hiroshi Sato (National Defense Academy, Japan)

This paper proposes a method for spatiotemporal prediction of crime occurrence locations based on previous data. In recent years, Japanese government has begun to release data on crime occurrences to improve the efficiency of policing. In addition, the development of maps that can manage patrol and assist residents' crime prevention has been planned. For statistical crime prediction, while several methods are invented abroad, it has just begun to develop a specific crime prediction model for a low-crime country, Japan. One of the known methods uses LSTM to predict crime occurrences only from a temporal perspective, but it cannot predict points of crime occurrences and is insufficient to generate a map. Therefore, we propose a method that combines this LSTM based method with CNN that can adopt geographic locations. As a result of computer experiments, this method seems to be able to make predictions with a tendency to capture actual characteristics.

OS23-4 Cross-View Image Geo-Localization using Multi-Scale Generalized Pooling with Attention Mechanism

Duc Viet Bui, Masao Kubo, Hiroshi Sato (National Defense Academy, Japan)

Cross-view image matching for geo-localization is the task of finding images containing the same geographic target across different platforms. This task has drawn significant attention due to its vast applications in UAV's self-localization and navigation. Given a query image from UAV-view, a matching model can find the same geo-referenced satellite image from the database, which can be used later to precisely locate the UAV's current position. Many studies have achieved high accuracy on existing datasets, but they can be further improved by combining different feature processing methods. Inspired by previous studies, in this paper, we proposed a new strategy by using a channel-based attention mechanism with a generalized mean pooling method to enhance the feature extracting process, which improved accuracy.



OS23-5 Recommendation a Emergency Patient Destinations by LightGBM

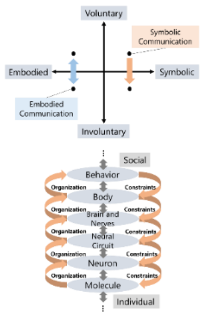
Ryota Kawaguchi[1], Masao Kubo[1], Hiroshi Sato[1], Daizoh Saitoh[2], Takao Oya[3]
([1]National Defense Academy, [2]National Defense Medical College, [3]Saitamaseibu Fire Bureau. Japan)

In this study, we aim to reduce the burden of selecting a destination hospital for a rescue team. We propose to adopt the LightGBM method to recommend a destination hospital based on information such as a conversation between the patient and a call center. Previous studies have used detailed information such as patients' disease histories, whereas we use only simple information. Therefore, it can be implemented in regional medical organizations that do not have enough medical information systems. For our method, we used Doc2vec to convert the text of the conversation to a vector. As a result, the LightGBM method has an accuracy score of 70%. It is higher than the k-nearest neighbor, logistic regression, and neural network. In addition, by analyzing the result of our method, we found that age, injury or illness level, and location information are important factors for improving accuracy scores.

OS23-6 A Framework for Understanding the Neural Underpinnings of Symbolic and Non-Symbolic Communication Based on Global Synchronization in Human Brain Activity

Masayuki Fujiwara, Takashi Hashimoto (Japan Advanced Institute of Science and Technology, Japan)

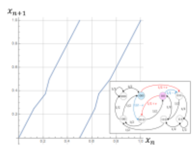
We propose a framework for understanding the neural underpinning of communication with electroencephalogram (EEG) synchronization. It consists of four stages: 1) two-dimensional space defined by symbolic/embodied (non-symbolic) vs. voluntary/involuntary to characterize the target communication, 2) ontological hierarchy to focus the level of synchronization analysis, 3) neurocognitive modeling to hypothesize neural mechanism, and 4) model-based EEG neurofeedback to empirically validate the hypothesis. We claim that following the framework makes it possible to advance our understanding of neural dynamics and mechanisms for communication. Moreover, we analyzed two EEG experiments, implementing the former two stages: the formation of symbolic communication changing from voluntary to involuntary and embodied communication competing between voluntary and involuntary. Their outcome is a hypothesis that three different brain regions are involved in interpreting symbols, motor intentions, and social coordination. A neural field model and a manipulation technique with EEG-based connectivity neurofeedback are also proposed for the latter stages.



OS23-7 Characterization of randomness tests by using tests results of weakly correlated chaotic sequences

Akihiro Yamaguchi (Fukuoka Institute of Technology), Asaki Saito (Future University Hakodate, Japan)

High quality pseudo-random number sequences are required in various fields of engineering, and the statistical test of randomness is one of the important subjects. A typical test suite of randomness, e.g., DIEHARD, NIST SP800-22, and TestU01, is defined as a set of several different kinds of random number tests. One problem here is that the similarity between the individual tests included in the test suite is not obvious, and it is difficult to make an argument for the optimality of a set of randomness tests. In this study, we propose a characterization method based on the test results of weakly correlated binary sequences generated by the piecewise linear chaotic map. Then, we characterize randomness tests included in the test suite of NIST SP800-22 and discuss the characterization performance of our proposed method.



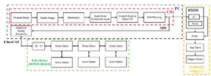
OS24 Artificial Systems and Life (5)

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

OS24-1 An EtherCAT Based Delta Robot Synchronous Control Application

Chung-Wen Hung[1], Yu-Hsuan Tseng[1], Chau-Chung Song[2], Guan-Yu Jiang[1] ([1]National Yunlin-University of Science and Technology, Taiwan, [2]National Formosa University, Taiwan)

The delta robot synchronous control based on the Ethernet Control Automation Technology (EtherCAT) protocol is proposed in this paper. Personal Computer (PC) is used as master and the delta robot motor drivers are used as slaves in this work. The Master sends command to slave base on the motion control profile CAN in Automation 402(CiA402). Subsequently, the program in C# perform the user's interface and EtherCAT communication. And the system is not only easy use, but also quickly high-precision. A complex painting application is proposed to show this system workable.



OS24-2 Web-based SCADA using MQTT Protocol and AES

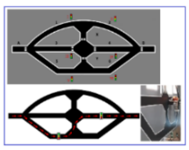
Jr-Hung Guo, Tzu-Yuan Lin, Kuo-Hsien Hsia
(National Yunlin-University of Science and Technology, Taiwan)

Internet of Thing (IoT) technology is a very popular research topic. Especially in the application of Industry 4.0, it is the most basic and important part. However, related security and application system development often have many problems. Therefore, this paper uses MQTT protocol and AES encryption technology to develop a Web-based SCADA system. This Web-based SCADA uses drag-and-drop operation, and users can quickly build a WYSIWYG (What You See Is What You Get) application system. And by collecting different communication protocols and integrating multiple communication interfaces, this SCADA can be connected to many PLCs and other equipment. Allow the industry to build application systems quickly and at low cost.

OS24-3 Smart Identification System of Teaching-type Autonomous Vehicles

Chun-Chieh Wang (Chienkuo Technology University, Taiwan)

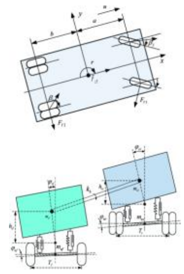
Today's image recognition technology has been used in many engineering fields. Especially the image recognition combined with the automatic driving system can bring greater traffic convenience to people. To improve the teaching efficiency of autonomous driving image recognition technology, a self-driving car intelligent identification system suitable for teaching has been developed in this article, so that teachers and students can easily use this system for experiments. As for the main controller of the car body is the Raspberry Pi microcomputer processor. It works with Python for image processing. Image processing techniques include grayscale, binarization, morphology, image cutting, etc. In addition, in order to facilitate teaching, the experimental road field is planned to include 9 paths for self-driving cars to drive autonomously, in this article. To realize the function of self-driving cars, there are four main functional tests in the context setting. It includes road identification, conversion of lane turning arc into front wheel turning angle, intersection identification, and traffic light identification. The experimental results confirm that the developed smart identification system and the experimental environment planning are helpful for autonomous driving related teaching and can enhance students' willingness to learn.



OS24-4 Automatic Anti-Lock Brake System for Anti-Rollover Control of Autonomous Heavy-Duty Truck

Chian C. Ho, Riki Umami Sanaz Ulfitria (National Yunlin-University of Science and Technology, Taiwan)

In recent years, there are more and more rollover accidents about autonomous heavy-duty trucks or autonomous ordinary vehicles in intelligent airports or seaports. These accidents leads to the hot research field about the prevention of rollovers in advance for autonomous vehicles, especially in autonomous heavy-duty trucks. This paper develops an automatic anti-lock brake system (ABS) as one way to stabilize the autonomous vehicle. Then, through monitoring both vehicle's and wheel's speeds, this paper helps the autonomous vehicle keep stability control with automatic ABS even when the wheels halt or slip on the road. This paper adopts TruckSim to model the vehicle safety dynamics and MATLAB/Simulink to simulate the vehicle stability control. Experimental results show that the elaborate automatic ABS proposed by this paper can smoothly keep the vehicle stable even under dangerous road conditions of sharp corner or hairpin turn.



OS24-5 Development of Intelligent Beehive and Network Monitoring System for Bee Ecology

Chau-Chung Song[1], Geng-Yi Lin[1], Chi-Chung Peng[2] and Chung-Wen Hung[3] ([1, 2]National Formosa University, Taiwan, [3]National Yunlin-University of Science and Technology, Taiwan)

In this paper, development of intelligent beehive and network monitoring system for bee ecology is focused on data acquis such as temperature, humidity, weight, and GPS positioning, combining the beehive with MCU, sensors, and ZigBee to implement front end sensing nodes to build a bee ecological network monitoring information system for real-time remote network monitoring, and assist beekeepers to establish a cloud-based real-time monitoring system and history traceability Bee product management system to enhance the convenience of beekeeper management and risk control, thereby effectively improving the efficiency of labor utilization. Also, clear production history information can increase consumers' trust in products and enhance the overall bee-related industry Economic benefits. This paper cooperated with beekeepers in Gukeng, Yunlin County, set up an Intelligent beehive system in the bee farm. Observations in the past month have shown that the activity and number of adult bees and larvae grown in the intelligent beehive are in good condition, user can also connect to the Intelligent beehive monitoring website through their mobile phone, tablet, or computer to analyze and monitor the status of each beehive.

OS25 Information Applications and Cybersecurity (6)

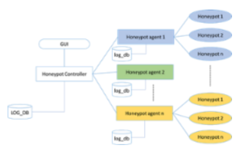
Chair I-Hsien Liu (National Cheng Kung University, Taiwan)

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

OS25-1 Extendable ICS Honeypot Design with Modbus/TCP

I-Hsien Liu, Jun-Hao Lin, Hsin-Yu Lai, Jung-Shian Li (National Cheng Kung University, Taiwan)

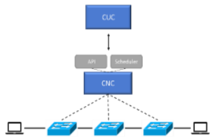
In order to protect the Cybersecurity of Industrial control system (ICS), we design a prototype of an ICS honeypot. All honeypots are controlled by a server, and using the description file to define honeypot's characteristics, to achieve our honeypot system with scalability and high interaction. We compare our honeypot system and Conpot. the results show that the responses of our honeypot system have more interaction. Even more, our honeypot obtained a perfect score in the honeypot scoring mechanism of Shodan.



OS25-2 Industrial Control System Cybersecurity Testbed with TSN Feature

I-Hsien Liu, Li-Yin Chang, Jung-Shian Li (National Cheng Kung University, Taiwan)
Chuan-Gang Liu (Chia Nan University of Pharmacy & Science, Taiwan)

Time sensitive networking (TSN) is the future industrial network trend, under the development of the future Industry 4.0. and a large number of Internet of Things. How to ensure the transmission delay and jitter of the packet will be a major issue. A key feature of TSN is the traffic scheduling mechanism,which can accommodate hard real time streams of critical data with bounded end to end delays. In this work we have setup a TSN testbed to implement TSN. IEEE 802.1 TSN standard includes precise clock to sync, bandwidth reservation, and traffic shaping, which provide high reliability, low latency and other industrial needs. IEEE 802.1Qbv standards use some mechanisms to handle scheduled traffic jitter and critical traffic latency. We focus on testing and implementing 802.1Qbv standard.



OS25-3 Using the Modified Delphi Method to Construct the Quality Indicators of the Counseling Service System

Li-Min Chuang, Hsiu-Hao Liu (Chang Jung Christian University, Taiwan)

The counseling service system is a case management tool used by professional guidance counselors and full-time teacher-counselor. However, the system satisfaction and willingness to use were generally low. This study used the modified Delphi method to construct the quality indicators of the counseling service system. The study found the quality indicators of the system can be divided into five major dimensions and 23 evaluation indicators to explore. There were three indicators most valued by experts and users, which were the functions are simple and convenient to operate, the project meets the needs of the work, and the efficiency and convenience of paper processing are improved. Based on the research results, this study puts forward substantive implications for academics and management.

OS25-4 The Key Success Factors of Introducing ERP System in Taiwan's Manufacturing Industry

Li-Min Chuang, Yu-Po Lee (Chang Jung Christian University, Taiwan)

This study examines the key success factors of ERP introduction in Taiwan’s manufacturing industry as a reference for enterprises to reduce the high and unforeseeable financial and time costs of ERP introduction. In the second phase, a questionnaire study was conducted using the Analytic Hierarchy Process to extract the relative weights of the distance between primary and secondary dimensions, and 5 primary dimensions were derived, including “Management/Organization”, “Introduction Process”, “Technical Support”, “Documentation”, and “Personnel”, as well as 15 secondary dimensions.

Dimension	Weight	Rank
A Management/ Organization	0. 1961	3
B Introduction Process	0. 2203	2
C Technical Support	0. 1206	4
D Documentation	0. 0413	5
E Personnel	0. 4207	1

OS25-5 The Fuzzy AHP approach for intelligent building assessment model

Li-Min Chuang, Yu-Po Lee (Chang Jung Christian University, Taiwan),
Chien-Chih Kuo (Chien Chang Construction Co., Ltd.)

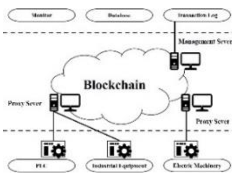
The main objective of this thesis is to probe into how Taiwanese building investment and development companies rate the analytical framework and weights of artificial intelligence buildings. Document Analysis, the Delphi method, and the Fuzzy Analytic Hierarchy Process (FAHP) are applied to conduct a FAHP questionnaire survey among 20 building investment and development companies in Tainan. Based on the calculation of composite weights, the findings are: (1) The most crucial evaluation indicator for security and hazard prevention is the "access control system". (2) The most crucial evaluation indicator for energy-saving management is "energy-saving technology". (3) The most crucial evaluation indicator for health and comfort is the "interior comfort system". (4) The most crucial evaluation indicator for intelligent innovation is the "intelligent innovation concept".

Dimension	Weight
Security & hazard prevention	0.375
Energy-saving management	0.303
Health & comfort	0.164
Intelligent innovation	0.158

OS25-6 Blockchain-based Verification Mechanism for Industrial Control System

Yao-Chu Tsai, I-Hsien Liu and Jung-Shian Li (National Cheng Kung University, Taiwan)

Industrial control systems (ICS) and critical infrastructure have become increasingly dependent on communication networks and cyber-physical systems. Since infrastructure is vulnerable to natural disasters, physical destructions, and adversarial attacks, the research on cybersecurity is vital in Industry 4.0. In order to secure the integrity of data in ICS, this paper proposes a blockchain-based network architecture implemented on physical industrial equipment. By the arrangement of the blockchain transaction process in the specialized client-server network model, industrial control signal transmission can be verified based on authority. The transaction logs would not be easily tampered with due to the characteristics of blockchain, and the data integrity could be assured.



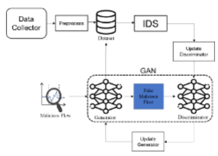
OS26 Intelligent Life and Data Analysis (6)

Chair I-Hsien Liu (National Cheng Kung University, Taiwan)
Co-*C*hair Chu-Fen Li ((National Formosa University, Taiwan)

OS26-1 Data Balanced Algorithm Based on Generative Adversarial Network

I-Hsien Liu, Cheng-En Hsieh, Wei-Min Lin, Jung-Shian Li (National Cheng Kung University, Taiwan), Chu-Fen Li (National Formosa University, Taiwan)

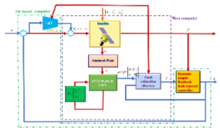
In order to defend against malicious attacks, intrusion detection systems have introduced machine learning as a protection strategy. However, machine learning algorithms and datasets have a great influence on the effectiveness of the machine learning model. This study uses five algorithms which are Naïve Bayes, CNN, LSTM, BAT, and SVM to train the IDS machine learning model. We use three datasets which are NSL-KDD, UNSW-NB15, and CICIDS 2017 to train and evaluate the model performance. We design a data-balanced method based on the GAN algorithm to improve the data imbalance problem of the IDS dataset. Also, we use the method to generate labels, these labels are used to explain the clustering effect of the unsupervised model.



OS26-2 Fault-Tolerant Control System Design for Nonlinear System with Actuator Faults

Ho-Nien Shou (Air Force Institute of Technology, Taiwan),
Hsin-Yu Lai (National Cheng Kung University, Taiwan)

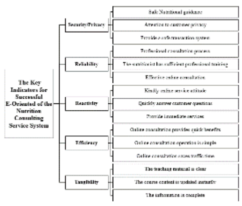
This article deals with observer-based integrated robust fault estimation and accommodation design problems for nonlinear system. The environmental disturbance torque, actuator faults, sensor faults and model uncertainties are considered. Firstly, we propose the augmented fault estimation observer (AFEO) to guarantee the convergence of $H\infty$ performance index of fault estimation and to restrict the influence of uncertainties with respect to the fault estimation error as well. We then design the fault accommodation which is based on the dynamic output feedback to keep the stability of the closed-loop system while malfunctioning. AFEO and dynamic output feedback fault tolerant controller (DOFFTC) are separately designed. Their performances are separately considered. Finally, we propose a simulation result of the micro-satellite attitude control system to demonstrate the effectiveness of the presenting method.



OS26-3 Key Indicators for Successful E-Oriented Operation and Management of the Nutrition Consulting Service System

Ling-Mei Hsu (Chang Jung Christian University, Taiwan)

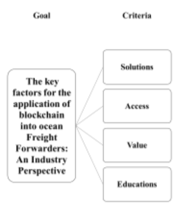
According to the application theory of information systems, it is a tendency that the e-management of the nutrition consulting service industry by the Internet generation, and service quality is the key indicator of operational performance, which is of great importance. This research purpose was to analyze the key dimensions of online service quality when customers were learning online. This research used the E-SERVQUAL to design five major dimensions and 15 criteria, and through the analytic hierarchy process (AHP) constructed a set of successful operation management modules. This research verified the order of the key indicators were: Security/Privacy, Reliability, Reactivity, Efficiency, Tangibility.



OS26-4 The key factors for the application of blockchain into ocean Freight Forwarders: An Industry Perspective

Chu-Ting Hsu, Ming-Tao Chou, Ji-Feng Ding (Chang Jung Christian University, Taiwan)

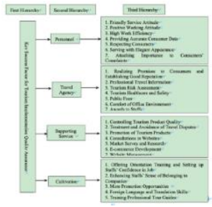
Blockchain is an emergent technology concept that enables the decentralized and immutable storage of verified data and is often considered to be applied to help the maritime industry to manage innovation. This study subdivided the four dimensions into 20 appropriate evaluation indicators. The indicators of key factors based on the results are as follows: 1. ”The application of blockchain will reduce intermediary costs and increase revenue when switching transportation vehicles.” 2. ”Using blockchain will save a lot of manpower and correspondence in the traditional model, and build trust, and fully secure information. ”Ocean freight forwarders can explore the standard service model as a reference to build competitive advantages and ensure sustainable business decisions.



OS26-5 Key Success Factors for Implementation Quality Assurance of Information Technology in Tourism Industry

Shuen-Huei Yao, (Chang Jung Christian University, Taiwan), Cheng Chung Yeh (National University of Tainan, Taiwan), Wen Jung Tsai (Chang Jung Christian University, Taiwan)

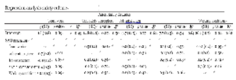
With the constant change nowadays, various enterprises are facing such a fierce competition that they have to seek for competitive advantages and core competitiveness. With respect to consumers, the implementation quality assurance based on the technology system platform in tourism industry. In order to figure out for implementation quality assurance of information technology in tourism, the Delphi method is applied in this research. For the weights and priorities of various measurement indicators are compared through AHP. According to the results, “training professional tour guides”, “foreign language and translation skills” are the most critical success factors for the implementation quality in tourism industry. Besides, it hoped results can provide suggestions for the actual implementation and management of the platform.



OS26-6 AI Big data analysis and application: Patient Safety Culture of Nursing Staff in an Operation Room

Su-Chiu Hsiao (Chang Jung Christian University, Taiwan)

The objectives of this study were to investigate patient safety attitudes among operation Room (OR) nursing staff by using the Safety Attitudes Questionnaire (SAQ) and to compare their safety attitudes with the entire hospital nurses'. This study investigated the factors affecting patient safety attitudes by the multivariate regression models. There was a significant association between resilience and safety climate after adjusting for age, gender, and career experience using the multiple-regression model. Perception of management was also significantly associated with job satisfaction. The survey data provide baseline information on OR nurses. The results can be used to follow-up on the effectiveness of quality improvement campaigns, caring Strategy, and patient safety education in the future. Establish a caring notification system. Achieve effective management.



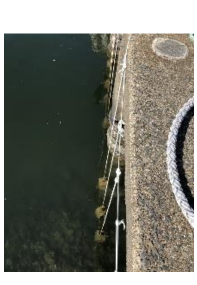
OS27 Environmental Monitoring (5)

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

OS27-1 Biofouling Monitoring Experiments of Underwater Concrete Samples for Offshore Platform Cleaning Robot Development

Keisuke Watanabe[1], Hiroki Goda[2], Koji Harada[3]
([1]Tokai University, [2]Kyusyu Institute of Technology, [3]Nishimatsu Construction Co., Ltd., Japan)

As the Japanese government decided to boost the carbon neutral power source development, offshore wind farm projects are emerging in Japan and hundreds of platforms will be constructed in the near future. Some of these platforms are possible to be floating structures and made of concrete-like material whose biofouling should be limited from the viewpoint of drag force reduction. An autonomous cleaning robot is one of possible solutions to minimize the effect of biofouling. To develop the cleaning device, we started field experiments to study biofouling process. In this paper, we introduce some results of biofouling monitoring experiments using several different concrete-like samples at sea.



OS27-2 Fall risk notification system using LiDAR sensor for the visually impaired people

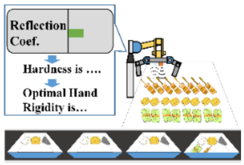
Daigo Katayama[1], Kazuo Ishii[1], Shinsuke Yasukawa[1], Satoshi Nakadomari[2], Koichi Wada[2], Akane Befu[2], Chikako Yamada[2] ([1]Kyushu Institute of Technology, Japan, [2]NEXT VISION, Japan)

We have developed the fall risk notification system using LiDAR sensors to reduce the number of fall accidents on platform involving visually impaired people. This system contains of iPhone 12 Pro with the small LiDAR sensor and the cane grip vibration device that notifies the user of fall risk using vibration. In this paper, we report the experiment results of the environment recognition algorithm for the fall risk notification system. In this system, the height map is generated from the depth information obtained from the LiDAR sensor and the posture of the iPhone. The system then performs the threshold process based on the height of the iPhone from the road surface during use. In the experiment, we evaluated the detection accuracy and responsiveness when entering risky area of falling, such as stairs and gaps.

OS27-3 Reflection Coefficient Estimation through the Modelling of Ultrasonic Transmission

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

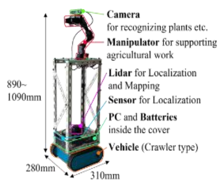
In food industry, shortage of working force is one of a critical problem. As a background, in Japan, increasing single households and demands for prepared dishes are relevant, thus, automation of food packing is a challenge. As the automation, food picking operation must be done safely without damaging food by robotic hand. We propose non-contact acoustic estimation with ultrasonic wave before the picking, having assumption of the relationship between hardness and acoustic impedance, that is, the product of density and sonic velocity. Furthermore, larger reflection coefficient can be observed from media of higher acoustic impedance. We built up ultrasonic transmission model considering attenuation by reflection and absorption, then, made an experiment with two overlapped media to configure the change of reflection coefficient at the surface of lower medium depending on its acoustic impedance. As the result, larger reflection coefficient changed according to rising acoustic impedance of the lower medium.



OS27-4 Evaluation of Maps Constructed by Crawler-type Agricultural Robot in Different Farms

Takuya Fujinaga, Tsuneo Nakanishi (Fukuoka University, Japan)

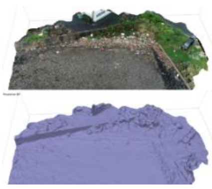
Various studies have been carried out with the aim of realizing smart agriculture. Many of them have targeted at large fields and large-scale facilities that are relatively easy to verify and implement robots and IoT devices etc. On the other hand, although the ratio of small-scale facilities is high in Japanese agriculture, there are few cases of study targeting small-scale facilities. We aim to develop an agricultural robot for supporting agricultural work that can move autonomously in the field for small-scale vinyl house. This paper evaluates maps of three different environments (strawberry farm, herb farm and vegetable farm) constructed using the developed agricultural robot. In addition, through verification experiments and evaluation results, we describe the requirements and concerns for operating robot in small-scale vinyl house.



OS27-5 An Estimation Method of Coastal Ocean Debris Using Aerial Drone

Kazuo Ishii[1], Kanako Shirahashi[1], Yuya Nishida[1], Moeko Tominaga[2], Yoshiki Tanaka[1], Dominic B. Solpico[1]
([1]Kyushu Institute of Technology, Japan, [2]Nishinippon Institute of Technology, Japan)

Although the actual situation of ocean debris has not been measured accurately, innumerable ocean garbages are drifting in the ocean. Especially, non-perishable waste such as microplastics continues to grow and is damaging marine life, including endangered species, and some are washed ashore and causing pollution damage to coastal areas. Microplastics incorporated into marine organisms, Arctic sea ice, and deep-sea seafloor sediments have also been detected. The Ellen MacArthur Foundation in the United Kingdom estimates that the total amount of marine debris exceeds 150 million tons, with more than 8 million tons of new inflow each year. We measured and compared the amount of ocean debris in coasts in Hirado and Matsuura cities, Nagasaki with manual count and an aerial drone observation.



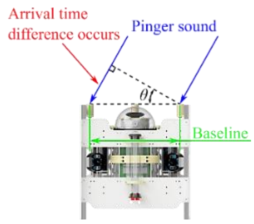
OS28 Robot Competitions and Education (6)

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

OS28-1 Underwater Acoustic Positioning Based on MEMS Microphone for a Portable Autonomous Underwater Vehicle

Irmiya R. Inniyaka, Dominic B. Solpico, Daiki Hamada, Akihiro Sugino, Rikuto Tanaka, Yuya Nishida, Kazuo Ishii (Kyushu Institute of Technology, Japan)

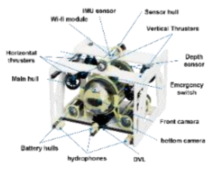
Autonomous Underwater Vehicle positioning is important for accurate control and high-quality data collection during operation. Conventional acoustic positioning systems are expensive. This paper describes the design and performance results of an inexpensive acoustic system for a lightweight AUV “Kyubic” used in Underwater Robotic competition, Okinawa 2021. The positioning method is based on Super-short baseline (SSBL) principle. The system design comprises of self-made hydrophone (3) module using MEMS microphone, ReSpeaker 4-mic array, Raspberry pi, and isolated power supply module. Using Python, the distance and angle of processed acoustic signals are integrated for dynamic control strategy of AUV to locate the position of a Pinger (acoustic pulse generator).



OS28-2 Autonomous Underwater Vehicle with Vision-based Navigation System for Underwater Robot Competition

Kazuki Harada, Riku Fukuda, Yusuke Mizoguchi, Yusuke Yamamoto, Kouta Mishima, Yoshiki Tanaka, Yuya Nishida, Kazuo Ishii (Kyushu Institute of Technology, Japan)

An Underwater Robot Competition (URC) is organized as the concurrent event of Techno-Ocean 2021 to advance underwater technology. The URC has two leagues, Autonomous Underwater Vehicle (AUV) and Free style robot leagues, and our team, Kyutech Underwater Robotics, jointed the AUV league using the developed AUV. The missions of AUV league consist of Gate Pass colored by green, Buoy Touch with the yellow and red colors, and Homing to a Pinger. Our AUV tackles the missions based on a vision-based navigation system. The paper explains the AUV system, the mission strategy and detail of image processing.



OS28-3 Tomato Harvesting in Greenhouse Considering the Effect of Sunlight

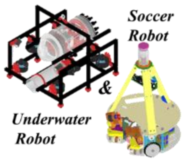
Kai Shioji, Shinsuke Yasukawa, Kazuo Ishii (Kyushu Institute of Technology, Japan)

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. For the operation in greenhouses, the recognition system with color constancy under sunlight is necessary. In the harvesting, tomatoes should be handled gently for less damages so that the soft handling end-effector is needed. In this paper, we introduce the system configuration of the robot and the experiments conducted to solve the problem in the greenhouse.

OS28-4 Inter-University Collaboration Aimed at Integrating Different Robotic Fields: Development of Underwater Robots and Soccer Robots Through these Competitions

Moeko Tominaga[1], Jonghyun Ahn[2], Yasunori Takemura[1], Kazuo Ishii[3]
([1]Nishinippon Institute of Technology, Japan, [2]Hiroshima Institute of Technology, Japan, [3]Kyushu Institute of Technology, Japan)

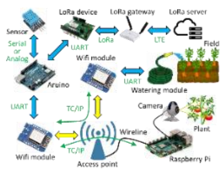
In robotics, the problems and solutions to be focused on may differ due to the different fields of robots to be developed. This is a phenomenon that appears depending on the operating environment of the robot, but it may be applicable to various fields. Also, the development factors that the students focus on differ depending on the university or laboratory the students belong to. In this research, we will verify the effect of exchanging opinions and sharing knowledge by collaborating with students aiming to participate in different robot competitions between underwater robots and soccer robots. Due to the influence of Covid-19, the robot competition was held online, and the cooperation between universities that had been continued online became possible on the day of the competition. This report was analyzed and evaluated based on the results of regular questionnaire monitoring of students' willingness to work on projects and activity records.



OS28-5 Exercise on Environmental Monitoring and Control of Greenhouse by IoT Devices toward Smart Agriculture

Yuya Nishida, Ryugo Mochizuki, Shinsuke Yasukawa, Kazuo Ishii (Kyushu Institute of Technology, Japan)

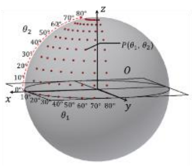
In the future estimation of 2050, the food demand increases 70% and the production decreases 15% caused by global warming, and farmer population be 1/5 of current workers. We need an agricultural system of twice production with the same farmland area and 5 times effective operations until 2050. For the sustainable society, smart agriculture including robot technology, AI, IoT is one of the solutions for food issues. As crops in greenhouses are widely distributed, IoT devices placed near the crops should be stand-alone and modular, and data from the devices are collected over the networks. Camera is often used for monitoring of growth status of the crops, and only resulted information by image processing should be transmitted. Smart agriculture requires knowledge of a wide range of fields including electricity, information, and image processing. We have designed an AI and IoT technology exercise on environmental monitoring and control of a greenhouse where we have been preparing for grow up of tomatoes and other vegetables.



OS28-6 Evaluation of roller arrangement of sphere by omnidirectional integral value

[1]Kenji Kimura, [2]Yusuke Abematsu,[3]Hirai Hiroyasu, [3]Kazuo Ishi
([1]Fukuoka Daiichi High School, Japan, [2]Kagoshima Gyokuryu High School, Japan,
[3]Kyushu Institute of Technology, Japan)

The conventional sphere moving mechanism driven by a constraining roller has a fixed roller rotation axis, so the angular velocity vector has two degrees of freedom. Here, if the degree of freedom is 3, the rotational diversity of the sphere increases. In this study, we propose a spherical mechanism with variable roller rotation axis and consider the problem of spherical transport with the best kinetic energy efficiency. At this time, using the kinetic energy integration in all directions (direction of travel of the sphere) as an evaluation function, the evaluation distribution of the roller contact position is considered, and the conventional mechanism is compared with the variable mechanism.



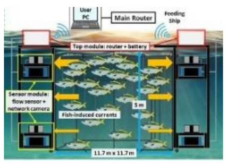
OS29 Advances in Marine Robotics and Their Applications (8)

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

OS29-1 A Sensor Network to Estimate Fish Activity and Assist Feeding Decisions in Marine Aquaculture

Dominic B. Solpico, Yuya Nishida, Kazuo Ishii (Kyushu Institute of Technology, Japan)

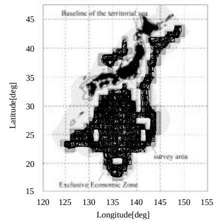
Fish feeding is one of most important issues in marine aquaculture. Satiety feeding causes high expenses and ocean pollution so that optimization of feeding and DX of expert farmers are necessary toward sustainable marine aquaculture. Currently, the decision-making has been based on subjective experience of expert farmers. This paper presents the development of a sensor network for estimating fish feeding behavior, which could help in optimize feeding by less experienced farmers. This network is made up of arrays of sensor modules consisting of current sensors, inertial measurement units and cameras. We introduced this system in fish cages and measured fish activity from start to end of feedings. Experiment results indicate that currents were highest at the surface at the feeding proper, rising and dropping around when feeding started and ended, respectively. Expert farmer finds a small fish behavior change and stop the feeding, whose timing is also suggested from obtained data.



OS29-2 Proposal of a Swarm Intelligent Underwater Glider System for a Long-term Three-dimensional Wide-area Ocean Observation

Kanako Kobatake, Masakazu Arima (Osaka Prefecture University, Japan)

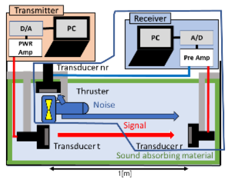
The ocean absorbs heat and carbon dioxide from the atmosphere, and it serves to mitigate climate change caused by human activities. However, ocean is currently facing serious threats due to global warming and ocean acidification. In order to conserve ocean, it is very important to have a correct understanding of the ocean environment conditions over a long term and wide area. In this research, the authors propose a swarm intelligent underwater vehicle system for a long-term three-dimensional wide-area ocean observation. Several cases of underwater cruise simulation using the autonomous underwater glider with independently controllable main wings, SOARER, were carried out for the validation of the effectiveness of the proposed system, and estimated the optimal required number of vehicle that cover the Japan's EEZ.



OS29-3 Underwater Acoustic Communication using QPSK Modulation Method

Yuya Nishida, Yuichiro Uemura, Rikuto Tanaka, Kazuo Ishii (Kyushu Institute of Technology, Japan)

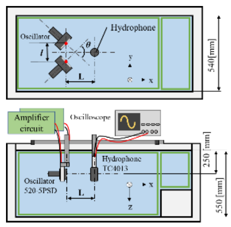
Acoustic transmission has less attenuation than those of radio or light and is commonly used for communication between underwater robots. However, acoustic communication has a lot of influence of noises, and difficulties in sending and receiving data correctly because of the noise from robot’s thrusters and other acoustic devices, and its baudrate is only 1/1000 of conventional Wi-Fi communication. To improve the acoustic communication performance, this research developed acoustic communication system against the thruster’s noise. The acoustic communication system consists of a transducer, two Amp circuits, device including A/D and D/A convertor, and PC for modulation. Communication message includes not only data (payloads) but also sync signal and error correction bytes, and the message is converted to acoustic wave by QPSK modulation which four phase data represent two bits.



OS29-4 Ultrasonic Cleaner using Two Transducers for Ship Hull Cleaning Robot

Yuya Nishida, Toshihiro Matsumura, Kazuo Ishii (Kyushu Institute of Technology, Japan)

Fuel consumption of the ship gets worse by barnacles and stains put on its bottom, so that the ship hull should be cleaned regularly. Usual methods such as cleaning by special divers and cleaning after pulling up to the dock are not used frequently due to those high cost and the heavy burden on cleaners. We proposed a cleaning method using underwater vehicle with brushes, however there is a possibility to remove paints on the hull. To improve cleaning performance of the underwater vehicle, the ultrasonic cleaner is developed by using cavitation occurred by sound waves of the acoustic transduces. The cleaner generates sound waves with sound pressure of 0.2MPa or more required to occur cavitation at the point where sound waves from two transducers overlap. In experimental results, the cleaner occurred cavitation enough to make two holes in the aluminum foil located 50mm away from transducers.



OS29-5 Motion Control of a Ship Hull Cleaning Robot

Hyoga Yamamoto, Yuya Nishida, Takayuki Matsuo, Kazuo Ishii (Kyushu Institute of Technology, Japan)

Recent trend towards increased cost of fuel will continue further, so the reduction of fuel consumption will be required more severely as well as the severe requirement of reduction of CO2 emissions. To achieve these requirements, the reduction of resistance is essential for efficient ship navigation. One of solution is the prevention of marine biofouling. Anti-fouling paint is effective to prevent biofouling to the ship hulls like barnacles, however, even the painted hull acquires slime-like biofouling caused by marine alga on its surface easily. The cleaning of the ship hull is generally carried out during inspection in dockyard once a year. Frequent cleaning while the ship is berthing is desirable to keep good fuel efficiency. If frequent ship hull cleaning is possible while ships dock at berth with ease, ships can keep good fuel efficiency that makes the transportation costs and CO2 production less. Cleaning by divers costs much and is a high risk task. One possible solution for this issue is to introduce underwater robots for ship hull cleaning. In this paper, the motion control system of the robot is described.

OS29-6 Development of a USV Testbed and Its System Check Experiments at Sea

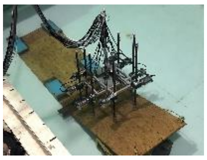
Keisuke Watanabe, Masatoshi Shimpo (Tokai University, Japan)

The catamaran type USV is useful as an autonomous platform in the field of marine engineering, such as collection of marine waste in ports, ocean observation platforms, and automated vessels for offshore wind farm maintenance. The authors have been developing an experimental testbed of USV platform to conduct basic studies on control algorithms by simulation as well as hardware systems to conduct sea experiments. In this paper, we report on the construction of an actual sea area experimental system and basic experimental results.

OS29-7 Development of a Seabed Walking Platform for Ore Sample Drilling in Deep Sea Mining

Keisuke Watanabe[1], Hideyuki Suzuki[2], Yoshiyasu Watanabe[1]
([1]Tokai University, Japan), ([2]The University of Tokyo, Japan)

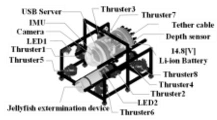
In deep-sea mineral resource development, exploratory drilling is indispensable for estimating the amount of resources. In order to reduce the cost of exploration, a system for drilling on the seabed is needed instead of a support vessel at sea. This exploration platform must be able to move along with the undulations of the seafloor and have a structure that supports the reaction force of the excavation. The authors have been studying an eight-legged drilling platform. This paper introduces the system configuration of an eight-legged walking robot.



OS29-8 Development of Remotely Operated Vehicle for Small-size Jellyfish Extermination and its Evaluation of Extermination Motion Control

Hiroyuki Yokota[1], Shinsuke Yasukawa[2], Jonghyun Ahn[1]
([1]Hiroshima Institute of Technology, Japan, [2]Kyushu Institute of Technology, Japan)

In recent years, increase in the number of jellyfish has caused damage in the fishery and tourism industries. Therefore, the extermination work of jellyfish is being carried out by human hands. However, conventional methods for extermination are required a lot of time and manpower. In this paper, we propose a method for extermination work of jellyfish using underwater robot. Also, we introduce developed ROV type underwater robot, which is called J.E.N.O.S., and its extermination motion control. The J.E.N.O.S. is developed in consideration of the attitude control during the extermination operation. Because, the attitude, such as surge and pitch angle, of J.E.N.O.S. becomes unstable when performing jellyfish extermination. Therefore, we equipped 8 thrusters to improve attitude stability during the jellyfish extermination. As a result, surge acceleration is reduced to about 30.0%, and pitch angle velocity is reduced to about 25.8%.



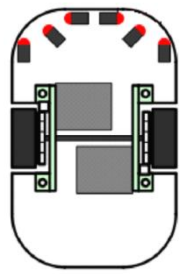
OS30 Intelligent Systems and Robotics -4-(6)

Chair Jichao Zhao (Tianjin University of Science and Technology, China)
Co-Chair Hucheng Wang (Tianjin University of Science and Technology, China)

OS30-1 A Design of Micromouse Control System

Hucheng Wang (Tianjin University of Science and Technology, China)

This paper improves the problems of the traditional computer mouse, such as the vehicle weight, inflexible turning, and the sensor can not accurately feed back the external environment information. In terms of machinery, the solid, durable and light-weight material is used as the vehicle body, and the transmission mode adopts belt indirect transmission, so that the computer mouse has a lighter, more stable and more flexible structure. The sensor module is optimized to make it more sensitive and accurate to perceive the external environment. The software part uses the potential value search algorithm, combined with the external information collected by the infrared sensor, so that the computer mouse can sprint from the starting point to the focus more quickly. This paper also innovatively adds the voltage detection module, which solves the problem that the traditional computer mouse can not know the remaining power in the competition, and ensures the stability of power supply.



OS30-2 A Driver Reaction Time Detection System Design

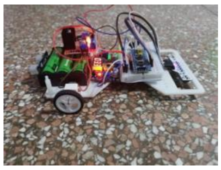
Yuhui Cheng[1], Mochi Li[2]
([1] Tianjin University of Science and Technology, China), ([2] Ocean University of China, China)

This paper proposes a design scheme of driver reaction time detection system, which collects electroencephalography (EEG) data and reaction time required by the driver to complete corresponding test actions by deploying sensors on the driving simulator. The scheme can be divided into control unit, data acquisition unit and interaction unit. The test performs corresponding operations according to the instructions issued by the interaction unit, and the system sorts out and saves the collected data. This system is of great significance to study the reaction time of drivers in different states and scenarios.

OS30-3 A PID Tracking Car Design based on STM32

Yande Xiang, Teng Zhang, Shixiang Zhao, Ling Zhou, Mingjuan Tian, Haoran Gong
(Tianjin University of Science and Technology, China)

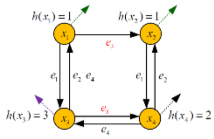
Intelligent car is a high-order intelligent product integrating intelligent integrated control, visual terminal, automatic control output and cognitive computing. This paper introduces an innovative intelligent tracking car based on STM32 embedded chip. The vehicle adopts fuzzy PID algorithm to control the vehicle operation, and adopts the innovative grayscale card designed by the team hardware designer. This car fully caters to the global environmental protection trend, the new concept of green development. On the premise of low cost and low power consumption, it also ensures the smooth, smooth and high precision operation of the vehicle.



OS30-4 Matrix Approach to Current-state Detectability of Discrete-event Systems

Jinliang Wang, Jiawei Wei, Xiaoguang Han (Tianjin University of Science and Technology, China)

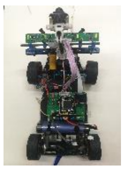
In our previous work, a matrix-based framework is proposed to tackle the problem of verifying strong detectability in the context of partially-observed nondeterministic discrete event systems (DESs). Two key concepts, namely, unobservable reach and detector, are redefined therein. Also, the dynamics of a detector, under the frameworks of the Boolean semi-tensor product of matrices, are converted equivalently into an algebraic representation. In this paper, we extend our previous work to other versions of detectability, including strong periodic detectability, weak detectability, and weak periodic detectability. Several necessary and sufficient conditions are derived for verifying aforementioned three types of detectability, respectively. Compared with the existing ones, the proposed methodology is easier to be implemented in software in the sense that it avoids the symbolic manipulations. Finally, an example is given to illustrate the theoretical results.



OS30-5 Hardware Circuit Design Of Tracking Car Based On K60

Peng Jia, Yande Xiang (Tianjin University of Science and Technology, China)

The development of the robot industry is booming. For many repetitive and difficult jobs, using industrial robots to complete them can greatly save money, manpower and time costs. Industrial tracking vehicle AGV has been applied in many fields and has shown good results. Taking the industrial AGV as the model, this paper analyzes the actual needs, explores, designs and builds a set of hardware model of intelligent tracking car based on K60. The main work includes hardware circuit scheme design, component selection, schematic diagram and PCB design, PCB board welding.



OS30-6 Detachable IoT Garbage Sorting Device Based on Machine Vision

Tao Zhu,Yang Su, Zhiqing Xiao, Fengzhi Dai (Tianjin University of Science and Technology, China)

China is promoting a garbage sorting system, but people have to spend more time changing their habits, and the garbage bins on the road do not have the function of automatic sorting. Therefore, a detachable IoT waste sorting device based on machine vision is proposed. Through machine vision and a unique mechanical structure design, it can be directly installed on existing trash cans, and is suitable for two and four classifications that meet the Chinese classification standards. Inside the device, it is fixed with an existing trash can with a mechanical electronic lock, and the switch lock operation can be performed through the mechanical key and the Internet of Things applet, which is convenient for the replacement of the device.

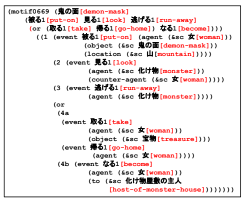
OS31 Approaches to Post-Narratology that Combines AI and Cognitive Science with Narratology (9)

Chair Jumpei Ono (Aomori University, Japan)
Co-Chair Hiroki Fxyma (Tainan University of Technology, Taiwan)
Co-Chair Takashi Ogata (Iwate Prefectural University, Japan)

OS31-1 Story Units of the Types of Japanese Folktales and the Combination with a Noun Conceptual Dictionary

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

Story units in this study mean units described by Common Lisp based on the types of folktales analyzed by Seki and his group. In particular, we have been developing story units to use in our narrative generation system, namely an Integrated Narrative Generation System (INGS), as a type of narrative techniques that generate a narrative structure based on the synthesis, transformation, expansion, etc. of a narrative structure. Story units function in INGS through the combination with the conceptual dictionaries in INGS. We have attempted the combination of story units with the verb conceptual dictionary. The objective of this paper is the combination with the noun conceptual dictionary. As the main constitutional elements of each story unit are verb and noun concepts, their combinations with the noun conceptual dictionary enable the substantial function of story units based on the types of Japanese folktales as a group of narrative techniques in INGS.



OS31-2 Visualization of the Unconscious in Quality Inspection in Manufacturing

Jun Nakamura (Chuo University, Japan)

In quality inspection, which is the final stage of the manufacturing process, there is an operational manual on how the object should be inspected. In this study, we followed the gaze of the workers to search for the "way of looking" at the objects to be inspected, which is not described in the manual. As a result of experiment, we found a difference between the reciprocating eye movements that occur in skilled worker and the static state of newcomers. It is suggested that this indicates a way of looking at inspection that has been unconscious until now.

OS31-3 The Study on the Relationship Between the Comic Artists' Styles and the Visual Languages: From the Stylistic Changes in the Work of Japanese Comic Artists

Kaori Otsuru (Tainan University of Technology, Taiwan (ROC))

As a first step for the machine-based comics analysis, we propose approaching the stylistic aspects of comics from the framework of linguistics and narrative theory. Machine learning has been applied to paintings and text mining, but comics, a synthesis of these works, contain elements that cannot be analyzed simply by integrating pictures and texts. In this study, I propose a broad definition of the "style" of comics as covering the following concepts; the micro/macro-level development of the story and composition of comics, the synchronous/diachronic changes in drawing, and the impressions of the work obtained through the manga works. This study is intended as a basis for future research on manga style using machine learning.

OS31-4 Theoretical Backgrounds toward Text Mining for a Phenomenological Model of Taste Perception

Hiroki Fxyma (Tainan University of Technology, Taiwan)

In this presentation, I discuss the temporal aspects of cognitive content generation and appreciation of taste. The cognitive content of taste evolves chronologically, as does the appreciation of music. However, the time which is reconstructed and given meaning as a story is considered to have a different time axis from the physical time (i.e., linear time, point time, cyclic time, and so on). Toward a chronological model for the taste phenomenon, I examine the language usage in the tasting comments of wine and sake. A lot of tasting phrases potentially include the concept of time. For example, "drifting" indicates continuous time. By examining how temporal terms co-occur with taste elements, I will discuss the model of taste phenomenon.

OS31-5 Why is the Early Detection of Dementia Failed?

Yuki Hayashi (Chiba University/National Institute for Japanese Language and Linguistics, Japan)

To prevent dementia, the early detection is important. However, it's often failed because it's difficult to articulate "dementia" as "dementia" verbally. In this paper, I analyzed semi-structured interview data stored in DIPEx-Japan, a database of personal experiences of health and illness, and discussed the difficulty of early detection of dementia.

OS31-6 Relationship Between World-view and Advertising Techniques

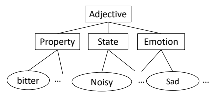
Yoji Kawamura (Kindai University, Japan)

The "world-view" plays an important role in advertising. In this research, we first extracted the keywords of the world-view related to advertising. Next, those keywords were classified by factor analysis, and a regression analysis was performed on the relationship between the classified genres and advertising techniques. 66 keywords related to the world-view were extracted and classified into 8 genres by factor analysis. The characters & product movements, atmosphere, and products (functions, naming & logos) had a great influence on the world-view of the 8 genres. We also quantitatively clarified the advertising techniques related to each genre. These relationships provide implications for planning advertising creatives.

OS31-7 Adjective and Adjective Verb Conceptual Dictionaries in an Integrated Narrative Generation System

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

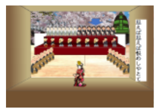
Although main conceptual types of our previous narrative generation study using conceptual dictionaries were verb and noun concepts, other conceptual types are necessary for the implementation of more precise narrative generation functions. In this paper, we prepare the frameworks for adjective concepts and adjective verb concepts in our narrative generation system called Integrated Narrative Generation System (INGS). Furthermore, we define the information of opposite meaning for each adjective and adjective verb concept. For example, this function will contribute to introduce a type of contrasting rhetoric into narrative generation.



OS31-8 Prototyping Animation System that Combines a Kabuki Work and its Background Story: *Kyōganoko Musume Dōjōji* and the Legend of Dōjōji

[1]Miku Kawai, [2]Shunta Kudo, [3]Jumpei Ono, [2]Takashi Ogata
([1]The Open University of Japan, [2]Iwate Prefectural University, [3]Aomori University, Japan),

Kyōganoko Musume Dōjōji, a famous kabuki dance work, is an adaptation based on the legend of Dōjōji. A series of research regarding Kyōganoko Musume Dōjōji includes many themes, such as the analysis and simulation as an animation system of the stage-performing structure, the survey and analysis of the relationships between Kyōganoko Musume Dōjōji and the legend of Dōjōji, and the design and experimental system development reflected the above relationships. Based on these studies, in this paper, we present an animation-based mechanism that flexibly associates the narrative flow of the stage-performing structure to the story of the legend of Dōjōji.



OS31-9 How Will Art Appreciations Change According to Information Change?

Akinori Abe (Chiba University, Japan),

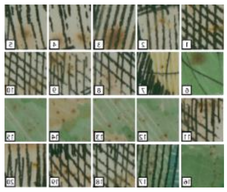
For the art appreciation in museums, usually a certain information will be provided as a caption. Visitors usually read the description to help his/her understanding. Thus, such a help will be necessary for ordinal person's understanding. Previously, we conducted an experiment to determine viewers' artwork understandings by gradually adding information in caption. In the previous experiment, we offered information only with the official information. In this paper, we will generate information with artists' own writings (artist's explanation). In addition, in this experiment, we used a rather abstract painting and representational but rather strange and difficult to understand paintings. In this paper, I will show how viewers' understanding or interpretation of artworks, and sense of value and preference of artworks change according to the changing information.

OS32 Human-Machine Interface and Automation (9)
Chair Norrima Mokhtar (University of Malaya, Malaysia)
Co-Chair Fakhrul Hazman Yusoff (UITM, Malaysia)

OS32-1 A Derivative Oriented Thresholding Approach for Feature Extraction of Mold Defects on Fine Arts Painting

Hilman Nordin, Bushroa Abdul Razak, Norrima Mokhtar, Mohd Fadzil Jamaludin
(Universiti Malaya, Malaysia)

Identification of mold defects is an important step in the restoration of damaged paintings. The process is usually lengthy and depends heavily on the qualitative visual judgement of an expert restorer. This study proposes an automatic mold defect detection technique based on derivative and image analysis to assist in the restoration process. This new method designated as Derivative Level Thresholding (DLT) method combines binarization and detection algorithms to rapidly and accurately detect mold from scanned high-resolution images of a painting. The proposed method is compared to existing binarization methods of Otsu's Thresholding Method, Minimum Error Thresholding and Contrast Adjusted Thresholding Method. Experimental results from the analysis of 20 samples from high-resolution scans of 2 mold-stained painting have shown that the DLT method is the most robust with the highest sensitivity rate of 84.73% and 68.40% accuracy.



OS32-2 Imaginary Finger Control Detection Algorithm Using Deep Learning with Brain Computer Interface (BCI)

Suresh Gobee[1], Norrima Mokhtar[1], Hamzah Arof[1], Noraisyah Md Shah[1], Wan Khairunizam[2]
([1]Universiti Malaya, Malaysia), ([2]University Malaysia Perlis, Malaysia)

Before the advancement of deep learning technology, the brain signals are to be analysed manually by the neuroscientist on how the brain signals reacts in proportion with human body. This process is very time consuming and unreliable. Therefore, this project aims to develop a brain signal detection based on deep learning algorithm in response to the output of EEG device on the imagery finger movements. These fingers include thumb, index, middle ring and little of right hand. There are four CNN classification models being developed in this project. The different between the models are the pre-processing requirements and neural network architecture. The best results for offline classification obtained in this project are 69.07% and 82.83% respectively in terms of average accuracy from 6-class ad 2-class tests. The developed work can be applied for proof of concept for online study.

OS32-3 Investigating the Effect of Individuality Factors in Measuring Aggresion Induced by Human Brain

Wan Khairunizam[1], Kai Xu tung[1], Lugieswaran[1], Wan Azani Mustafa[1], Hashimah Ali[1], Zuradzman M. Razlan[2], Shahriman AB[2], Norrima Mokhtar[3]
([1]University Malaysia Perlis, Malaysia), ([2]Universiti Malaya, Malaysia)

Aggression is a behavior of human that may cause physical or emotional harm to others. Several factors that cause aggressive behavior such as physical health, mental health and socioeconomic. Many previous researchers reported that aggression could be measured through either questionnaire or the brain signals. This paper proposes the experimental studies to collect the brain signal of the human subject for investigating the effect of individuality in aggression. Ten subjects are selected to perform the aggression activities. The experimental protocol for inducing aggression is proposed. In general, there are four tasks which is collecting brain data in relaxing state before and after the experiments, and data collection while playing game in muted and maximum volume levels. In the experiments, the subject are required to play a popular non-violence smart phone game named “Subway Surfers” and at the same time the EEG signals are recorded from the subject’s brain. In the signal pre-processing stage, a Butterworth filter is used to remove the noises contain in the signals. A windowing technique is employed for extracting significant features. A Pearson correlation technique is used to reduce and remain the less and most significant features. In the methodologies, the aggressiveness level A, is defined to investigate the effect of individuality in inducing the aggression signals. The proposed experimental protocol and signal processing techniques are seen able to generate level of aggression.


Figure 1. Mindset 24 EEG amplifier

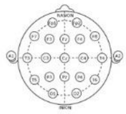
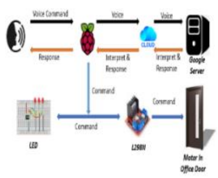


Figure 2. EEG electrode’s location according to 10-20 system

OS32-4 Voice User Interface(VuI) Smart Office Door Application in the Context of Covid-19 Pandemic

Muhammad Zharif Aiman Alias, Wan Norsyafizan W. Muhamad, Suzi Seroja Sarnin, Darmawaty Mohd Ali
(University Technology MARA, Malaysia)

Nowadays, all countries around the globe are trying their best to prevent the spread of Covid-19 from reaching their people. However, the virus spread is through the transmission of close contact with the infected person and touching a surface that already contains the virus. This project presents Voice User Interface (VUI) Smart Office Door System that interact with the Internet of Things (IoT) to execute the command given by the user. This project use Raspberry Pi as microcontroller which Thonny Python software used for running the coding script of the system. The electrical components such as dc motors and LEDs are connected to General-Purpose Input Output (GPIO) pin of Raspberry Pi and motor driver. One of the dc motors used for controlling the lock and another one used for controlling the door. Blue LED used as locked door notification while green LED used as unlocked notification. Microphone and speaker connect to the Raspberry Pi through usb port and Bluetooth. This system used Google Assistant as its VUI to control the smart door contactless. The user can give the command to the system through microphone to control the output of the system. The purpose of this project is to invent the smart door with IoT technology that help prevent the spread of Covid-19 at the workplaces.



OS32-5 Endometrial Cell Images Segmentation: A Comparative Study

Wan AzaniMustafa[1], Nurul Umaira Salim[1], Wan Khairunizam[1], Shahrina Ismail[1], Hiam Alquran[2] ([1]Universiti Malaysia Perlis, Malaysia; [2]Yarmouk University, Jordan)

Uterine cancer, also known as endometrial cancer, is a form of cancer that affects the female reproductive system. Nowadays, there are 2 step methods that the physician or health care provider tend to use to diagnose cancer, which is using ultrasound technique and endometrial biopsy. The biopsy procedure is used to extract the cell and sent to the pathologist for histopathological image analysis. The histopathological image analysis is the crucial step in all the procedures because it determines the situation for the patient, whether positive or negative. They are two types of cell images known as high grade squamous intraepithelial lesion (HSIL) and low grade squamous intraepithelial lesion (LSIL). The problem occurs when both LSIL and HSIL are different, needing different medical treatment techniques but showing slighter differences in nucleus size cell histopathological image analysis. Therefore, the pathologist usually requires more time to identify whether it is LSIL or HSIL. Based on the limitation, the paper aims to compare a few popular detection methods, which are the Wolf method, Bernsen method, Otsu method and Feng method. Based on the Image Quality Assessment (IQA), the Wolf method shows good performance compared to the others. In a precise term, this finding could benefit the health care community to reduce the diagnosis time to categorize the cell and lead to early treatment of endometrial cancer.

OS32-6 Temperature Control Using Fuzzy Controller for Variable Speed Vapor Compression Refrigerator System

Siti Qurrata Ain Suhaimi, M. Saifizi, S.M. Othman, Azri A. Aziz, Wan Azani Mustafa, Wan Khairunizam (Universiti Malaysia Perlis, Malaysia)

Keeping the cold chain vaccine is crucial to a stable immunization programme; however, faulty processes may occur more frequently than are often thought in developing nations. This paper discusses the quick and accurate control process for designing fuzzy controllers for variable speed vapor compression refrigerator system. The suggested controller is based on the fuzzy logic intended to improve performance while keeping the cooler's constant internal temperature and increasing the refrigerator efficiency. Despite the external changes such as the outside temperature change or the volume change in the refrigerator vaccine, the fuzzy logic controller is utilised to maintain the interior temperature. However, a variable speed compressor (VSC) must be used to control the thermophysical characteristics, which dramatically alter the temperature with a small pressure change. In this case, fuzzy rules of the sort developed by Mamdani are used to build up the system. The programming platforms utilised to implement the model include MATLAB, SIMULINK, and Fuzzy Logic Toolbox (FLT). The efficiency of fuzzy logic controller design membership will be compared to ensure that the refrigerator temperature is more accurate and until it achieves the best performance, maintains a temperature of 5°C, and adapts to its surroundings. From the research done, the membership 2 with load shows the near accurate temperature of 5°C with steady-state error ±1.97°C.

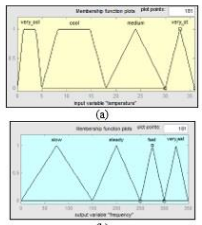
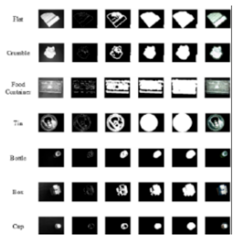


Fig.6. Membership 2 of Fuzzy Controller

OS32-7 Automatic Dry Waste Classification for Recycling Purpose

Muhammad Nuzul Naim Baharuddin[1], Hasan Mehmood Khan[1], Norrima Mokhtar[1], Heshalini Rajagopal[2], Tarmizi Adam[3], Wan Amirul Wan Mahiyiddin[1], Jafferi Jamaluddin[4]
([1]Universiti Malaya, Malaysia). ([2]Manipal International University, Malaysia),
([3]Universiti Teknologi Malaysia, Malaysia), (UMPEDAC Universiti Malaya, Malaysia)

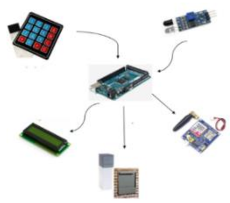
There has been a serious increment in solid waste in the past decades due to rapid urbanization and industrialization. Therefore, it becomes a big issue and challenges which need to have a great concern, as accumulation of solid waste would result in environmental pollution. Recycling is a method which has been prominent to deal with the problems, as it is assumed to be economically and environmentally beneficial. It is important to have a wide number of intelligent waste management system and several methods to overcome this challenge. This paper explores the application of image processing techniques in recyclable variety type of dry waste. An automated vision-based recognition system is modelled on image analysis which involves image acquisition, feature extraction, and classification. In this study, an intelligent waste material classification system is proposed to extract 11 features from each dry waste image. There are 4 classifiers, Quadratic Support Vector Machine, Cubic Support Vector Machine, Fine K-Nearest Neighbor and Weighted K-Nearest Neighbor, were used to classify the waste into different type such as bottle, box, crumble, flat, cup, food container and tin. A Cubic Support Vector Machine (C-SVM) classifier led to promising results with accuracy of training and testing, 83.3% and 81.43%, respectively. The performance of C-SVM classifier is considerably good which provides consistent performance and faster computation time. Further classification process is improved by utilization of Speeded-Up Robust Features (SURF) method with some limitations such as longer response and computation time.



OS32-8 A Low Cost Smart Parcel Box System with Enhanced Security

Ahmad Luqmanulhakim, Wan Norsyafizan W.Muhamad, Suzi Seroja Sarnin, Meor Mohd Azreen Meor Hamzah ([1]Universiti Teknologi MARA, Malaysia)

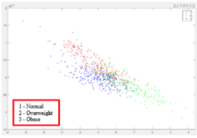
A global pandemic of covid-19 has hastened the growth of online shopping or E-commerce. Nowadays, E-commerce transactions provide various products from luxury goods and services to everyday necessities. While the popularity of online shopping has grown fast, there are several common issues that shoppers experience. Common problems that customers experienced are failed delivery, missing parcels and even the criminal case. Malaysia has introduced Pos Laju Ezi Box Kiosk, but still contemporary Kiosk type approach for the delivery system has some problems such as high initial installation cost, the expense of management system and especially security weaknesses in the wireless communication. To overcome the aforementioned problems, this paper proposes a low-cost smart parcel box system that will be installed at individual homes with enhanced security. This system used Arduino Mega 2560 to control all the processes of the developed system. The system will be initiated when couriers send the parcel's tracking number as a message to the user via applications to get the password. For security purpose, password will be provided once the courier's message is the same as the message specified by the user. Couriers then can enter the password provided and insert the parcel into the smart parcel box. The proposed low-cost smart parcel box system ensures that parcels are delivered safely and securely to the customer's door.



OS32-9 Classification of Body Mass Index Based Face Images Using Facial Landmarks Approach and PCA plus LDA

Hasimah Ali[1], Ho Yong Kang[1], Wan Khairunizam Wan Ahmad[1], Mohamed Elshaikh[1], Norrima Mokhtar[2]
([1]Universiti Malaysia Perlis, [2]Universiti Malaya, Malaysia)

Human faces contain rich information. Recent studies found that facial features have relation with human weight or body mass index (BMI). Decoding “facial information” from the face in predicting the BMI could be linked to the various health marker. This paper proposed the classification of body mass index (BMI) using facial landmark approach based on facial images. In this framework, Discriminative Response Map Fitting (DRMF) method has been used as feature extraction technique to detect and locate the facial landmark points on the facial images. About sixty-six (66) facial landmark points were identified. Only nineteen (19) of facial landmark points have been employed to extract the facial features in terms of distance and ratio features. A total of 221 facial landmark features were obtained and used as feature vector to classify the BMI classes. The rationale of using 221 facial landmark features is because these features were able to exhibit the unique characteristic of the BMI classes, which are normal, overweight and obese. Then, the extracted features were further reduced using Principal Component Analysis (PCA) plus Linear Discriminant Analysis (LDA) to map high dimension features into low dimensional feature with maximize between class scatter and minimize within class variations. Later, the reduced features were subjected to k-NN classifiers. A series of experiments has been conducted on MORPH II database using the reduced facial landmark features to classify the three BMI classes. Based on the experimental results, it shows that the reduced features using PCA plus LDA based on k-NN classifier has achieve the highest recognition rate with accuracy of 83.33 %. The obtained results show that the reduced facial landmark features were able to discriminate the three BMI classes of normal, overweight and obese, thus shows the promising results



OS33 Signal Processing and Chaotic System (5)

Chair Huailin Zhao (Shanghai Institute of Technology, China)
Co-Chair Fengzhin Dai (Tianjin University of Science and technology, China)

OS33-1 A Visual Measurement Algorithm of Approaching Vehicle Speed Based on Deep Learning

Yurong Zhu, Huailin Zhao, Liu Junjie, Zhang Jinping, Ji Xiaojun (Shanghai Institute of Technology, China)

With the urbanizational process expediting and the national economy developing rapidly and healthily, the amount of private cars is on the rise, and traffic accidents occur frequently due to speeding and other reasons, and the difficulty of traffic supervision has also increased. This topic will use semantic segmentation and feature extraction and matching. Based on the video data of the traffic surveillance camera, an algorithm is designed to quickly calculate the matching of feature points in adjacent frames with low computing power to achieve the calculation. The same vehicle moves within the two frames of the target, so as to calculate the speed of the vehicle. Firstly, performing semantic segmentation based on deep learning, we choose a fully convolutional network to achieve semantic segmentation of depth maps, and distinguish the picture's principal part. After that, we can realize features extraction and mapping. The HOG algorithm is used on the matching step, the target's relative movement is calculated based on these matched point pairs to measure the moving speed of the vehicle. The experiment and the test prove that the system can realize the efficient speed measurement of moving vehicles.

OS33-2 Target Search Based on Scene Priors

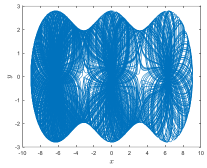
Shengyang Lu, Lanjun Liang, Huailin Zhao, Fangbo Zhou, Feng yao
(Shanghai Institute of Technology, China)

Aiming at the problems of reinforcement learning algorithm in target search tasks, such as low accuracy and low fault tolerance, this article mainly introduces a method of reinforcement learning target search based on scene prior in simulation environment. This method mainly uses graph convolutional neural network to extract the current object relationship as the input of prior knowledge. Secondly, it uses the actor-critic algorithm to take the agent's vision, position and prior knowledge as input to decide the agent's next navigation. Finally, use path planning to navigate to the target point to find the target. Through experiments conducted in Habitat and compared with the previous algorithm, the experiment shows that this method is better than the previous algorithm in target search accuracy and navigation efficiency.

OS33-3 A Generalized Hamiltonian Conservative Systems with Multi-scroll Chaotic Flows

Jingwen Liu, Zhonggao Chen (Tianjin University of Science and Technology, China)

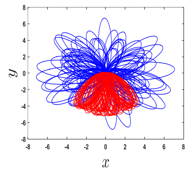
By analyzing mechanics and energy of a three-dimensional volume conservative chaotic system proposed by Vaidyanathan and Volos, a new generalized conservative chaotic system with multi-scroll chaotic flows is found based on the corresponding Hamiltonian energy. The new system satisfies both volume conservation and energy conservation. By analyzing the equilibrium characteristics of the system, equilibrium points of the new system are found to be a line. In addition, the number of scrolls of conservative chaotic flows of the new system depend on the corresponding Hamiltonian energy. The paper provides a new conservative chaotic model for chaos application.



OS33-4 Multi-stability and FPGA Implementation of a Conservative Chaotic System

Minghan Song (Tianjin University of Science and Technology, China)

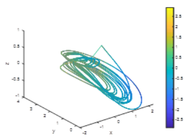
The paper first studies the reason why a three-dimensional volume conservative chaotic system proposed by Vaidyanathan and Volos can generate chaos by analyzing mechanics and energy. Then, based on numerical methods including balance characteristics, Lyapunov exponents, bifurcation diagrams, phase trajectories and so on, multi-stability of the three-dimensional volume conservative chaotic system are discovered. In addition, the three-dimensional volume conservative chaotic system is realized by using FPGA, and all the results from FPGA implementation are consistent with those from numerical analysis. Which further verify multi-stability of the three-dimensional conservative chaotic system from physical characteristics.



OS33-5 A New Hyperchaotic Financial System

Lei Gong (Tianjin University of Science and Technology, China)

In this paper, a new hyperchaotic financial system is obtained based on a financial system. It is first transformed into Kolmogorov model, which is composed of conservative torque, dissipative torque and external torque, to study the reason why the new system can generate chaos. Then, by studying energy exchange and combining different torques, dynamics of the new system is analyzed, the external torque is found to be the main reason the new system generate chaos. The paper provides a new method of analyzing chaotic dynamics in financial system, and further promotes new strategies are found to control chaos in financial market.



OS34 Robotics Navigation and Control (3)
Chair Jiwu Wang (Beijing Jiaotong University, China)
Co-Chair Shilong Zhen (Beijing Jiaotong University, China)

OS34-1 Research on Path Planning Algorithms of Multiple Mobile Robots in Intelligent Warehousing

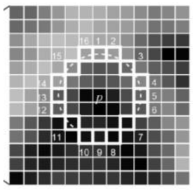
Jiwu Wang, Shilong Zheng (Beijing Jiaotong University, China)

In recent years, with the rapid development of the e-commerce industry and the increasing number of orders for commodities, the demand for intelligent automated warehousing systems has become more and more urgent. Therefore, the study of intelligent storage systems based on multiple mobile robots has proved to be a hot spot and development direction of the intelligent manufacturing industry. Among them, the collaborative path planning between multiple mobile robots has become a key issue to be solved urgently. In order to solve this problem, this paper will propose an improved A* algorithm based on the reservation table to solve the collision and traffic jams between robots and realize multi-robot collaborative path planning. The effectiveness of the algorithm is verified by simulation.

OS34-2 Research on the effectiveness of improved ORB depth estimation in monocular vision slam

Jiwu Wang, Weipeng Wan (Beijing Jiaotong University, China)

The application of monocular vision to measure the depth information of image feature points is one of the key points of monocular vision slam.Triangulation is often used for monocular vision to measure the depth information of image feature points, but in actual applications, the uncertainty of feature point matching will cause greater depth uncertainty.This article mainly optimizes the extraction and matching of ORB feature points, improves the key point extraction strategy in ORB, combines quad-tree and RANSAC to complete the feature point extraction and matching of the image, and then uses the final matching result to estimate the depth.Experiments show that the improved feature point extraction and matching technology can effectively improve the depth estimation accuracy.



OS34-3 Research on Research on Corner Detection Algorithm Based on Edge Contour in Automatic Loading Positioning

Jiwu Wang, Junwei Fu (Beijing Jiaotong University, China)

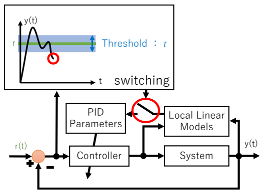
At present, automatic loading machines generally have problems such as low efficiency and low degree of flexibility. In the loading process, the use of visual methods to accurately position the cargo platform is one of the key links to improve loading automation. In the actual application process, The corner detection has the problems of low positioning accuracy, high false detection and high missed detection rate. This paper proposes a corner detection method based on the edge contour, and performs the corner detection algorithm. Optimize to improve the false detection rate and missed detection rate in the detection process, and obtain the required corner information. Experiments show that this method can effectively improve the accuracy of corner detection in the positioning link.

GS1 Control System Applications (6)
Chair Hiroaki Wagatsuma ([1]Kyushu Institute of Technology, [2]RIKEN CBS, Japan)

GS1-1 Design of local linear models using Self tuning Control System for PID Tuning According to error

Shinichi Imai (Tokyo gakugei university, Japan)

PID control is widely used in process systems represented by chemical processes and petroleum refining processes. The reason is that PID control has a simple structure. However, most of the existing systems are non-linear systems, and it is difficult to always obtain good control results with fixed PID control. Therefore, in this study, we propose a method of tuning the PID gain according to the deviation (control error) of the control result, and verify the effectiveness of this method through experiments. For self-tuning PID control using a local linear model, we propose a program that performs PID tuning only when the deviation occurs with a certain magnitude. A simulation is performed on the Hammerstein model, which is a non-linear system. As a result of the experiment, the number of PID gain changes could be significantly reduced.



GS1-2 A Systematic Analysis of the Knee Support Exoskeleton Based on Multibody Dynamics Toward Personalization with 3D Printed Spring-Damper Components

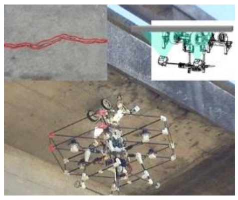
Shintaro Kasai[1], Pancho Dachikinov[1], Kohei Tanaka[1], Hiroaki Wagatsuma[1,2]
([1]Kyushu Institute of Technology, Japan; [2]RIKEN CBS, Japan)

Exoskeleton-type assistive devices have been studied a long time focusing on the universal design and a simplification for mass production, and recently the concept is extended to the personalization according to the advancement of 3D printing, which allows to embed spring-damper systems in the form of compliant mechanisms. Therefore, a sophisticated kinematic and kinetic analysis is highly important for the realization of integrative systems and theories of multibody dynamics enhance the capability to find best parameters that are suitable for target body requirements. We analyzed a knee support exoskeleton in the form of the linkage system as the rigid-body dynamics and estimated necessary spring-damper components in the system to reduce burden on joint motions, especially persons with joint dysfunctions.

GS1-3 A Drone-Based Concrete Crack Inspection System by Using Morphological Component Analysis and Sub-Pixel Width Estimation

Ankur Dixit[1], Wataru Oshiumi[1], Hiroaki Wagatsuma[1,2]
([1]Kyushu Institute of Technology, Japan; [2]RIKEN CBS, Japan)

Social infrastructure inspections were relying on human experts, and then a recent topic is a possible implementation to realize an automated inspection based on machines and sophisticated software algorithms. We have studied an advancement of Morphological Component Analysis (MCA) to apply the concrete-crack position estimation especially for submillimeter-width cracks, which are highly difficult for traditional methods to detect finely. We demonstrated a concrete crack detection from images obtained from proximity cameras attached a specialized multi-copter, by using the MCA-based crack position estimation and the linear regression-based sub-pixel width estimation. It will contribute to the actual field work not only for the concrete crack detection but also various social infrastructure inspections.



GS1-4 A Systematic Geometric Design Method of Flexible Bars Available for Personalized Knee Orthoses with Spring-Damper Functions

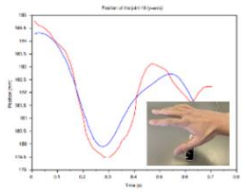
Pancho Dachikinov[1], Shintaro Kasai[1], Kohei Tanaka[1], Hiroaki Wagatsuma[1,2]
([1]Kyushu Institute of Technology, Japan; [2]RIKEN CBS, Japan)

3D printed compliant mechanisms were recently highlighted not only in the traditional way, but also in the reverse engineering of human joint supportive devices. However, a systematic design principle is unclear for solving of the issue how it can be modifiable to fit to target body requirements in the sense of the personalization. We proposed a geometrical approach providing the target design by using a morphological replacement at the concentration of stress. This concept can be verified in orthoses to improve the knee joint function by 3D printed flexible prototypes. It allows larger deformations of the bars to control joint's motion. Theoretical analysis and experiments demonstrate the flexibility and support during the flexion and extension of the knee. It implies the impact of the geometry in orthosis designs.

GS1-5 State-space modeling of fingers motion measured by the Leap Motion Controller

Ryuichi Usami, Hideyuki Tanaka (Hiroshima University, Japan)

This paper studies modeling of the motion of experts in a sport to demonstrate their motion to beginners. It deals with the dynamics of transient response from one position to another, such as a crouching start of a short-distance race. A modeling algorithm is developed to remove personal habits of experts from their motion and to present simple models to learners according to their learning stages. The algorithm is based on a step response of a state-space representation and the deterministic realization theory. It is applied to fingers motion captured by the Leap Motion Controller, and simplified models are obtained by reducing the order of the state-space representation.



GS1-6 Applicability Verification of iWakka Game to Children with Developmental Coordination Disorder

Masakazu Nomura[1], Moe Nishiya[1], Yoshifumi Morita[1], Hideo Yamagiwa[2]
([1]Nagoya Institute of Technology, Japan) ([2]Tokyo Metropolitan Tobu Medical Center, Japan)

In our previous study, we developed a testing and training device of adjustability for grasping force “iWakka”. Moreover, we developed “iWakka Game” including software for autistic patients by improving the game quality of the original one. In this paper, we investigated the applicability of iWakka Game to children with developmental coordination disorder (DCD). Based on the results of the preliminary experiment, we developed a small-size grasping body “Wakka” and improved the evaluation method for children with DCD. We applied iWakka Game to 4 children aged 5 to 9 years with DCD. As a result, all the children could complete the evaluation task with iWakka Game. The proposed evaluation method made it possible to extract the characteristics of hand dexterity.

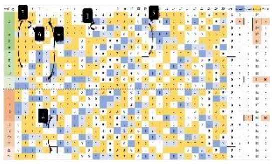
GS2 Learning Methods (5)

Chair Masato Nagayoshi (Niigata College of Nursing, Japan)

GS2-1 A basic study of how to exchange work shifts using reinforcement learning on a constructive nurse scheduling system

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

Various studies have been conducted on the nurse scheduling problem, which is the creation of a work schedule for nurses. However, for practical use, adjustments including various constraints and evaluation values are required, and the created work schedule is often not practical as it is, so many head nurses still feel burdened by creating work schedules. In this paper, we propose a work revision method using reinforcement learning for a constructive nurse scheduling system. The constructive nurse scheduling system has the characteristic of having easy to understand shift schedule creation procedures and rules because the system does not use the evaluation value for the entire shift schedule. We have confirmed the possibility of improving the quality of the shift schedule by the proposed method.



GS2-2 Developing Machine Learning and Deep Learning Models for Customer Churn Prediction in Telecomunication Industry

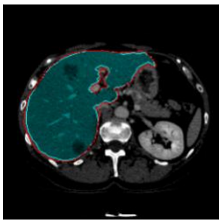
Teoh Jay Shen, Abdul Samad Shibghatullah (UCSI University, Malaysia)

Customer churn is always a significant problem and one of the biggest concerns of telecommunication companies. The companies are attempting to create and design an approach to predict customer churn. This is why determining factors that causes the customer to churn is significant. The proposed models constructed in this work apply both the machine learning and deep learning algorithms. Those models was constructed and run under the Python environment and it used an open sources dataset that are available to everyone on www.kaggle.com. This dataset contained 7043 rows of customer's data with 21 features, and it was applied in the training and testing process of the models development. These models used four different types of machine learning and deep learning algorithms, which are the Artificial Neural Network, Self-Organizing Map, Decision Tree and a hybrid model with the combination of the Self-Organizing Map and Artificial Neural network algorithms.

GS2-3 Liver Segmentation in CT Images Using Deep-Learning and 3D CRF

Shuntaro Nagano[1], Guangxu Li[2], Tohru Kamiya[1]
([1]Kyushu Institute of Technology, Japan), ([2]Tiangong University, Tianjin, China)

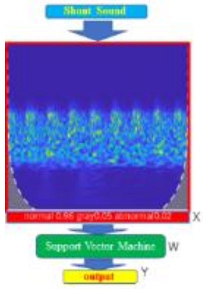
In recent years, the development of CAD systems aimed at reducing the burden on doctors and improving diagnostic accuracy has been promoted. In this paper, we propose a segmentation method for the liver site on abdominal CT images as a pretreatment for a CAD system using Dynamic CT images. The method consists of two stages. First, we segment the liver with a model based on U-net, a segmentation model using CNN. Next, the 3D CRF (Conditional Random Field) is used to make corrections that take into account the three-dimensional characteristics of the liver to improve the accuracy of segmentation. In the experiment, the accuracy was evaluated for CT images of 20 cases.



GS2-4 Judgement on Shunt Sounds from Vascular Access using YOLO Deep Learning Model

[1]Kyosuke Fujiwara, [1]Takayuki Yamamoto, [1]Lindsey Tate, [2]Kazuya Kibune, [1]Hiroki Tamura
([1]University of Miyazaki, Japan). ([2]Tokatsu Dialysis Hospital & Clinics, Japan)

A shunt is a blood vessel that connects a vein directly to an artery to allow a large volume of blood to pass through for hemodialysis. The sound of a shunt is the sound made at the junction. There is a risk of stenosis or blood clots in the shunt. The test is usually done by a doctor using an echo. However, an echocardiogram can only be performed with a special machine. Therefore, we thought that we could link the results of the echocardiography to the shunt sound and use machine learning to accurately diagnose the shunt as long as we had the shunt sound data and a computer. We used the wavelet transform to image the shunt sounds before and after stress, and trained YOLOv2tiny to classify the images with the labels "abnormal," "gray(uncertain)," and "normal. In this case, YOLOv2tiny alone did not improve the discrimination rate. Therefore, we tried to improve the discrimination rate by performing multivariate analysis using the multivariate calculated in the judgment of YOLOv2tiny. In this paper, we compared the discrimination rate of the result of YOLOv2tiny alone and the result of the combination of YOLOv2tiny and multivariate analysis, and explain how to achieve a higher discrimination rate.



GS2-5 Research of Classification of Palmprint Based on Deep Learning

Kunyu Yu, Hiroshi Matsuki (Ashikaga University, Japan)

After many years of clinical research in traditional Chinese medicine, it was found that the large thenar part of the palm was related to allergic reactions such as the asthma. This paper classifies the thenar part of the palm according to the characteristics of the wrinkles through the transfer learning in the depth learning, so as to play an assistant role in the diagnosis. The palmprint of the thenar palm can be divided into two categories (positive and negative). In this paper, the palm data collected will be preprocessed to intercept the large thenar part. The classification is mainly based on the deep learning convolution neural network, using the TensorFlow framework and inception V3 model. Simulation results show that the inception V3 model recognition algorithm has high accuracy in the classification of negative and positive palmprints.

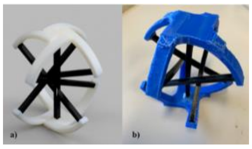
GS3 Robotics (3)

Chair Hideyuki Tanaka (Hiroshima University, Japan)

GS3-1 A Three-Dimensional Design of the Multi-material Joint System to Realize a Structural Spring-Damper Compliant Mechanism with Versatility in Engineering Fields

Pancho Dachkinov[1], Anirudha Bhattacharjee[2], Bishakh Bhattacharya[2], Hiroaki Wagatsuma[1,3]
([1]Kyushu Institute of Technology, Japan; [2]Indian Institute of Technology Kanpur, India; [3]RIKEN CBS, Japan)

Design of a 3D printed cross-spring compliant joint is an emerging topic for its multipurpose applications in various fields due to its realization from a combination of flexible materials with different mechanical properties. It performs the motion by deformation in the elastic region and is suitable for precision engineering applications and instruments. The proposed concept is a modification of a traditional cross-spring pivot, which effectively provides frictionless and wear free in-plane motion. The joint's behavior is analyzed based on a non-linear FEA simulation and the properties were investigated with various loading conditions. Compliant joints are envisaged to bring paradigm shift in the design of high-precision actuators and robotic manipulators.



GS3-2Haptic Device that Presents Sensation Corresponding to Palm on Back of Hand for Teleoperation of Robot Hand Report 5: Verification of development device specifications

Kyosuke Ushimaru, Noritaka Sato (Nagoya Institute of Technology, Japan)

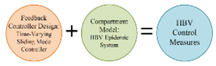
Recently, research and development of disaster response robots gained momentum. However, it is known that teleoperation of a robot hand equipped on a rescue robot is difficult. So, in this research, we propose and develop a new tactile presentation device for tele-operation of a robot hand. Proposed device is presented the tactile sensation on the back of the operator's hand. The required parameters for device development, Interval between stimulus points(i), Diameter of the stimulus point(d), Force of the stimulus point(f), were determined by subject experiments. The device was developed based on the optimized values (i, d, f) = (30mm, 6mm, 0.9kgf). A suction stimulus was used as the stimulus presentation. In this paper, we confirm that the developed device can present tactile sensation correctly and discuss the results.

.

GS3-3 HBV Epidemic Control Using Time-Varying Sliding Mode Control Method

Arsit Boonyaprapasorn[1]. Suwat Kuntanapreeda[2], Parinya Sa Ngaimsunthorn[2],
Thunyaseth Sethaput[3], Tinnakorn Kumsaen[4]
([1]Chulachomklao Royal Military Academy, Thailand)
([2]King Mongkut's University of Technology North Bangkok, Thailand)
([3]Thammasat University, Thailand) ([4]Khon Kaen University, Thailand)

Hepatitis-B (HBV) disease is one of the life-threatening diseases due to causing cirrhosis and liver cancer in the infected person. Setting the policy to control the HBV epidemic is an important issue that can be achieved by using feedback controller design procedure through the compartment mathematical model. In this article, the sliding mode controller with a time-varying sliding surface was employed to set the multiple measures control policy for controlling the HBV epidemic. The stability of the control HBV epidemic system was investigated. The simulation of the control system was conducted to confirm the feasibility of applying the time-varying sliding mode controller for setting the HBV control policy. The simulation results showed that the designed control policy can drive the target subpopulation to the desired levels. Thus, the time-varying sliding mode controller is a feasible approach to set the measures for controlling the HBV epidemic.



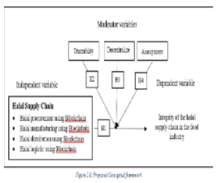
GS4 Applications (3)

Chair Noritaka Sato (Nagoya Institute of Technology, Japan)

GS4-1 Blockchain Technology for Halal Supply Chain Management

Kadeer Zulihuma, Abdul Samad Shibghatullah, Chit Su Mon
(Institute of Computer Science & Digital Innovation, UCSI University, Malaysia)

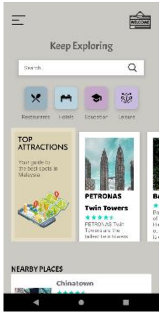
Blockchain technology (BT) is a distributed and decentralized database that store transaction information in a network. Due to providing better visibility and transparency, this technology has gained a considerable attention in the recent years. This research is carried out with the purpose of exploring the potential of blockchain technology to increase supply chain integrity in halal food industry. Three features of smart contract including traceability, decentralized, and anonymity are added to the model as moderators to explore their influence on integrity of halal supply chain. According to the survey of 251 halal supply chain practitioners in Kuala Lumpur, the proposed model was tested. The results demonstrated that BT has a positive and significant direct effect on integrity of halal supply chain, and traceability has a significantly positive influence on integrity of halal supply chain. However, decentralization and anonymity have insignificant effect on integrity of supply chain.



GS4-2 Smart Tourism Guide Application Using Location-Based Services – Go.Travel

Wong Yit Meng, Abdul Samad Bin Shibghatullah, Kasthuri Subaramaniam (UCSI University, Malaysia)

Smart tourism technology is rapidly advancing, opening new prospects for tourism, and providing higher-quality, all-around travel services to the general population. This research paper presents the creation of an assistive application that allows travellers to use tourism-related services in Malaysia to improve their travel experience. The objective of this project is to develop and enhance a Smart Tourism Guide Application, Go.Travel based on the limitations discovered. Tourists can use the proposed application to view nearby destinations, get information about them, view nearby places in Google Maps, share their stories in travel forum, read the latest articles, and save their favourite spots. This project uses an iterative development strategy. Android Studio was utilized extensively for development. It is best hoped that tourists will be able to travel more efficiently and effectively in Malaysia with the use of the proposed application.



GS4-3 Gesturenomy: Touchless Restaurant Menu Using Hand Gesture Recognition

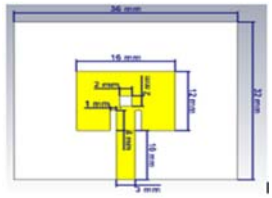
Ian Christian Susanto, Kasthuri Subaramaniam, Abdul Samad Bin Shibghatullah (UCSI University, Malaysia)

Hand gesture recognition is one of the more modern methods of human-computer interaction. However, study of its application in real world setting is sparse, especially its use in digital restaurant menus. We discuss various types of computer vision-based hand gesture recognition and decide on hand pose estimation as the method of recognition used. An analysis into the requirements show that respondents are concerned with the hygiene of touchscreen digital menus. We provide a description of the system being developed. The system passes the user acceptance tests given. This project contributes to the knowledge base by providing one case study of the use in hand gesture recognition in digital restaurant menus, as research in this application of gesture recognition is currently minimal. The authors hope that this project can open the doors for more research into this area.

N-1 Compact Ultra-Wideband Slotted Microstrip Patch Antenna for 5G, IoT and RFID Applications

Gurney Nga, Mastaneh Moakyef, Shahid Manzoor, Manickam Ramasamy (UCSI University, Malaysia)

With new emerging technologies in 5G era, the need of compact and cost-effective antenna holds significant promise for delivering contributions to society. The compact UWB omnidirectional microstrip patch antenna of 36mm × 32mm × 1.6mm for IoT, 5G and RFID applications is proposed and simulated. The proposed antenna covers frequency of 5.1256 GHz to 6.4391 GHz with UWB (bandwidth > 500MHz) characteristics. The proposed antenna has an FR4 substrate that made the antenna a cost-effective candidate in 5G era. The directivity of 5.3 dBi with gain of 3.51 dBi are obtained from the simulation results.



N-2 Drone Performance Analysis Based on SNR Factor

Gershom Phiri, Mastaneh Moakyef, Sew Sun Tiang, Wong Chin Hong (UCSI University, Malaysia)

In this paper, a depth analysis of a drone performance based on the signal to noise ratio (SNR), has been provided. The MATLAB software is used to simulate the effect of frequency, power and distance on the performance of drones. The SNR optimization has been proposed as well. The light has been shed on the drone performance in wildfire scenario.

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]			
A. Aziz	Azri	OS32-6	44/114
AB	Shahriman	OS32-3	44/113
Abbyasov	Bulat	OS17-1	37/79
		OS17-4	38/80
Abdul Latef	Tarik	OS22-3	25/91
Abdul Razak	Bushroa	OS32-1	43/112
Abe	Akinori	OS31-9	45/111
Abematsu	Yusuke	OS28-6	30/104
Aburada	Kentaro	OS10-2	26/67
		OS10-3	25/67
		OS10-4	25/67
Abu Talip	Mohamad Sofian	OS22-2	24/91
Adam	Tarmizi	OS32-7	44/115
Ahn	Jonghyun	OS28-4	29/103
		OS29-8	40/107
Ain	Mohd Fadzil	OS22-3	25/91
Akiyama	Kouichi	OS10-1	25/66
Al-Talib	Ammar A.	OS20-1	42/85
		OS20-2	42/86
		OS20-6	42/87
Ali	Darmawaty	OS32-4	44/113
Ali	Hashimah	OS32-3	44/113
		OS32-9	44/116
Alias	Muhammad Zharif	OS32-4	44/113
Aljunid Merican	Amir Feisal Merican	OS22-2	24/91
Alquran	Hiam	OS32-5	44/114
Ang	Chun Kit	OS22-1	24/90
Ang	Koon Meng	OS22-1	24/90
		OS22-4	25/92
Antonov	Vladimir	OS17-7	38/81
Aoki	Kenji	OS13-1	30/73
Arima	Masakazu	OS29-2	40/105
Arof	Hamzah	OS32-2	44/112
Asada	Taro	OS7-3	42/61
		OS7-4	42/60
Azam	S. M. Kayser	OS22-3	25/91
Azizul Hasan	Zati Hakim	OS22-2	24/91
[B]			
Baharuddin	Muhammad Nuzul Naim	OS32-7	44/115
Befu	Akane	OS27-2	39/101
Bhattacharjee	Anirudha	GS3-1	26/124
Bhattacharya	Bishakh	GS3-1	26/124
Boonyaprapasorn	Arsit	GS3-3	26/125
Bose	Rishav	IS2-1	21/48
		IS2-2	21/49
Bui	Viet	OS23-4	27/93
[C]			
Cao	Shengmin	OS12-7	46/72
Chang	Li-Yin	OS25-2	28/97
Chaw	Kam Heng	OS20-2	42/86
Chen	Linhui	OS9-6	43/65
Chen	Wangzi	OS11-3	32/68
Chen	Xiaoyan	OS11-1	32/68
		OS11-2	32/68
		OS11-3	32/68
		OS11-4	32/69
		OS11-5	33/69
		OS11-6	33/69
		OS11-7	33/70
		OS11-8	33/70
		OS12-1	45/71

		OS12-2	45/71
		OS12-3	46/71
		OS12-4	46/71
		OS12-5	46/72
		OS12-6	46/72
		OS12-7	46/72
		OS12-8	46/73
Chen	Ya-Chun	OS17-8	38/82
Chen	Yun-ju	OS2-3	22/52
Chen	Zhonggao	OS33-3	24/118
Chen	Zhihui	OS11-1	32/68
		OS12-1	45/71
Cheng	Wy-Liang	OS22-4	25/92
Cheng	Yuhui	OS3-5	41/54
		OS30-2	23/107
Chit	Su Mon	GS4-1	26/125
Chou	Ming-Tao	OS26-4	37/99
Chuang	Li-Min	OS25-3	28/97
		OS25-4	28/97
		OS25-5	28/98
Chumkamon	Sakmongkon	OS19-1	34/83
		OS19-2	34/83
		OS19-3	35/84
		OS19-4	35/84
		OS19-5	35/84
		OS19-6	35/85
		OS19-7	35/85

[D]

Dachikinov	Pancho	GS1-2	38/120
		GS1-4	39/121
		GS3-1	26/124
Dai	Fengzhi	OS3-1	40/52
		OS3-2	41/53
		OS3-3	41/53
		OS3-4	41/54
		OS3-5	41/54
		OS3-6	41/54
		OS5-1	31/56

		OS30-6	24/109
Ding	Ji-Feng	OS26-4	37/99
Dixit	Ankur	GS1-3	38/121
Du	Meng	OS12-3	46/71

[E]

Edgar A.	Martínez-García	OS17-1	37/79
		OS17-2	37/80
Elshaikh	Mohamed	OS32-9	44/116

[F]

Fang	Jiayi	OS11-9	33/70
Fawzi	Tarek	OS20-2	42/86
Fei	Shumin	OS5-1	31/56
Feng	Hao	OS11-2	32/68
Fu	Junwei	OS34-3	28/119
Fuchi	Masako	OS21-2	31/88
Fuengfusin	Ninnart	OS15-1	36/76
Fujinaga	Takuya	OS27-4	39/101
Fujisawa	Ryusuke	OS19-3	35/84
		OS19-4	35/84
Fujiwara	Kyosuke	GS2-4	34/123
Fujiwara	Masayuki	OS23-6	27/94
Fukuda	Riku	OS28-2	29/102
Fukumoto	Mana	OS7-3	42/61
Fxyma	Hiroki	OS31-4	45/110

[G]

Gamolped	Prem	OS19-1	34/83
		OS19-3	35/84
		OS19-7	35/85
Gao	Longyu	OS3-4	41/54
Gao	Yiting	OS4-4	23/56
		OS4-5	23/56
		OS8-2	41/62
Ge	Yang	OS9-4	43/64
Geiser	Sylvain	OS19-4	35/84
		OS19-5	35/84

Gobee	Suresh	OS32-2	44/112
Goda	Hiroki	OS27-1	39/100
Gong	Haoran	OS30-3	23/108
Gong	Lei	OS33-5	24/118
Guo	Chomgxu	OS9-4	43/64
Guo	Jr-Hung	OS24-2	36/95
Gurchinsky	Mikhail	OS17-6	38/81

[H]

Hamada	Daiki	OS28-1	29/102
Hamzah	Meor	OS32-8	44/115
Han	Xiaoguang	OS30-4	23/108
Hanajima	Naohiko	OS5-4	31/57
Harada	Kazuki	OS28-2	29/102
Harada	Kensuke	OS16-1	28/78
		OS16-4	29/79
Harada	Koji	OS27-1	39/100
Hashimoto	Takashi	OS23-6	27/94
Hayashi	Eiji	OS19-1	34/83
		OS19-2	34/83
		OS19-3	35/84
		OS19-4	35/84
		OS19-5	35/84
		OS19-6	35/85
		OS19-7	35/85
Hayashi	Katsuki	OS21-2	31/88
Hayashi	Yuki	OS31-5	45/110
Hayashida	Yuki	OS6-4	33/59
		OS6-5	34/60
Hirai	Hiroyasu	OS28-6	30/104
Hiramitsu	Tatsuhiro	OS16-2	29/78
		OS16-3	29/79
Hiraoka	Toru	OS1-1	22/50
Hirota	Masaharu	OS1-2	22/50
Hisamitsu	Shota	OS6-2	33/59
Ho	Chian C.	OS24-4	36/96
Ho	Yong Kang	OS32-9	44/116
Hsia	Kuo-Hsien	OS17-3	37/80
		OS17-8	38/82
		OS24-2	36/95
Hsiao	Su-Chiu	OS26-6	37/100
Hsieh	Cheng-En	OS26-1	37/98
Hsu	Chu-Ting	OS26-4	37/99
Hsu	Ling-Mei	OS26-3	37/99
Hu	Jintao	OS5-2	31/57
		OS5-3	31/57
Hu	Lintao	OS9-1	42/63
Hung	Chung-Wen	OS24-1	36/95
		OS24-5	36/96
Hussin	Eryana	OS20-3	42/86
		OS20-4	42/86
		OS20-5	42/87
		OS22-1	24/90

[I]

Ian	Christian Susanto	GS4-3	26/126
Iizumi	Shinya	OS21-5	32/89
Ikeda	Satoshi	OS13-1	30/73
		OS13-2	30/74
		OS13-3	30/74
		OS13-4	30/74
Ikeda	Tsubasa	OS21-2	31/88
Illias	H. A.	OS22-3	25/91
Imai	Shinichi	GS1-1	38/120
Inniyaka	Irmiya R.	OS28-1	29/102
Ishii	Kazuo	OS27-2	39/101
		OS27-3	39/101
		OS27-5	39/101
		OS28-1	29/102
		OS28-2	29/102
		OS28-3	29/103
		OS28-4	29/103
		OS28-5	29/103
		OS28-6	30/104
		OS29-1	39/104
		OS29-3	40/105
		OS29-4	40/105

		OS29-5	40/106
Ishikawa	Seiji	OS21-1	31/88
		OS21-4	32/89
		OS21-5	32/89
Ishikawa	Shinnosuke	OS6-5	34/60
Ito	Kota	OS21-1	31/88
Ito	Takao	PS3	30/47
		OS2-1	22/51
		OS2-2	22/51
		OS2-4	22/52
		OS13-1	30/73
		OS13-2	30/74
		OS13-3	30/74
		OS13-4	30/74
Ito	Tsutomu	OS2-1	22/51
		OS2-2	22/51
		OS2-4	22/52
		OS13-1	30/73
		OS13-2	30/74
Iwanaga	Saori	OS23-1	26/92
Iwanaga	Takuro	OS13-3	30/74

[J]

Jamaluddin	Jafferi	OS32-7	44/115
Jamaludin	Mohd Fadzil	OS32-1	43/112
Ji	Xiaojun	OS33-1	24/117
Jia	Peng	OS4-3	23/55
		OS30-5	23/108
Jia	Yingmin	OS5-2	31/57
		OS5-3	31/57
Jia	Yongnan	OS5-5	31/58
Jiang	Guan-Yu	OS24-1	36/95
Jin	Kosuke	OS14-2	27/75

[K]

Kai	Shioji	OS28-3	29/103
Kai	Xu tung	OS32-3	44/113
Kamasaka	Taketo	OS13-1	30/73
		OS13-2	30/74
Kamiya	Tohru	GS2-3	34/123
Kasai	Shintaro	GS1-2	38/120
		GS1-4	39/121
Katayama	Daigo	OS27-2	39/101
Katayama	Tetsuro	OS10-2	25/67
		OS10-3	25/67
		OS10-4	25/67
Kawai	Miku	OS31-8	45/111
Kawaguchi	Ryota	OS23-5	27/94
Kawaguchi	Yuki	OS21-4	32/89
Kawamura	Motoki	OS7-4	42/61
Kawamura	Yoji	OS31-6	45/110
Kibune	Kazuya	GS2-4	34/123
Kilin	Mikhail	OS17-5	38/81
Kimura	Kenji	OS28-6	30/104
Kita	Yoshihiro	OS1-3	22/50
		OS10-2	25/67
		OS10-3	25/67
		OS10-4	25/67
Kloomklang	Nattapat	OS19-1	34/83
Kobatake	Kanako	OS29-2	40/105
Kobayashi	Sota	OS7-2	42/61
Kohno	Takashi	OS15-2	36/77
Komori	Mochimitsu	OS21-2	31/88
Kononov	Kirill	OS17-1	37/79
Koyama	Keisuke	OS16-1	28/78
Kubo	Masao	OS23-2	26/93
		OS23-3	27/93
		OS23-4	27/93
		OS23-5	27/94
Kudo	Shunta	OS31-8	45/111
Kumagai	Motoki	OS31-1	45/109
Kumsaen	Tinnakorn	GS3-3	26/125
Kuntanapreeda	Suwat	GS3-3	26/125
Kuo	Chien-Chih	OS25-5	28/98

[L]

Lai	Hsin-Yu	OS25-1	27/96
		OS26-2	37/99

Laosiripong	Ranatchai	OS19-1	34/83
		OS19-3	35/84
Lavrenov	Roman	OS17-2	37/80
		OS17-5	38/81
Lee	Yu-Po	OS25-4	28/97
		OS25-5	28/98
Li	Chu-Fen	OS26-1	37/98
Li	Guangxu	GS2-3	34/123
Li	Hongbing	OS17-4	38/80
Li	Jianliang	OS12-6	46/72
Li	Jiaxin	OS4-2	23/55
Li	Jung-Shian	OS25-1	27/96
		OS25-2	28/97
		OS25-6	28/98
		OS26-1	37/98
Li	Mochi	OS30-2	23/107
Li	Rui	OS9-10	43/66
Li	Ruitao	OS9-8	43/65
Li	Yaxin	OS5-2	31/57
		OS5-3	31/57
Liang	Lanjun	OS33-2	24/117
Liang	Xiwen	OS11-2	32/68
		OS12-2	45/71
Lim	Wei Hong	OS22-1	24/90
		OS22-4	25/92
Lin	Geng-Yi	OS24-5	36/96
Lin	Jun-Hao	OS25-1	27/96
Lin	Tzu-Yuan	OS24-2	36/95
Lin	Wei-Min	OS26-1	37/98
Linar	Zagidullin	OS17-3	37/80
Liu	Chuan-Gang	OS25-2	28/97
Liu	Hsiu-Hao	OS25-3	28/97
Liu	Huating	OS8-1	41/62
Liu	I-Hsien	OS25-1	27/96
		OS25-2	28/97
		OS25-6	28/98
		OS26-1	37/98
Liu	Jingwen	OS33-3	24/118
Liu	Junjie	OS33-1	24/117

Liu	Qingliang	OS9-5	43/64
Liu	Qunpo	OS5-4	31/57
Liu	Sidan	OS9-10	43/66
Liu	Zilong	OS3-4	41/54
Liu	Ziyue	OS21-6	32/89
Ismail	Shahrina	OS32-5	44/114
Ishichi	Tatsuya	OS16-3	29/79
Lu	Peng	OS3-3	41/53
		OS4-3	23/55
		OS4-4	23/56
		OS4-5	23/56
		OS8-3	41/62
Lu	Shengyang	OS33-2	24/117
Lugieswaran		OS32-3	44/113
Lund	Henrik	PS1	21/47
		IS1-1	21/48
		IS1-2	21/48
		IS2-1	21/48
		IS2-2	21/49
Luo	Wendinig	OS9-3	43/64
[M]			
M. Razlan	Zuradzman	OS32-3	44/113
Magid	Evgeni	OS17-1	37/79
		OS17-3	37/80
		OS17-4	38/80
		OS17-5	38/81
		OS17-8	38/82
Maike	He	OS16-2	29/78
Makarenko	Sergey	OS17-6	38/81
		OS17-7	38/81
Mao	Zhen	OS11-7	33/70
		OS11-8	33/70
Matsuki	Hiroshi	GS2-5	34/123
Matsumura	Toshimune	OS29-4	40/105
Matsuo	Takayuki	OS29-5	40/106
Matsuno	Seigo	OS2-1	22/51
		OS2-2	22/51
		OS2-4	22/52

Md Shah	Noraisyah	OS32-2	44/112
Mehmood Khan	Hasan	OS32-7	44/115
Meshcheryakov	Roman	OS17-2	37/80
		OS17-3	37/80
Miao	Xia	OS11-4	32/69
		OS12-6	46/72
Mishima	Kouta	OS28-2	29/102
Miyamoto	Kodai	OS13-1	30/73
		OS13-2	30/74
Mizoguchi	Yusuke	OS28-2	29/102
Mochizuki	Ryuugo	OS27-3	39/101
		OS28-5	29/103
Mohamad	Mahazani	OS22-2	24/91
Mohamad	Zeeda Fatimah	OS22-2	24/91
Mohamed	Rizon	OS22-1	24/90
		OS22-4	25/92
Mohd Salleh	Mohd Faiz	OS22-2	24/91
Mokhtar	Norrima	OS32-1	43/112
		OS32-2	44/112
		OS32-3	44/113
		OS32-7	44/115
		OS32-9	44/116
Morita	Akito	OS6-1	33/58
		OS6-3	33/58
Morita	Yoshifumi	GS1-6	39/122
Motoda	Tomohiro	OS16-1	28/78
Mowshowitz	Abbe	OS19-6	35/85
Mustafa	Wan Azani	OS32-3	44/113
		OS32-6	44/114
Muto	Takafumi	OS10-3	25/67
[N]			
Nagano	Shuntaro	GS2-3	34/123
Nagayoshi	Masato	GS2-1	34/122
Nakadomari	Satoshi	OS27-2	39/101
Nakamura	Akira	OS16-4	29/79
Nakamura	Jun	OS31-2	45/109
Nakanishi	Tsuneo	OS27-4	39/101
Nakaoka	Iori	OS2-3	22/52
Nanami	Takuya	OS15-2	36/77
Ng	Joon Wen	OS20-4	42/86
Ngaimsunthorn	Parinya Sa	GS3-3	26/125
Nagato	Shusei	OS16-1	28/78
Nishida	Yuya	OS27-3	39/101
		OS27-5	39/102
		OS28-1	29/102
		OS28-2	29/102
		OS28-5	29/103
		OS29-1	39/104
		OS29-3	40/105
		OS29-4	40/105
		OS29-5	40/106
Nishiya	Moe	GS1-6	39/122
Nitta	Masuhiro	OS21-1	31/88
		OS21-3	32/88
		OS21-5	32/89
		OS21-6	32/89
		OS21-7	32/90
Niu	Hong	OS8-4	41/63
Nomura	Masakazu	GS1-6	39/122
Nordin	Hilman	OS32-1	43/112
[O]			
Oda	Tetsuya	OS1-2	22/50
Ogata	Takashi	OS31-1	45/109
		OS31-7	45/111
		OS31-8	45/111
Ohnaka	Kento	OS13-5	31/75
Okada	Ryosuke	OS6-5	34/60
Okazaki	Naonobu	OS10-2	25/67
		OS10-3	25/67
		OS10-4	25/67
Okuno	Hirotsugu	OS6-1	33/58
		OS6-2	33/59
		OS6-3	33/59
Ono	Jumpei	OS31-1	45/109
		OS31-7	45/111

		OS31-8	45/111
Ono	Yuta	OS21-3	32/88
		OS21-5	32/89
Oshiumi	Wataru	GS1-3	38/121
Othman	Mohamadariff	OS22-2	24/91
		OS22-3	25/91
Othman	S.M.	OS32-6	44/114
Otsuru	Kaori	OS31-3	45/110
Ouchi	Tomohito	OS7-1	41/60

[P]

Pagliarini	Luigi	IS1-1	21/48
		IS1-2	21/48
Pan	Li	OS22-1	24/90
		OS22-4	25/92
Park	Yousin	OS2-3	22/52
Peng	Chi-Chung	OS24-5	36/96
Peng	Yizhun	OS9-1	42/63
		OS9-2	43/63
		OS9-4	43/64
		OS9-5	43/64
		OS9-6	43/65
		OS9-7	43/65
		OS9-8	43/65
		OS9-9	43/66
		OS9-10	43/66
Petrenko	Vyacheslav	OS17-6	38/81
		OS17-7	38/81
Piyavichayanon	Chanapol	OS19-1	34/83
		OS19-3	35/84
		OS19-7	35/85

[Q]

Qasaymeh	Yazeed	OS22-3	25/91
Qin	Jiang	OS15-3	36/77

[R]

Rajagopal	Heshalini	OS32-7	44/115
Rahman	Hameedur	OS22-4	25/92
Redzuan	Muhammad R.	OS20-6	42/87
Ren	Keying	OS11-3	32/68
		OS11-4	32/69
		OS12-3	46/71
Rusli	Ahmad	OS32-8	44/115
Ryabtsev	Sergey	OS17-6	38/81
		OS17-7	38/81
Ryu	Yukiho	OS21-2	31/88

[S]

Safarova	Liaisan	OS17-4	38/80
Safin	Ramil	OS17-2	37/80
Saifizi	M.	OS32-6	44/114
Saito	Asaki	OS23-7	27/94
Saito	Masaya	OS1-4	22/51
Sakai	Daimon	OS15-2	36/77
Sakai	Nobuo	OS21-2	31/88
Sakamoto	Makoto	OS2-1	22/51
		OS2-2	22/51
		OS2-4	22/52
		OS13-1	30/73
		OS13-2	30/74
		OS13-3	30/74
		OS13-4	30/74
		OS13-5	31/75
Sakata	Koichiro	OS10-1	25/66
Sakolchik	Artur	OS17-7	38/81
Sakurai	Kiyotaka	OS1-3	22/50
Salim	Nurul Umaira	OS32-5	44/114
Sanaz Ulfitria	Riki Umami	OS24-4	36/96
Sarnin	Suzi	OS32-4	44/113
		OS32-8	44/115
Saruchi	Sarah 'Atifah	OS20-1	42/85
Sato	Hiroshi	OS23-2	26/93
		OS23-3	27/93
		OS23-4	27/93
		OS23-5	27/94
Sato	Noritaka	GS3-2	26/124

Sethaput	Thunyaseth	GS3-3	26/125
Sasaki	Masaru	OS14-1	27/75
Seki	Hiroaki	OS16-2	29/78
		OS16-3	29/79
Shen	Luqi	OS3-4	41/54
Shen	Yaxin	OS9-9	43/66
Sheng	Chang	OS4-1	23/55
Sheng	Yueqin	OS5-4	31/57
Shi	Peng	OS9-7	43/65
Shi	Songyun	OS9-4	43/64
		OS9-9	43/66
Shibanoki	Taro	OS14-1	27/75
		OS14-2	27/75
		OS14-3	27/76
		OS14-4	27/76
Shibghatullah	Abdul Samad	GS2-2	34/122
		GS4-1	26/125
		GS4-2	26/126
		GS4-3	26/126
Shimpo	Masatoshi	OS29-6	40/106
Shinomiya	Takashi	OS21-4	32/89
Shirai	Daiki	OS15-4	36/77
Shirahashi	Kanako	OS27-5	39/102
Shou	Ho-Nien	OS26-2	37/99
Solpico	Dominic B.	OS27-5	39/102
		OS28-1	29/102
		OS29-1	39/104
Song	Chau-Chung	OS24-1	36/95
		OS24-5	36/96
Song	Minghan	OS33-4	24/118
Song	Yunzhong	OS3-1	40/53
		OS5-1	31/56
Struchkov	Igor	OS17-6	38/81
Su	Bo	OS5-4	31/57
Su	Wei	OS12-4	46/71
		OS12-5	46/72
Su	Yang	OS30-6	24/109
Su	Zirui	OS11-9	33/70
Subaramaniam	Kasthuri	GS4-2	26/126
		GS4-3	26/126
Suga	Kensuke	OS10-2	25/67
Sugino	Akihiro	OS28-1	29/102
Suhaimi	Siti Qurrata Ain	OS32-6	44/114
Sun	Shengli	OS12-2	45/71
Sun	Yue	OS11-4	32/69
Suzuki	Hidekazu	OS29-7	40/106
Suzuki	Yasuhiro	OS18-1	35/82
		OS18-2	35/82
		OS18-3	35/83
Syed Taha Tahir	Syarifah Nabilah	OS22-2	24/91
Svinin	Mikhail	OS17-5	38/81
[T]			
Tabuse	Masayoshi	OS7-1	41/60
		OS7-2	42/61
		OS7-3	42/61
		OS7-4	42/61
Tadokoro	Masaya	OS14-3	27/76
Tai	Junjie	OS9-6	43/65
Takagi	Tomohiko	OS10-1	25/66
Takaki	Ryosuke	OS1-1	22/50
Takei	Amane	OS13-1	30/73
		OS13-2	30/74
		OS13-5	31/75
Takemura	Yasunori	OS28-4	29/103
Tamaki	Hisashi	GS2-1	34/122
Tamukoh	Hakaru	OS15-1	36/76
		OS15-5	36/78
Tamura	Hiroki	GS2-4	34/123
Tan	Joo Kooi	OS21-1	31/88
		OS21-3	32/88
		OS21-4	32/89
		OS21-5	32/89
		OS21-6	32/89
Tanaka	Hideyuki	GS1-5	39/121
Tanaka	Kohei	GS1-2	38/120

		GS1-4	39/121
Tanaka	Rikuto	OS28-1	29/102
		OS29-3	40/105
Tanaka	Yoshiki	OS27-5	39/102
		OS28-2	29/102
Tanikawa	Kiyotaka	OS1-4	22/51
Tasaki	Tetsufumi	OS6-5	34/60
Tate	Lindsey	GS2-4	34/123
Tebueva	Fariza	OS17-6	38/81
		OS17-7	38/81
Tengku Mohmed Noor Izam	Tengku Faiz	OS22-2	24/91
Teoh	Jay Shen	GS2-2	34/122
Tian	Mingjuan	OS30-3	23/108
Tiang	Sew Sun	OS22-1	24/90
		OS22-4	25/92
Ting	Hui Chong	OS22-1	24/90
Ting	Kee Yuan	OS20-1	42/85
Tominaga	Ayumu	OS19-3	35/84
		OS19-4	35/84
		OS19-5	35/84
Tominaga	Moeko	OS27-5	39/102
		OS28-4	29/103
Tomokawa	Takumi	OS19-4	35/84
		OS19-5	35/84
Toshimori	Tomoyasu	OS13-4	30/74
Tonooka	Hideyuki	OS14-1	27/75
		OS14-3	27/76
		OS14-4	27/76
Tsai	Wen Jung	OS26-5	37/100
Tsai	Yao-Chu	OS25-6	28/98
Tseng	Yu-Hsuan	OS24-1	36/95
Tsoy	Tatyana	OS17-1	37/79
		OS17-3	37/80
		OS17-4	38/80
Tsuji	Tomofumi	OS19-1	34/83
		OS19-3	35/84
		OS19-7	35/85
Tsuji	Tokuo	OS16-2	29/78
		OS16-3	29/79
Tsuji	Toshio	OS14-1	27/75
[U]			
Ueda	Kazuki	OS1-3	22/50
Ueda	Takahiro	OS10-4	25/67
Uemura	Yuuichiro	OS29-3	40/105
Ul Husnain	Anees		
Usami	Ryuichi	GS1-5	39/121
Ushimaru	Kyosuke	GS3-2	26/124
[W]			
W. Muhamad	Wan Norsyafizan	OS32-4	44/113
		OS32-8	44/115
Wada	Koichi	OS27-2	39/101
Wagatsuma	Hiroaki	GS1-2	38/120
		GS1-3	38/121
		GS1-4	39/121
		GS3-1	26/124
Wan	AzaniMustafa	OS32-5	44/114
Wan	Khairunizam	OS32-2	44/112
		OS32-3	44/113
		OS32-5	44/114
		OS32-6	44/114
		OS32-9	44/116
Wan	Weipeng	OS34-2	28/119
Wan	Weiwei	OS16-1	28/78
Wan Mahiyiddin	Wan Amirul	OS32-7	44/115
Wang	Chun-Chieh	OS24-3	36/95
Wang	Dechao	OS9-5	43/64
Wang	Di	OS12-8	46/73
Wang	Hucheng	OS30-1	23/107
Wang	Jinliang	OS30-4	23/108
Wang	Jiwu	OS34-1	28/119
		OS34-2	28/119
		OS34-3	28/119
Wang	Qian	OS12-7	46/72
		OS12-8	46/73

Wang	Sicong	OS9-2	43/63
Wang	Shiming	OS12-1	45/71
Wang	ZiChen	OS11-4	32/69
		OS12-8	46/73
Watanabe	Keisuke	OS27-1	39/100
		OS29-6	40/106
		OS29-7	40/106
Watanabe	Yoshiyasu	OS29-7	40/106
Wei	Jiawei	OS30-4	23/108
Wen	Haokang	OS4-1	23/55
		OS8-3	41/62
Wen	Xiran	OS11-9	33/70
Wong	Chee Ming	OS20-5	42/87
Wong	Chin Hong	OS22-4	25/92
Wong	Jie Jian	OS20-3	42/86
Wong	Yit Meng	GS4-2	26/126
Wu	Dashuo	OS11-5	33/69

[X]

Xi	Guangyong	OS12-5	46/72
Xiang	Yande	OS30-5	23/108
Xiang	Yawned	OS30-3	23/108
Xiao	Huimin	OS5-1	31/56
Xiao	Zhiqing	OS30-6	24/109
Xie	Yong	OS5-5	31/58
Xu	Changhua	OS21-6	32/89
Xu	Jiarun	OS3-4	41/54
Xu	Nenghua	OS11-2	32/68
Xu	Qiqi	OS11-9	33/70
Xu	Zhuofan	OS9-8	43/65

[Y]

Yagi	Tetsuya	OS6-5	34/60
Yaji	Kaede	OS23-3	27/93
Yamaba	Hisaaki	OS10-2	25/67
		OS10-3	25/67
		OS10-4	25/67
Yamabe	Takayuki	OS16-3	29/79
Yamada	Chikako	OS27-2	39/101
Yamagiwa	Hideo	GS1-6	39/122
Yamaguchi	Akihiro	OS23-7	27/94
Yamaji	Yuuki	OS6-3	33/59
Yamamoto	Hidehiko	PS2	21/47
Yamamoto	Hyoga	OS29-5	40/106
Yamamoto	Takayuki	GS2-4	34/123
Yamamoto	Yusuke	OS28-2	29/102
Yamasaki	Haruna	OS19-2	34/83
Yamawaki	Akira	OS15-3	36/77
		OS15-4	36/77
Yamazaki	Yugo	OS14-4	27/76
Yan	Xiaoning	OS11-1	32/68
		OS11-2	32/68
		OS11-3	32/68
		OS11-4	32/69
		OS11-5	33/69
		OS11-6	33/69
		OS11-7	33/70
		OS11-8	33/70
Yan	Yihan	OS9-9	43/66
Yang	Bai	OS17-2	37/80
		OS17-5	38/81
Yang	Kunzhi	OS11-5	33/69
		OS11-6	33/69
Yang	Xinpeng	OS9-10	43/66
Yao	Feng	OS33-2	24/117
Yao	Pengyu	OS9-5	43/64
Yao	Shuen-Huei	OS26-5	37/100
Yeh	Cheng Chung	OS26-5	37/100
Yoshimoto	Yuma	OS15-5	36/78
Yoshitomi	Yasunari	OS7-3	42/61
		OS7-4	42/61
Yasukawa	Shinsuke	OS6-4	33/59
		OS27-2	39/101
		OS28-3	29/103
		OS28-5	29/103
		OS29-8	40/107
Yin	Junhui	OS9-10	43/66
Yokkampon	Umaporn	OS19-6	35/85

		OS19-7	35/85
Yokomichi	Masahiro	OS13-2	30/74
Yokota	Hiroyuki	OS29-8	40/107
Yong	Jonathan	OS20-2	42/86
Yotsumoto	Takaaki	OS21-3	32/88
Yotsumoto	Takanori	OS6-4	33/59
Yu	Kunyu	GS2-5	34/123
Yucheng		OS9-5	43/64

[Z]

Zeng	Xin-Ying	OS17-8	38/82
Zhang	Chuang	OS3-2	40/53
Zhang	Dongyang	OS11-3	32/68
		OS11-6	33/69
Zhang	Jinping	OS33-1	24/117
Zhang	Teng	OS30-3	23/108
Znang	Tianyi	OS3-6	41/54
		OS8-1	41/62
		OS8-2	41/62
Zhang	Weicun	OS5-1	31/56
		OS5-5	31/58
Zhao	FuChen	OS9-3	43/64
Zhao	Jichao	OS3-1	40/53
		OS3-2	40/53
Zhao	Huailin	OS33-1	24/117
		OS33-2	24/117
Zhao	Shixiang	OS30-3	23/108
Zhao	Yuwei	OS11-7	33/70
Zheng	Shilong	OS34-1	28/119
Zheng	Shuangwu	OS11-1	32/68
		OS11-8	33/70
Zhong	Binfu	OS9-8	43/65
Zhou	Fangbo	OS33-2	24/117
Zhou	Ling	OS30-3	23/108
Zhu	Tao	OS30-6	24/109
Zhu	Yurong	OS33-1	24/117
Zulihuma	Kadeer	GS4-1	26/125

Road To Cyber Physical Factory
(Application Examples of Intelligent Factory and its Technology)

*Hidehiko Yamamoto
Department of Engineering, Gifu University,

1-1, Yanagido, Gifu, 501-1193, Japan yam-h@gifu-u.ac.jp

Abstract

The future factory by using Autonomous System, we call Autonomous Decentralized Flexible Manufacturing Systems, is presented. The virtual factory and several kinds of simulations for production systems and scheduling problems is presented. The simulations application examples for automobile production lines are presented. IoT production and Cyber Physical Factory which is the near future manufacturing model are presented. Its application example for automobile parts production is also presented.

Keywords: intelligent manufacturing, cyber physical factory, GA, scheduling, process planning, robot assembly .

1. Introduction

Since 1980, Computerized machine tools and robots have been developed. This paper describes the history of the Intelligent manufacturing systems and their technology including Artificial Intelligence and GA [1][2]. The technology are as follows. The future factory by using Autonomous System, we call Autonomous Decentralized Flexible Manufacturing Systems, is presented. The virtual factory and several kinds of simulations for production systems and scheduling problems is presented. The simulations application examples for automobile production lines are presented. IoT production and Cyber Physical Factory which is the near future manufacturing model are presented. Its application example for automobile parts production is also presented.

2. Manufacturing history

Figure1 shows the history of Manufacturing automation or intelligent technology.

The first step of Manufacturing Automation is Mass Production. The mass production has a kind of conveyer and each operator does his or her jobs of the processes flowed from the conveyer. The first mass production started in the world was Ford motor company.

The second step of Manufacturing Automation is the development of Numerical control machine tools (NC Machine). By the NC machine, variety of parts can be manufactured without setup changes.

The third step of Manufacturing Automation is the development of Transfer production line. The line is based on the mass production and some kinds of variety parts can be manufactured because of computer control and NC machine.

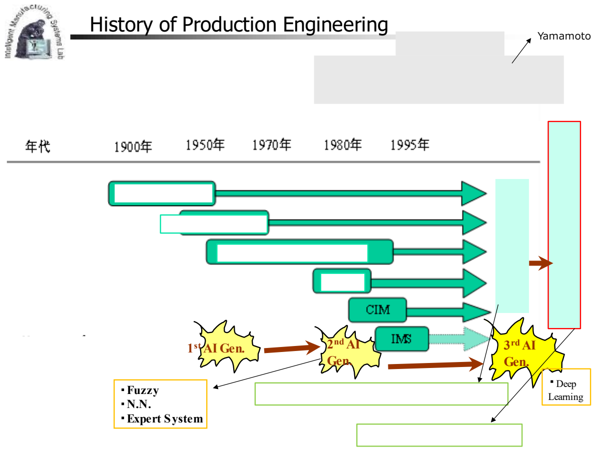


Figure1. Factory automation history

The fourth step of Manufacturing Automation is the development of Flexible Manufacturing Systems (FMS). FMS does not include the conveyer but include Automated Guided Vehicles (AGV), Machining Centers that can operate some processes and Automated warehouse.

The fifth step of Manufacturing Automation is the development of Computer Integrated Manufacturing (CIM). CIM is the computerized technology of not in factory but in all departments of industrial company such as departments of sales, product development, experiment department and product plan department.

The sixth step of Manufacturing Automation is Intelligent Manufacturing Systems (IMS). IMS is called as the 21st century technology.

3. Autonomous Decentralized Factory

One of the IMS technology is an autonomous decentralized factory (ADF) [3][4]. ADF does not need a computer host. ADF has some elements (agents) that represents NC machine tools, Machining Centers, AGVs and warehouse like Figure 2. Without orders of a computer host, each agent includes each knowledge and each agent exchanges other agents' knowledge. Thanks to the knowledge exchanges, ADF can efficiently operate productions. The functions by the knowledge exchanges are negotiations and cooperation. The functions order each agent which AGV takes and sends it to which machine tool. In this way, ADF can give each agent efficient orders and a total factory output efficiency can keep high without production scheduling system. ADF needs a real-time decision making. To carry out the real-time decision making corresponding real-time scheduling, Reasoning to Anticipate Future that the author proposed is necessary.

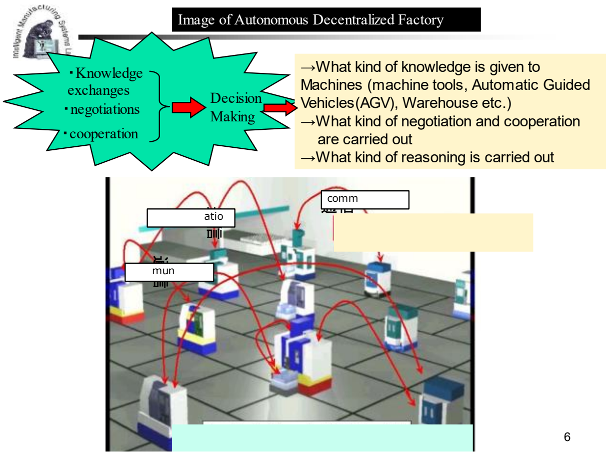


Figure 2. Autonomous decentralized factory

4. AGV and its Mind

When ADF carries out, unexpected problems happen. One of the problems is AGVs collisions or AGV interference.

In order to avoid the problems, the AGV moving control needs some kinds of if-then rules. Although, the number of AGVs are small, the rules take an effect to avoid AGVs interference. When the number of the AGVS moving in the factory, unexpected situations happen and the rules cannot be beforehand expressed.

Human beings can avoid the unexpected happening not always using rules but using a mind. I proposed the AGV mind and applied it into the ADF.

The AGV mind has two kinds of mind, one is the modest mind and the other is the arrogant mind [5][6]. The AGV that has the modest mind gives a way to other AGVs. The AGV that has the arrogant mind forcibly moves (See Figure 3). AGVs sometimes exchange the two minds just like humans do.

By using the AGVs with a mind and AGVs that do not have a mind, many virtual production simulations were carried out.

The results show that the collision (interference) number of AGVs with a mind are 0. The production outputs of using AGVs with a mind are bigger than those of using AGVs that have no mind (See Figure 4).

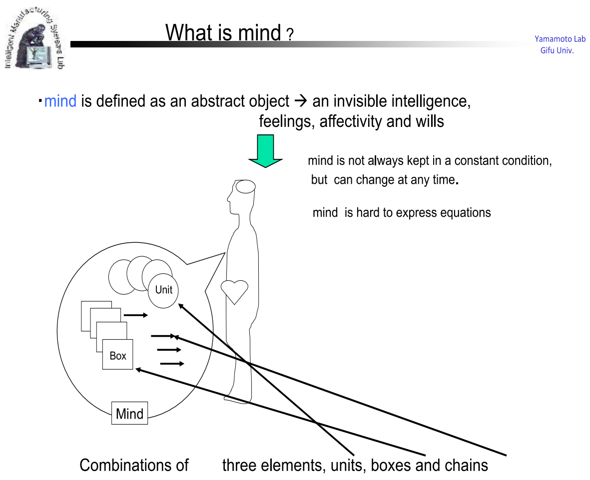


Figure 3. AGV mind

Results

Conditions	1-1	1-2	1-3	2-1	2-2	2-3	3-1	3-2	3-3
MC average efficiency (%)	35.3 21.2	22.9 41.7	25.8 46.8	11.7 19.3	11.4 25.7	16.1 30.7	9.3 16.7	11.6 22.0	13.6 26.9
Collisions avoidance number	161 0	396 0	423 0	105 0	276 0	465 0	123 0	234 0	477 0
Outputs	93 264	86 311	141 349	48 170	45 225	56 272	36 146	48 195	61 235

Without mind / With mind

Figure4. Simulation results for AGV mind

5. Parts layout Decision system by GA

In an assembly line, we need parts supply for an assembly line. To quick parts supply, we need to decide the efficient parts stock area. The efficient means how efficient the operator can pick up the parts from the parts shelves. Because of it, we need to decide which part is put on which shelf like Figure 5 and 6.

In order to solve the problem, the system to find the efficient parts layout system by using GA. We call the system Virtual Assembly Cell-production system (VACS) (See Figure 7)

Figure 5. Parts shelf

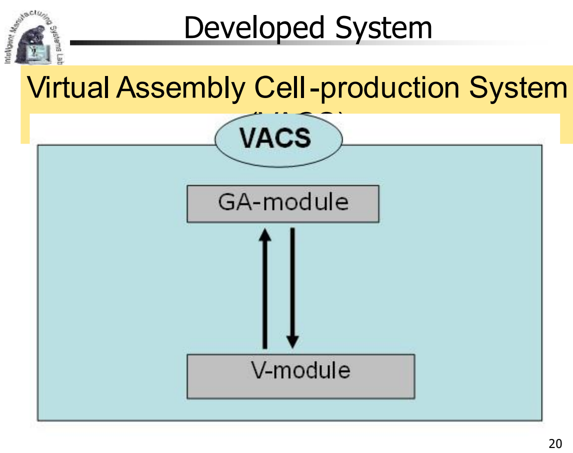
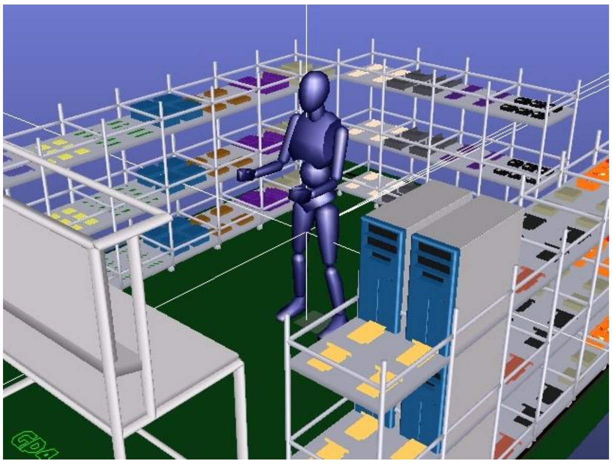


Figure 6. Parts layout and cell production



Fugure7. Outline of VACS

6. Cyber physical factory

Recently, IoT production has been developed. However, The characteristic of the IoT production is to gather information from the production line and machine tools, to analyze the gathered information and to return the analyzed information results. The period between the time og gathering information and the time of returning the results need a few days. That means the IoT production does not carry out the real-time scheduling.

To carry out the real-time scheduling, we need to develop the cyber physical factory.

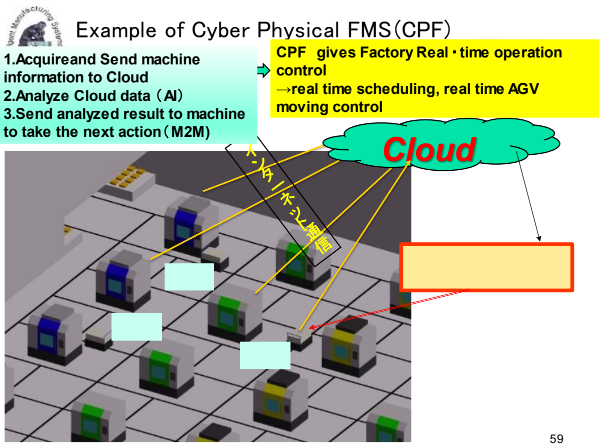


Figure 8. Outline of cyber physical factory

To carry out the real-time scheduling, it is important to use RAF mentioned above. The gathered information is uploaded into the cloud, RAF in the cloud finds the next AGV action and to send it to the agent. The process is carried out immediately.

7. Conclusions

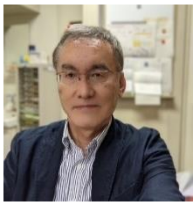
This paper describes the history of the automation manufacturing from the 1900 to the near future. The next important of the industrial companies need Intelligent technology and the cyber physical factory.

References

1. H. Yamamoto, "One-by-One Parts Input Method by Off-Line Production Simulator with GA", European Journal of Automation, Hermes Science Publications, Eds. A. Artiba, 1173-1186 (2000)
2. H. Yamamoto and D. Moldovan, "Robot Path Planning for Visiting FA Working Points by Obstacle Avoiding using GA", Proceedings of the 6th International Conference on Artificial Life and Robotics (AROB 6th '01), Tokyo, 524-527 (2001).
3. H. Yamamoto and E. Marui, "Decentralized Autonomous FMS Control by Hypothetical Reasoning including Discrete Simulator", Proceedings of The Fourteenth International Conference on Industrial & Engineering Applications of Artificial Intelligence & Expert Systems (IEA/AIE-2001), Lecture Notes in the Computer Science/Lecture Notes in Artificial Intelligence series, Springer-Verlag (eds. L. Monostori, J. Vancza and M. Ali), Budapest, 571-581(2001).
4. H. Yamamoto and E. MARUI, "Intelligent Communication Between Agents of Autonomous Decentralized FMS", Proceedings of 2005 IEEE International Symposium on Computational Intelligence in Robotics and Automation (CD-ROM) , Espoo Finland, June 27-30 (2005).
5. H. Yamamoto, K. Hiroyuki, T. Yamada and M. A. Jamali, "Moving Robots' Mind of Autonomous Decentralized FMS and Mind Change Control", The Journal of Artificial Life and Robotics, Vol. 14, No.1, 34-38 (2009), Springer..
6. H. Yamamoto and T. Yamada, "Cooperation of AGVs' Head-on Collision Avoidance by Knowledge Exchange in Autonomous Decentralized FMS", Lecture Notes in Computer Science,14th International Conference, KES 2010Cardiff, UK, September 8-10, 2010 Proceedings

Authors Introduction

Dr. Hidehiko Yamamoto



He is Specially Appointed Professor, Gifu University.

Education:

• 1980 year Master's Degree of Mechanical Engineering (Nagoya Institute of Technology)

• 1991 year Docter's Degree of Engineering (Nagoya Institute of Technology).

Employment:

1980-1992: Toyota Industrial Corporation Ltd.
1992-2000: Associate Professor, Wakayama University, Japan.
2000-2021: Professor, Gifu University, Japan.
2014: The Fellow of The Japan Society of Mechanical Engineers
2021-present: Specially Appointed Professor, Gifu University, Japan.
2021: Honorary Professor of Gifu University.

Robotics for Growing Life

Henrik Hautop Lund

[1] Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
[2] University College Copenhagen (KP), 1799 Copenhagen V, Denmark
E-mail: hhl@playware.dtu.dk

Abstract

We present a direction of artificial life robotics in which we use robotics to control the growth of real, natural life. The concept of using robotics to grow life presents itself as a proposal for a potential sustainable solution for food production, allowing an optimization of food quality and outcome. The design of these artificial life robotic systems allow for urban farming where food production happens in the close vicinity to the consumer, avoiding the long transportation of food to the consumer. We illustrate this concept with our development of the GrowBot, which is a tabletop size robotic green house for growing edible food plant. The GrowBots use sensors such as humidity, CO_2, temperature, water level and camera sensors, and actuators such as full spectrum LEDs, IR LEDs, UV LEDs, fertilizer and water pumps, air change and air circulation fan. The software acts as recipes for the plant growth in the robotic greenhouse adjusting the environmental condition for the growth of the living plants such as salad, parsley and basil. Changing the recipes, one may experiment and investigate easily to search for optimization for volume production, taste, etc. We illustrate this concept and implementation of artificial life robotics with the growth of Italian basil, *Ocimum basilicum*.

Keywords: Artificial Life, Robotics, Biology, Urban Farming.

1. Introduction

Artificial life is defined as life-as-it-could-be, as compared to life-as-we-know-it [1]. Within artificial life, computer simulations are used to study the development of life in many different forms. In order to study such development in the real, physical world, artificial life robotics developed as a field in which the simulated artificial life organisms were transferred to the physical robots. Artificial life robotics includes biomimetic robotics, in which biology is used as inspiration for creating robots, which in turn can be used as experimental tools to verify biological hypotheses. For example, Lund et al. [2] recorded male cricket *Gryllus bimaculatus* calling song, built a robot with the hypothesised control mechanism for phonotaxis behaviour and verified that it could account for the behaviour by observing the robot doing phonotaxis to the real cricket song. The robot model was implemented as a simple spiking neural network that is less complex than the controllers traditionally hypothesized for cricket phonotaxis and syllable rate preference [3]. Thereby, the robot experiment could be used as an existence proof that such a simple neural network model is enough to account for the complex cricket behavior. The interaction was done between the robot and the live animals, i.e. live crickets / recorded live cricket songs. Also, artificial life robotics is used to study life development in terms of ontogenetic and phylogenetic development, i.e. development attributed to individual life-time experiences with the environment (ontogenesis) and the development attributed to evolutionary process of species over generations (phylogenesis). Ontogenesis is studied with different robot learning methods, and phylogenesis is studied with evolutionary robotics.

In this way, artificial life robotics is used to study the development of life. With the present work, we want to open a direction of artificial life robotics in which artificial life robotics is used not only to study the development of life, but also as a tool to directly grow life. It is not a simulated or virtual representation of life which develops, but it is the actual real life which develops. This allows for new studies and insight into the growth of life and the environmental influence on the growth of life. The robotic system can be used to control environmental conditions, which influence the growth of life. In this way, one can parameterize the growth conditions and with the robotic system investigate how to optimize this parameterization.

The optimization can, for instance, be for height of living plants, canopy size of living plants, taste of harvested leaves and fruits, nutritious value of living plants and insects, energy expense, etc. Applications of this may lead towards an optimization of green food production, towards food production near the consumer in urban areas, and towards food production in otherwise environmental hostile locations for food production (polluted areas, deserts, arctic, etc.)

In this work, we will present an example of robotics for growing life with the development of the GrowBot, which is a tabletop robotic greenhouse aimed for educational use in schools.

2. GrowBot Hardware and Software

The first iterations of our development were inspired by the MIT Media Lab OpenAg food computer [4, 5], before we made extensive changes. The OpenAg met criticism that it was difficult to make the food computer work, to make plants grow, to achieve robustness, and to make the food computers work in school settings [6], and the project was shut down in 2020. In order to address these inadequacies, we used a comprehensive system design method from the area of playware [7] to develop the GrowBots to bring the smart farming approach into use in the educational sector.

The GrowBot is designed as a small, robotized greenhouse for hydroponic farming. The GrowBot measures 450mm x 450mm x 600mm, and consists of a 150mm top part, which holds the electronics – see Table 1 – and the growth-chamber. There is a water tray at the bottom of the chamber. It has a lid with 4 x 4 holes, in which the plants can grow. On the top, there is a 5" touchscreen display and a connector for a micro:bit [8]. See Table 1 for the technical specifications.

Table 1. Main components of the GrowBot.

Description	**Specification**
Processor	Raspberry Pi 4 Model B 2GB
I/O interface processor	16Mhz Arduino NANO V3.0 Atmega328 UNO IO Shield
Wireless communication	Raspberry Pi 4G/LTE with base HAT and antenna on the Quectel EC25 Mini PCle 4G/LTE Module
Display	5" Raspberry Pi LCD touchscreen with 800*480 resolution and 108×64.8mm display area (SKU DFR0550)
Power	12 V Power supply and a 3.3V linear regulator (ADP3338)
PWM controller	PCA9685PW PWM Controller
Air circulation fans	405FH DC Axial fans 9m3/h
LED drivers	LDH45 LED driver
Full spectrum light	LM301B white LEDs
IR light	XLamp XP-E2 IR 730nm
UV light	XLamp XP-E UV 458nm
Heater	Ceramic (PTC) air heating element
PWM Controller	PCA9685PW, PWM Controller
Pumps	DC 12V Air Diaphragm Pump 1.5-2 L/Min
Aeration	Oxygen disk
Humidity and temperature sensor	DHT 22 relative humidity and temperature sensor
CO_2 sensor	MHZ16 infrared CO_2 sensor
Water level sensor	DFROBOT SEN0193 capacity soil moisture sensor
Camera	Raspberry Pi 4 camera 150 degree wide angle 5 megapixels (OV5647 sensor)
micro:bit	BBC micro:bit board

Sensors in the GrowBot include humidity sensor, CO_2 sensor, temperature sensor, water level sensor, and RGB camera. Actuators include full spectrum lights, IR lights, and UV lights, fertilizer pump, water pump, air pump to pump air to the oxygen disk placed in the water tray, air change pump, and air circulation fan.

Fig. 1. The GrowBot

We have designed the GrowBot so that it can be programmed easily with simple graphical block programming. This can happen (i) programming a micro:bit, (ii) programming on the GrowBot display, or (iii) programming on a remote web-interface. With the simple graphical block programming, one can program the GrowBot to control the different environmental conditions, which may influence the growth of the plants, such as daily light cycle, light color, temperature, nutrient level, etc. A web-interface allows remote monitoring of sensory values and camera images.

With the programming of the GrowBot for the experimentation with growth of edible plants such as parsley, salad, and basil, the students can work on optimizing the environmental conditions in the GrowBot for the growth of the plants, e.g. in terms of plant height, leaf size, health of the plants, nutritional value, etc. Their programs can be viewed as recipes for the growth of specific plant species, and the students can exchange these recipes with each other.

3. Experimentation

We performed a number of experiments with the GrowBots to validate the functioning before bringing the GrowBots to schools. In one experiment, we planted seeds for parsley, basil, red lettuce and green lettuce in nine GrowBots, and ran the experiment of 43 days. All GrowBots had a light cycle of 18 hours of light and 6 hours of no light during each 24 hour period. They were programmed to different colored light. Figure 2 shows the reading of the sensory values over a 2-day period from the web-interface. It can be observed that during the night when lights are turned off, the temperature falls, while the humidity and CO_2 rises.

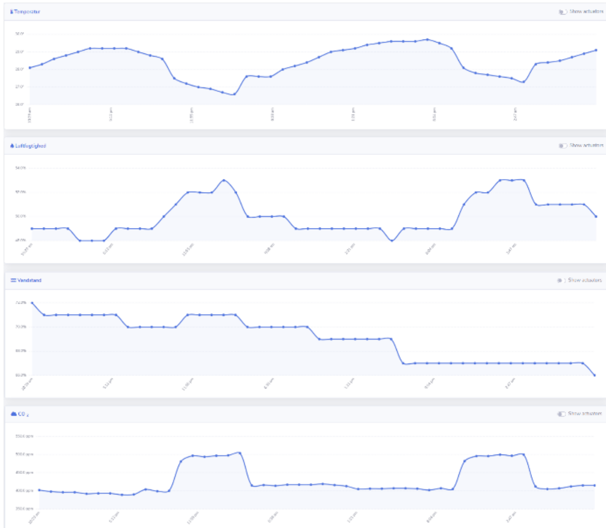


Fig. 2. The web-monitoring of four sensory values over 2 days of experimentation (temperature, humidity, water level, CO_2). It can be noticed that during the night when lights are turned off, the temperature falls, the humidity and CO_2 rises.

Figure 3 shows an example of some of the basil plants harvested at the end of the experiment. Roots were carefully dissected, and the plant weight, height and leaves sizes were measured in the experiment.

Fig. 3. Harvested basil plants from one of the nine Growbots.

With the robotic greenhouse and programming that allows for the careful control of the environmental conditions for growth of life, it is possible to perform numerous experiments that can give us further insight into the growth of life. In the experiment mentioned above, it was found that blue light condition (458nm) gave higher yield than red light condition (720nm) [9].

Figure 4 shows an example from another experiment, in which we investigated the consequences of stress on plant growth with heat stress and nutrient deficiencies. Here, the temperature was set to 29ºC - 35ºC. The nutrient composition was made to lack different chemicals in different experiments, e.g. lack of magnesium (Mg), potassium (K) or nitrogen (N). In the example in figure 4, it is clearly visible how the heat stress and potassium deficient nutrient combination leads to conditions on the leaves such as necrotic spots and chlorosis after merely 6 days [10, 11].

Fig. 4. Heat and potassium stress experiment at day 1 (left) and day 6 (right). Necrotic spots and chlorosis is clearly visible on many leaves after 6 days (right).

4. Discussion and Conclusion

This direction of artificial life robotics uses robotics to control the growth of real, natural life. In the GrowBots, we can for instance control the growth of plants to optimize for food production. The GrowBot robotic system allows us to parameterize life growth conditions for the natural plants through the recipes in terms of temperature, humidity, day light cycle, color of grow light, fertilizing, etc. However, this novel artificial life robotics concept can as well be used to optimize the growth of life in the form of drosophila farms, butterfly farms, and worm farms. Indeed, these can be insect food farms, e.g. for larvae mealworms, crickets, locust, and ants, optimizing the nutritional value and volume outcome by controlling the variables such as temperature, humidity, feed, water sources, and lighting depending on the insect species. Further, the artificial life robotics concept lends itself to studies within synthetic biology and biomedical engineering, e.g. biopharmaceutical production.

Acknowledgements

The authors would like to thank project partners at DTU Food, DTU Skylab, and KP. Thanks to the funding bodies Nordea Foundation and Novo Nordisk Foundation.

References

1. Langton, C. G. (1989). Artificial life. In Proceedings of the Santa Fe Institute. Studies in the sciences of complexity. Vol 6. Addison-Wesley.
2. Lund, H. H., Webb, B., & Hallam, J. (1997). A robot attracted to the cricket species gryllus bimaculatus. In 4th European Conference on Artificial Life (pp. 246-255).
3. Lund, H.H., Webb, B. & Hallam, J. (1998). Physical and temporal scaling considerations in a robot model of cricket calling song preference. Artificial Life, 4(1), 95-107.
4. C. Harper and M. Siller, "OpenAG: A Globally Distributed Network of Food Computing," in IEEE Pervasive Computing, vol. 14, no. 4, pp. 24-27, Oct.-Dec. 2015.
5. Ferrer, E. C., Rye, J., Brander, G., Savas, T., Chambers, D., England, H., & Harper, C. (2018). Personal Food Computer: A new device for controlled-environment agriculture. In Proceedings of the Future Technologies Conference (pp. 1077-1096). Springer.
6. Goldstein, H. (2019, 24 Oct 2019 (updated 3 Feb 2020)). MIT Media Lab Scientist Used Syrian Refugees to Tout Food Computers That Didn't Work, IEEE Spectrum.
7. Lund, H. H. (2014). Playware Research – Methodological Considerations. Journal of Robotics, Networks and Artificial Life, 1(1), 23-27.
8. http://www.microbit.org

Authors Introduction

Henrik Hautop Lund

Professor, Ph.D. Henrik Hautop Lund, Center for Playware, Technical University of Denmark, is World Champion in RoboCup Humanoids Freestyle 2002, and has more than 200 scientific publications and 5000 citations. He was awarded the Most Outstanding Healthcare Innovator in the World in 2019. Over the last year, he received international awards in Tokyo, Singapore, and London. He has developed shape-shifting modular robots, presented to the emperor of Japan, and has collaborated closely on robotics and AI with companies like LEGO, Kompan, BandaiNamco, Microsoft, Mizuno. He is the inventor of the Moto Tiles (www.moto-tiles.com), which are used by seniors all around the world.

Tiler. A physical to virtual control system implementing an art-based game

Luigi Pagliarini[1,2], Henrik Hautop Lund[1]

[1] Centre for Playware, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
[2] Academy of Fine Arts of Macerata, P.za Vittorio Veneto 5 - 405150, Macerata, Italy

l.pagliarini@abamc.it

Abstract

We hereby present Tiler, an art-based game where a virtual world made of tiles and controlled through a set of electronic cubes where players can/should gather the decoration of a floor based on aesthetical criteria. Such a tool allows projecting and designing bi-dimensional shapes by physically manipulating tri-dimensional objects. Besides that, Tiler, is an application that brings to life a clear example on how we can achieve new ways of interfacing the physical world with virtual ones. In the following article we introduce the logical and technical aspects of this real-to-virtual interface and show its potential applications in different fields.

Keywords: Interactive Art, Multisensory Rooms, Playware, Games.

1. Introduction

In this paper we report about *Tiler,* a research prototype for a Virtual-To-Real tool that, implementing the knowledge described in Pagliarini's Code [1], allows users to get engaged in a sort of art-based game and allows them by manipulating with their own hands a set of electronic cubes so to obtain aesthetical results either under the visual art and musical point of view. Strong of our long-lasting experience on such modular and interactive electronics artifacts as Moto Tiles [2] and Music Cubes [3], as well as interactive art [4] and gaming [5], we began to develop Tiler, a complex but easy to use device, by which participants might evolve dynamic geometry and might insert sound effects within it.

Basically, thought to be used in multisensory rooms, the whole project is meant to provide a dynamic, interactive and smart manipulation of physical object in order to obtain personalized aesthetical results while consciously gaming, and unconsciously eliciting and pushing on different cognitive skills, such as attention, stimuli responsiveness, memory, etc. Therefore, the Tiler project, itself, might be reviewed and reported under different angles and perspectives, since it can be considered either entertainment or therapy or, under a scientific point of view, a step forward in the research world of human oriented interfaces.

The choice of using an art-oriented output was taken to enlarge the range of the target of users as much as possible, so to be inclusive in respect of all the possible ages and gender and, possibly, generate a cross generational interaction media.

Despite of the fact that, at the very moment, the Tiler algorithm and related electronics are both prototypical (i.e. many possible variables as interface size, gaming speed, etc. are still to be properly tested), we are engineering the gaming experience so that it should be noise and errors resistant. We are also working at comprehensibility since we believe that it should be quick to learn. In the same fashion usability, that should be as easy as possible. Even the flow of the interaction has been conceived so to reduce the amount of eventual frustration at the minimal terms, so that game-dynamic returns in any case a certain degree of satisfaction uncaringly of the users' response and the final output outcome.

There is not a single specific goal in the game and the aim, which is assembling the floor of tiles that appears on the tablet and can be projected on to the screen, is surely aesthetics of shape and sound. Nevertheless, since there are several options that can be manipulated (e.g. speed, shapes, colors, etc.) to achieve an evidently good result might be either tricky or conditionate by cleverness, and psychological factors as well as physical ones. Certainly, individual view and taste for aesthetics will also direct the interaction with the game.

2. Technical Implementation

The basic game set consists in a group of physical electronic cubes, whose number can be customizable and depends on the number of variables introduced in the game itself. Roughly, each cube (Fig. 1) is provided with an inner electronical circuit with a microprocessor that controls an accelerometer and Bluetooth system, which is able to send information about cube rotation to the Android or iOS App that controls the game either on tablet or smartphones.

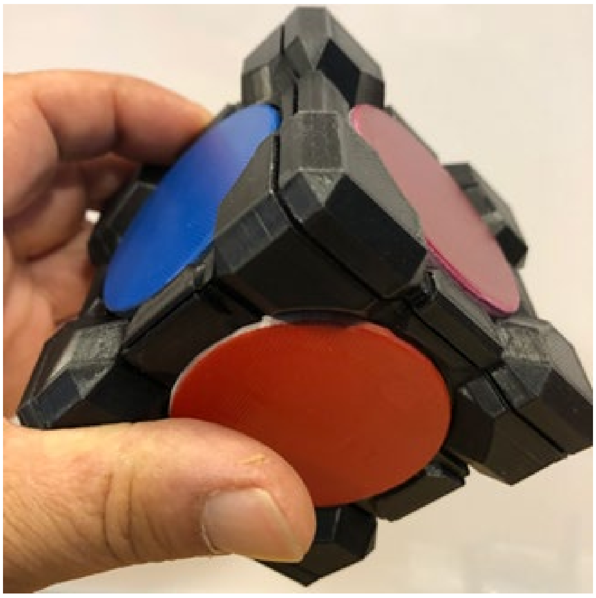


Figure 1. One cube example.

Each cube represents a specific variable and on each side of a specific cube, alias variable, the virtual representation of a possible value of the variable itself is coded. Therefore, tilting an active cube results in changing the denotation of a value in the system and as a consequence it will modify the state of tile in the virtual system.

The different cubes embody the variables introduced in the virtual system – at the moment they are Color, Shapes, Orientation, Sound, Speed, Direction, etc. – and, of course, they can be in a potentially quite large number whose only limit derives from the potentialities of the Bluetooth communication capabilities.

When starting a game session, users have to go through a quick initialization process that allows the app pairing the different cubes with the itself and recognize how many active cubes there are and which variables might be manipulated. The Tiler's App also forsees the not obvious capability of dynamically reallocating the number of cubes to use at each and every moment, so allowing the user to change the dynamic of the game instantly and constantly.

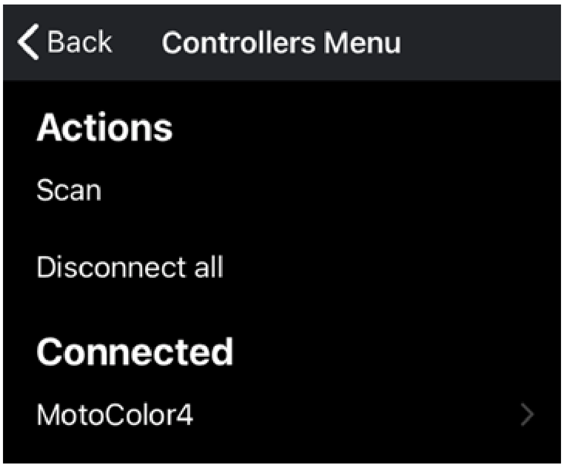


Figure 2. Pairing cubes in Tiler App.

The other side of the game consists of an application for mobile devices, which embody explicit techniques that allow to mutate a 3D based logic, consequent of the cube manipulation, to a 2D one, subsequent of the virtual tiles' management. Under many aspects, this was the most innovative and challenging task we did encounter and solve. Indeed, as even envisioned in many sci-fiction movies, one of the most promising evolutions of computer science is about interfacing 3D object manipulation in the physical world in order to obtain a 2D virtual representation of humans' actions.

To bend the binary logic imposed by screens and other informatic tools within a three-dimensional world we had to proceed carefully and, step-by-step, defining a correspondent meaning, variable by variable, while keeping and an eye on the whole, so that the integrity of the process could be coherent.

Therefore, each single cube needs (and has) its own symbolic representation that recalls the variable is calling and within it has a set of variation that allows to represent six (i.e. the sides of a cube) states that interfere with the output in a fulfilling or satisfying way. In the above figure (Fig. 1) it is shown the cube that allow to change the output colors, which is an easy task, which is made transparent to the user by the color coding with color patches on the sides of the cube. In other cases, such as orientation, speed, direction and mainly shape it might result more difficult.

Figure 3. Tiler App appearance.

The basic configuration of our output set was that of a chessboard in which the user can manipulate each single square while the flow of the focus on the single squares is by default automatized or can be manipulated by the user itself by tilting a specific cube. The square under manipulation is constantly highlighted (see Figure 3). In the above Figure 3 the Tiler App is using the famous Truchet tiles [1] schemata (see Figure 4) with the addition of two states (i.e. full foreground and full background) so to adapt it to a cubic logic. Hence, the color, shape or rotation underneath the cursor will change based on a rotation of the related physical cube.

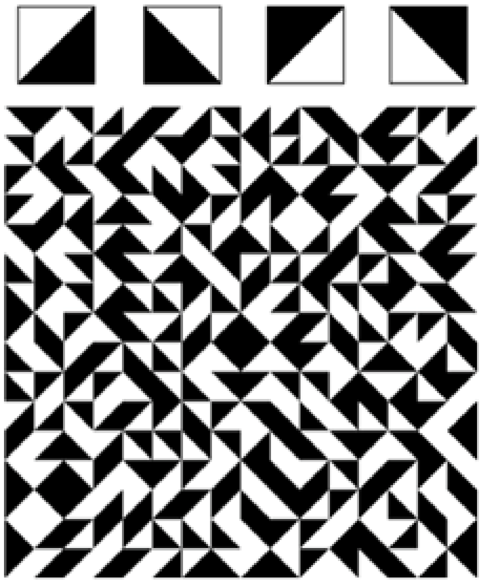


Figure 4. Truchet tiles schemata.

3. Discussion and Conclusion

The prototypical implementation of Tiler presented here has focused on the visual art, and how to allow the user to become creative with the visual art. It is currently implemented to run on an Android or iOS smart device, e.g. a tablet. This facilitates that the system can easily be set up anywhere as a component of a multisensory room. For the Tiler visual art part, this can easily be achieved by projecting the tablet visual representation onto a wall in the room, see Figure 5.

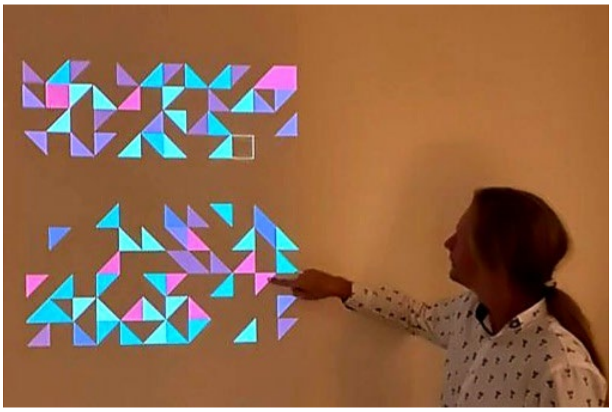


Figure 5. Tiler projected onto the wall as part of a multisensory room.

This may allow us to easily use Tiler as one component in multisensory rooms based on modular playware, in which modules with different sensorial modalities can easily be added and removed to create flexible multisensory environments (see example Fig. 6).

Figure 6. A playful creative multisensory room in the HC Andersen children's hospital in Denmark, in which the patient and the parent are using a modular playware cube to play with the visual patterns of the wall-mounted Moto Tiles [6].

Indeed, in our concept, we aim to develop *playful creative multisensory rooms*, in which we allow the user to become creative in the interaction with the elements in the multisensory room. The foundation of the concept is modular playware [7,8] and the Playware ABC [9]. The flexibility of the modular playware allows it to be a component for the Playware ABC and for our development of playful creative multisensory environments. The modular approach facilitates the work towards systems that are easy to set-up as environments by anybody, anywhere, anytime – in contrast to many classical multisensory rooms with a complex installation process, which demands a large infrastructure. The work on Tiler is seen as a component for this.

Some of the lessons learned from more than two decades of research and development of such flexible multisensory rooms relate to the design, feedback, sensory stimuli, variation and playfulness, and this gave direction to the design and development of Tiler.

Acknowledgements

Authors wish to thank Massimiliano Leggieri and Marin Holi from the Technical University of Denmark for their precious work and collaboration.

References

1. Pagliarini, L. "Codice Pagliarini.". *Rivista di Psicologia dell'Arte. Anno XXXIX n. 29 Dicembre 2018 Ed. Sergio e Giuliano Lombardo*.
2. Lund, H.H., 2015. Play for the elderly-effect studies of playful technology. In International Conference on Human Aspects of IT for the Aged Population (pp. 500-511). Springer, Cham.
3. Lund, H.H. and Ottesen, M., 2008. RoboMusic: a behavior-based approach. Artificial Life and Robotics, 12(1), pp.18-23.
4. Pagliarini, L. and Lund, H.H., 2014. ALife for Real and Virtual Audio-Video Performances. J. Robotics Netw. Artif. Life, 1(1), pp.33-38.
5. Pagliarini, L., Lund, H. H., Miglino, O., & Parisi, D. (1996). Artificial Life: A New Way to Build Educational and Therapeutic Games. In Proceedings of Artificial Life V. Cambridge, MA: MIT Press/Bradford Books.
6. Lund, H.H., Henningsen, A. and Nielsen, R., 2009. Modular robotic system as multisensory room in children's hospital. In Proceedings of 14th International Symposium on Artificial Life and Robotics (ISAROB).
7. Lund, H.H. and Marti, P., 2009, September. Designing modular robotic playware. In RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication (pp. 115-121). IEEE.
8. Lund, H.H. and Marti, P., 2005. Designing manipulative technologies for children with different abilities. Artificial Life and Robotics, 9(4), pp.175-187.
9. Lund, H.H., 2017. Playware ABC: Engineering Play for Everybody. J. Robotics Netw. Artif. Life, 3(4), pp.283-286.

Authors Introduction

Luigi Pagliarini

Luigi Pagliarini is an artist, psychologist, software designer, expert in robotics, AI and Artificial Life. He is currently Professor at the Academy of Fine Arts of Macerata (Italy) and Consultant Professor at DTU Center for Playware (Denmark). He has published in different international books, journals, congresses and conferences proceedings and has been rewarded with international prizes more than once. He has exhibited his work in different museums and institutions all over the world. Luigi Pagliarini has also worked for many different institutes and universities as professor or researcher and, as consultant, with many enterprises and multinational factories. His work has often been reported on many international newspapers, magazines and televisions.

Henrik Hautop Lund

Professor, Ph.D. Henrik Hautop Lund, Center for Playware, Technical University of Denmark, is World Champion in RoboCup Humanoids Freestyle 2002, and has more than 200 scientific publications and 5000 citations. He was awarded the Most Outstanding Healthcare Innovator in the World in 2019. Over the last year, he received international awards in Tokyo, Singapore, and London. He has developed shape-shifting modular robots, presented to the emperor of Japan, and has collaborated closely on robotics and AI with companies like LEGO, Kompan, BandaiNamco, Microsoft, Mizuno. He is the inventor of the Moto Tiles (www.moto-tiles.com), which are used by seniors all around the world.

Creative Multisensory Environments

Henrik Hautop Lund[1], **Luigi Pagliarini**[1,2]

[1] *Technical University of Denmark, 2800 Kgs. Lyngby, Denmark*
[2] *Academy of Fine Arts of Macerata, P.za Vittorio Veneto 5 - 405150, Macerata, Italy*
E-mail: hhl@playware.dtu.dk

Abstract

We outline the concept of playful creative multisensory environments. Multisensory environments are characterized by facilitating multimodal interaction by users through a composition of different objects. These objects allow manipulation of the environment. Multisensory rooms are often used to enhance users' sensations and emotions. In order to explore the development of such multisensory environments, which are playful and provide possibilities to become creative, we suggest the use of modular playware as part of the Playware ABC concept. This allows for the development of flexible, portable environments to be used by anybody, anywhere, anytime, and facilitates that the users can construct, combine and create. In this way, the users can become creative in seamless interaction with aesthetically pleasing environments. The paper outlines the lessons learned from the development and testing of a number of such playful creative multisensory rooms. Main findings are that such multisensory environments need to provide explicit immediate feedback, be simple in design, build on high quality aesthetics, provide variation, and be playful to result in intrinsic motivation.

Keywords: Multisensory Room, Snoezelen, Immersive, Playware, Modularity.

1. Introduction

Multisensory environments (or Snoezelen) originated in Holland in the late 1980s at the Hartenberg Institute. The word Snoezelen is derived from the Dutch words for sniff and doze. The aim of such multisensory environments is to induce leisure, enjoyment and relaxation in people through enhancing their sensations and emotions [1].

As a therapy, the environments are used as multisensory therapy, aiming to provide sensory stimulation for people who due to their learning disability would not spontaneously seek such stimulation. The environment prompts sensorial and emotional exploration. This is achieved in a constructed environment, usually a purpose designed room equipped with dimmed lighting various colours, visual displays, fiber-optic lighting, projectors with wheels for wall/floor displays, tactile objects, bubble tubes, olfactory stimulants, equipment for sound production and furnishing to relax on, such as floor cushions and water or air beds [1].

Even when multisensory rooms are used widely in therapeutic settings, there is debate as to whether multisensory environments should be considered a therapy as a primary purpose. Indeed, the original concept opposed the therapeutic framing of the concept and emphasized the aesthetic and playful qualities:

"The emphasis is on the provision of a leisure environment, rather than a therapeutic one, which includes equipment specifically designed to provide enjoyment." [2] When these environments are used with people with cognitive and/or physical disabilities, the original Snoezelen concept outlines the playful enjoyment as the primary aim of the activities (rather than a therapeutic aim):

"When considering leisure activities for people with profound handicaps there is often an emphasis on the learning or therapeutic dimension of recreation; for example, the acquisition of a new skill or the remediation of a behavior problem. The Snoezelen concept, however, stresses the fact that people with such handicaps have the same right to leisure time as everyone else and, whilst a therapeutic outcome or learning of a skill may arise from Snoezelen activities, the major emphasis is placed firmly on pleasure." [2]

This is in line with work on play and playful environments. Play is defined as a free and voluntary activity that we do for no other purpose than play itself. We do not play to achieve a certain outcome or product, but we play for the pleasure and enjoyment that we feel while playing. Nonetheless, under various circumstances, we may observe certain effects of play. For the one who plays, these effects are not the primary reason to engage in play. Therefore, we term such effects the *collateral effects of play*. The collateral effects of play can be educational achievements, motor skill enhancement, cognitive and physical rehabilitation, etc. These collateral effects of play can be significant and important, but it is essential to understand that play is a self-sustaining phenomenon which carries its purpose in itself. [3]

It is important in the design process to distinguish between (i) the design for therapeutic or educational outcome as the primary purpose, or (ii) the design for play and enjoyment as the primary purpose and therapeutic and/or educational outcome (only) as potential collateral effects. When designing for play, one aims to create user interaction arising from intrinsic motivation based on the enjoyment and free will of the user. The design will be evaluated by its ability to foster this enjoyment and play dynamics of the user. The therapeutic and/or educational outcome takes a secondary role and can/will only be evaluated after the design has been evaluated positively to foster the enjoyment and play dynamics.

These collateral effects of the play and enjoyment can be quite remarkable, and it has been asserted that the use of multisensory environments have positive therapeutic effects on disabled users [4-6]. Kewin [6] suggests that the therapeutic value is recognized in that it is believed to enhance exploration and development.

2. Playful Creative Multisensory Room

In our concept, we aim to develop *playful creative multisensory rooms*, in which we allow the user to become creative in the interaction with the elements in the multisensory room. The foundation of the concept is modular playware and the Playware ABC.

Playware is defined as intelligent hardware and software which aim at the creation of play and playful experiences [7], and hence the playware should work as a play force which pushes the user into a play dynamic. An instance of such playware is modular playware, in which the intelligent hardware and software takes a modular form. Modular playware is described by Lund and Marti [8] as playful modular objects that are visible, manipulable, sharable and interactive and imply construction, active participation, creativity for assembling, mastery of the parts by the users who play with them. These technologies allow a range of play from simple exercise play up to construction play requiring sensory-motor skills as well as coordination/ manipulation of objects.

The *Playware ABC* concept addresses the challenge of facilitating interaction by anybody providing quality of play and its collateral effects for many different user groups. The Playware ABC concept is formed by:

A: Anybody, Anywhere, Anytime

B: Building Bodies and Brains

C: Construct, Combine, Create

The Playware ABC concepts works for creating technology solutions for *anybody, anywhere, anytime* by using embodied artificial intelligence *building bodies and brains*, which facilitates that users can themselves manipulate with the technology solutions to *construct, combine and create* their own solutions. [9]

Modular playware technology can be well appropriated for the Playware ABC, if the modules are designed and developed according to this concept to allow anybody to easily construct and create with the modules. Indeed, Lund and Marti [8] outline a number of design features for modular playware to become flexible in both set-up and activity building for the end-user. Key features of this design approach are modularity, flexibility, and construction, immediate feedback to stimulate engagement, activity design by end-users, and creative exploration of play activities.

The flexibility of the modular playware allows it to be a component for the Playware ABC and for our development of playful creative multisensory environments. The modular approach facilitates the work towards systems that are easy to set-up as environments by anybody, anywhere, anytime – in contrast to many classical multisensory rooms with a complex installation process, which demands a large infrastructure. The Playware ABC concept suggests that such infrastructure and installation demands may be avoided to a large degree when the starting point of the engineering and IT system design is transformed from the optimal system performance to become a focus on creating a solution that can be used by anybody, anywhere, anytime. [10]

3. Implementations

We have combined different technological platforms to explore the modular playware concept for different user sensory modalities by combining heterogeneous building blocks (i.e. modular playware), creating multisensory environments. For instance, we combined the modular interactive tiles and cubic I-Blocks in the creation of a multi-sensory room in the HC Andersen children's hospital [11], we combined rolling pins and light&sound cylinders in the creation of a multi-sensory room for elderly with dementia [12], and we combined modular tiles, rolling pins, and light&sound cylinders for the first RoboMusic concert [13] (See Fig. 1-3).

These are *playful creative multisensory environments*, which in an easy manner allow any user to become creative with sound, light, visual images, and materials. The manipulation in terms of physical interaction with modular playware components such as the interactive tiles, blocks, pins and cylinders results in creative manipulation of the output modalities of the multisensory environments in terms of sound, light, vibration and images.

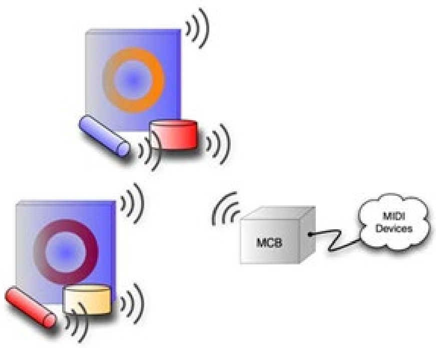


Fig. 1. Tiles, rolling pins, and cylinders used for Playful Creative Multisensory Environments for RoboMusic concerts, and rooms for seniors with dementia.

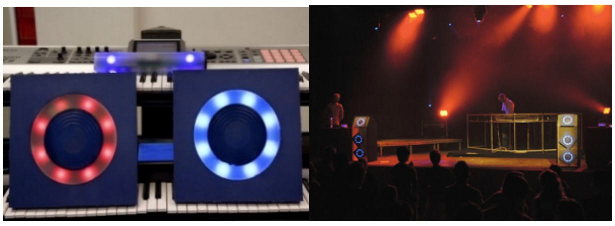


Figure 2. Left: Two Tiles and a RollingPin used as robotic instruments. Right: The RoboMusic live concert set-up, with Funkstar De Luxe and his control station in the center, and the robotic instruments on the left and right side of the stage.

Figure 1 and 2 shows the use of a few such modules for making a Playful Creative Multisensory Environment. The modules were easily set up as a multisensory room in a concert hall for allowing the audience to interact with the live performing artist (see Fig. 2 right). The modular playware technology was further refined a decade later for the interactive live performances on stage by Peter Gabriel. Similar modular playware was adopted to playful interaction as a multisensory room in the HC Andersen children's hospital (see Fig. 3).

Figure 3. An early implementation of a Playful Creative

Multisensory Room in the HC Andersen childrens hospital with a child turning and shaking a cube to create patterns and games on the modular tiles wall.

Some of the lessons learned from more than two decades of research and development of such flexible multisensory rooms relate to the design, feedback, sensory stimuli, variation and playfulness.

- **explicit immediate feedback**: it is important to make the feedback explicit and immediate for the user to appreciate the consequences of the physical actions.
- **simplicity**: the design must be simple to increase robustness and allow for easy set-up anywhere. Also, interactivity should be kept simple for anybody to understand the interaction modalities and consequences within seconds.
- **quality**: the sensory stimuli must have a high aesthetical quality, both for the visual and audio, and games must have a high game quality for users to appreciate manipulation with the objects to construct, combine and create (hence we engage professional artists, musicians, football players, etc. in the design)
- **variation**: it is important the ensure variation for both short-term and long-term interaction. We appreciate such changeability in nature, and can do so also when imposing variation and adaptivity in the multisensory room output.
- **playfulness**: when the system is made playful and the modular playware helps pushing the user into a play dynamic, we observe that the users easily forget fears and limitations, and continue to engage in the environment due to their intrinsic motivation.

4. Discussion and Conclusion

Building on these developments and insights about playful creative multisensory rooms, as artistic and aesthetical expressions, we are suggesting multisensory rooms in which any user can create and perform both graphics and music by manipulating simulation parameters and by manipulating a number of interactive playware modules (e.g. tiles or cubes). The multisensory room is a cross-media platform, which points towards cross-modal and cross-sensory artistic results. As indicated by Pagliarini and Lund [14], the playful creative multisensory room combines different input and output methodologies, systems and tools that might lead to a broader vision of software and robotic systems with an articulated, fluid and bidirectional flow between the physical and virtual environment. It is believed that such a fluid and multifaceted representation of a single activity may widely enhance the user immersion into a reality that combines multi-sensory activation (physical/virtual).

References

1. E. Slevin & A. McClelland. Multisensory Environments: Are They Therapeutic? A Single-Subject Evaluation of the Clinical Effectiveness of a Multisensory Environment. Journal of Clinical Nursing, 8: 48-56, 1999
2. Haggar, L.E. and Hutchinson, R.B., 1991. Snoezelen: An approach to the provision of a leisure resource for people with profound and multiple handicaps. Journal of the British Institute of Mental Handicap (APEX), 19(2), pp.51-55.
3. Lund, H.H., 2015, August. Play for the elderly-effect studies of playful technology. In International Conference on Human Aspects of IT for the Aged Population (pp. 500-511). Springer, Cham.
4. Hulsegge, J. and Verheul, A., 1987. Snoezelen: another world. Rompa, Chesterfield.
5. Hutchinson, R. and Haggar, L., 1991. The development and evaluation of a Snoezelen leisure resource for people with profound and multiple handicaps. The Whittington Hall Snoezelen Project: A Report from Inception to the End of the First Twelve Months. Chesterfield (England): North Derbyshire Health Authority, Community Health Care Service, Whittington Hall Hospital, pp.1-47.
6. Kewin, J., 1994. Snoezelen: the reason and the method. Sensations and Disability: Sensory Environments for Leisure, Snoezelen, Education and Therapy. Chesterfield: ROMPA UK Publications, pp.6-17.
7. Lund, H.H., Klitbo, T. and Jessen, C., 2005. Playware technology for physically activating play. Artificial life and Robotics, 9(4), pp.165-174.
8. Lund, H.H. and Marti, P., 2009, September. Designing modular robotic playware. In RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication (pp. 115-121). IEEE.
9. Lund, H.H., 2017. Playware ABC: Engineering Play for Everybody. J. Robotics Netw. Artif. Life, 3(4), pp.283-286.
10. Lund, H.H., 2012. Modular Playware Technology–A Brief Historical Review. In Proceedings of 17th International Symposium on Artificial Life and Robotics, ISAROB, Japan
11. Lund, H.H., Henningsen, A. and Nielsen, R., 2009. Modular robotic system as multisensory room in children's hospital. In Proceedings of 14th International Symposium on Artificial Life and Robotics (ISAROB).
12. Marti, P., Lund, H.H., Bacigalupo, M., Giusti, L. and Mennecozzi, C., 2007. A multi-sensory environment for the treatment of dementia affected subjects. International Journal of Gerontechnology, 6(1).
13. Lund, H.H. and Ottesen, M., 2008. RoboMusic: a behavior-based approach. Artificial Life and Robotics, 12(1), pp.18-23.
14. Pagliarini, L. and Lund, H.H., 2014. ALife for Real and Virtual Audio-Video Performances. J. Robotics Netw. Artif. Life, 1(1), pp.33-38.

15.

Authors Introduction

Henrik Hautop Lund

Professor, Ph.D. Henrik Hautop Lund, Center for Playware, Technical University of Denmark, is World Champion in RoboCup Humanoids Freestyle 2002, and has more than 200 scientific publications and 5000 citations. He was awarded the Most Outstanding Healthcare Innovator in the World in 2019. Over the last year, he received international awards in Tokyo, Singapore, and London. He has developed shape-shifting modular robots, presented to the emperor of Japan, and has collaborated closely on robotics and AI with companies like LEGO, Kompan, BandaiNamco, Microsoft, Mizuno. He is the inventor of the Moto Tiles (www.moto-tiles.com), which are used by seniors all around the world.

Luigi Pagliarini

Luigi Pagliarini is an artist, psychologist, software designer, expert in robotics, AI and Artificial Life. He is currently Professor at the Academy of Fine Arts of Macerata (Italy) and Consultant Professor at DTU Center for Playware (Denmark). He has published in different international books, journals, congresses and conferences proceedings and has been rewarded with international prizes more than once. He has exhibited his work in different museums and institutions all over the world. Luigi Pagliarini has also worked for many different institutes and universities as professor or researcher and, as consultant, with many enterprises and multinational factories. His work has often been reported on many international newspapers, magazines and televisions.

Using GrowBots to Study Heat and Nutrient Stress in Basil

Rishav Bose
Center for Playware, Technical University of Denmark, Building 326,
2800 Kongens Lyngby, Denmark

Henrik Hautop Lund
Center for Playware, Technical University of Denmark, Building 326,
2800 Kongens Lyngby, Denmark
E-mail: boserishav@gmail.com, hhl@playware.dtu.dk
https://www.elektro.dtu.dk/

Abstract

We present how a novel type of robot called a food computer can be used to simulate abiotic stresses and study their impact on hydroponically grown Italian basil, Ocimum basilicum. The food computer called the GrowBot is a tabletop sized robotic greenhouse for growing edible food plants. The GrowBot's actuators were used to alter the environmental conditions in the growth chamber to study different aspects of plant growth and food production in varying climate scenarios. The experiments show how we can used the LED lights to control the temperature to a certain, desired range (29ºC – 35ºC) for the heat stress experiments, while measurements show that we can simultaneously obtain the photosythetically active radition (PAR) values to be in the ideal range for growth of the basil plants.

Keywords: Artificial Life, Biology, Cyber Agriculture, Food Computing

1. Introduction

Abiotic stresses are the negative impact of non-living factors on plants in a specific environment. Abiotic stress factors, or stressors, are the naturally occurring, intangible, and inanimate factors such as intense sunlight, temperature, or wind. Extremities in temperature, inadequate or excessive water, etc. are hostile to plant growth and development, leading to crop yield penalties. Different abiotic stressors typically occur simultaneously in nature [1].

The growth and development of the plants are greatly affected by a series of morphological, biochemical, and physiological changes resulting from high-temperature stress. Heat stresses can be due to both chronic and abrupt heating. Heat stress reduces the number, mass, and growth of the roots. This curtails the supply of water and nutrients to the shoot system of the plant. High temperatures may also cause scorching of the twigs and leaves along with visual symptoms of sunburn, leaves senescence, growth inhibition, and discolouration of fruits and leaves [2].

A few studies have made attempts to replicate and understand the impact of heat and nutrient stresses in a controlled environment. For example Ukrit Watchareeruetai et al. performed a study in which nutrient deficiencies of calcium (Ca), iron (Fe), magnesium (Mg), nitrogen (N), and potassium (K) were induced on a set of black gram plants [3]. A study at Cornell University grew sweet basil 'Genovese' in nutrient solutions deficient in individual macro and micro-nutrients. The study recorded the visual symptoms of the different deficiencies [4-5]. Melis Mengutay et al. investigated the role of adequate magnesium (Mg) nutrition in mitigating the detrimental effects of heat and light stresses on wheat and maize [6].

2. GrowBot Hardware and Software

The GrowBot [7] (see figure 1) is a robotic system that can control the growth of natural life in the form of edible food plants. It allows its users to parameterize growth conditions for plants in terms of temperature, light cycle, the colour of grow light, fertilization, etc. Such a parameterization is called a recipe. Different recipes bring about different growth patterns in plants in the GrowBot. Hence the GrowBot can not only be utilized to test hypotheses in order to provide new knowledge about natural life, but it can contemporaneously control and optimize the growth of life.

Fig. 1. The GrowBot

Inside the growth chamber, there is room for a water tray with a lid with 16 holes arranged in a 4x4 grid (see figure 2). The seeds of the plants can be sown into small rockwool cubes that are laid down into cups. The cups are then placed into the aforementioned holes. On the front side of the chamber is a door, which can be tightened when closed to ensure an airtight space inside the chamber.

The control parameters are recorded as recipes, and figure 3 shows an example of a recipe. A recipe is composed of two types of rules - time rule and condition rule. In a time rule, we enter the time duration we want

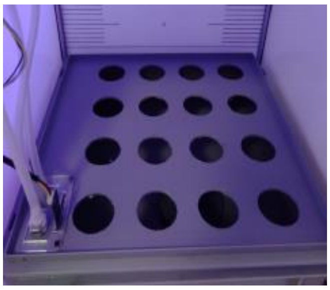


Fig. 2. Water tray in the growth chamber

an actuator to be on for. In figure 3, we are turning on the white light at 100% intensity for a period of 8 hours. In a condition rule, the value of a selected sensor is monitored,

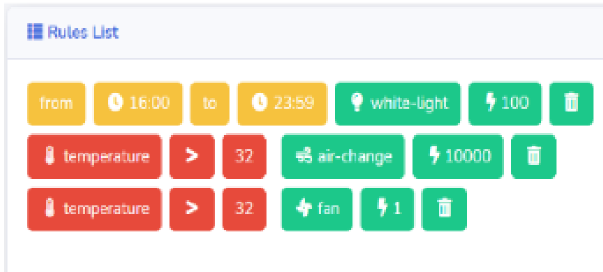


Fig. 3. Example of a recipe

and action is taken based on the sensor reading. In figure 3, we are monitoring the value of the temperature sensor, based on which the air change and fan actuators are turned on.

3. Experimental setup

In our experiments, we induced a combination of heat and nutrient stresses. Heat stresses were induced by using the LEDs. The full spectrum LEDs were used at 100% intensity, while a combination of the full spectrum and UV LEDs were used at 50% intensity each. We induced three types of nutrient stresses - Magnesium (Mg), Potassium (K), and Nitrogen (N). Table 1 shows the nutrients present in the potassium deficient solution and the quantities used in 10 litres of water. Three nutrient solutions, each deficient in the individual elements were created. The other parameters of the experiment were as follows:

- Amount of water in the growth tray: 10 litres
- pH range: 5.5 – 6
- Electric conductivity (EC) range: 1.0 - 1.6 mS/cm

- Temperature range: 29ºC – 35ºC
- Photoperiod - 24 hours
- Time duration of experiment: 7 days
- Number of plants in each GrowBot: 2

The growth chamber was maintained in the aforementioned temperature range because according to a study from Iowa State University, phenotypes like plant height, number of branches, fresh and dry mass accumulation rate, and internode length were negatively impacted at temperatures above 29ºC in different basil cultivars. The study noted that above 35ºC, the plants would end up dying [8]. The choice for the ranges pH and EC was based on our literature review [9]. We used a 24-hour photoperiod because it has been noted that constant daylight has a positive effect on basil growth [10]. Another factor taken into consideration while selecting the lighting intensity and photoperiod was that basil requires high photosynthetic photon flux density (PPFD) values of around 500 µmol.m^{-2}.s^{-1} [11].

Table 1: Solution to induce potassium (K) deficiency

Nutrients present	**Quantity**
N(17%)	1.5ml
N(2.2%)-Ca(2.8%)	1.5ml
P(17%)	1.5ml
Mg(7%)	0.5ml
Mg, S	0.5 grams
Ca(12%)	1.5ml
Ca, S	0.33 grams
Fe(0.1%)	1.5ml
B, Cu, Fe, Mn, Mo, Zn	4ml

For our study, we purchased pre-grown Italian basil plants from Irma's Infarm hydroponic system. This was done to save time on growing the plants from scratch, as we were interested in studying the impact of stresses on mature plants only.

We used BioBizz Bio pH Minus and BioBizz Bio pH Plus to maintain the pH of the nutrient solutions. To measure the pH and EC values of the nutrient solution, we used the HI98107 pH meter and HI98304 CE tester, respectively.

4. Results

In order to maintain the temperature within 29ºC - 35ºC, we performed a vast number of experiments varying the LED light intensities, as we noticed that the growth chamber temperature was affected by the intensity of the LEDs. In figure 4, we can see an example of how the growth chamber temperature was modified over a two hour period by using the LEDs at different intensities. In the graph at the top, we can see that the temperature range is not being maintained. However, when the intensities were changed, it was possible to maintain the temperature in the required range, as we can see in the bottom graph in figure 4.

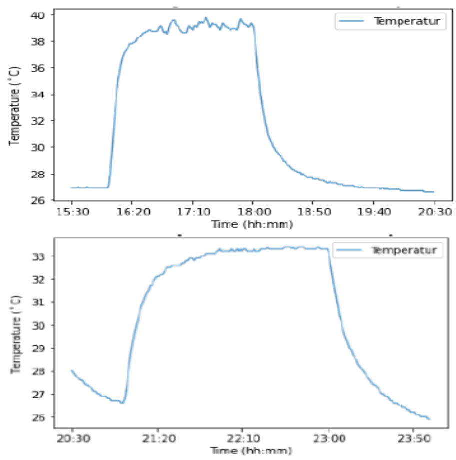


Fig. 4. Heating effect of white and blue light combination at 100% intensity each (top) and at 50% intensity each (bottom)

A number of such experiments helped us come to the conclusion that the following lighting settings should be used for keeping the temperature in the desired range:

- Mono-lighting
  - White (W) - 100% intensity
  - Blue (B) - 100% intensity
  - Red (R) - 100% intensity
- Multi-lighting
  - White-Blue-Red (WBR) - 33% intensity (each)
  - White-Blue (WB) - 50% intensity (each)
  - White-Red (WR) - 50% intensity (each)
  - Red-Blue (RB) - 50% intensity (each)

Our next step was to use an Apogee MQ-500 full-spectrum quantum photosynthetically active radiation (PAR) meter at different locations and heights in the growth chamber. This was done to ensure the PPFD values at different points in the GrowBots. We saw in the last section that basil grows best at an irradiance of around 500 µmol.m^{-2}.s^{-1}. Figure 5 shows the example of one such measurement in a GrowBot at 17.2cm from the bottom of the growth chamber. We measure the PPFD values at five locations - lower left (LL), top left (TL),

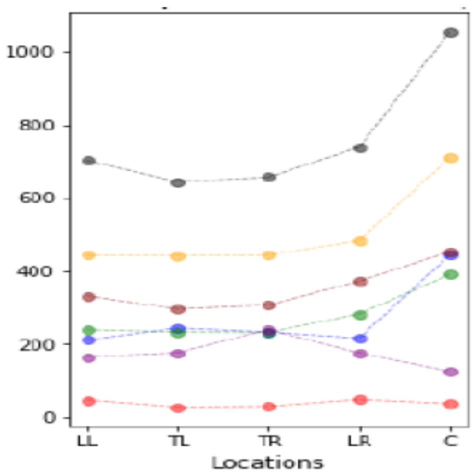


Fig. 5. Variation of PPFD values in one of the GrowBots

top right (TR), lower right(LR), and centre (C). From the graph, we can see that the PPF requirements are approximately satisfied by white light at 100% and a combination of white and blue light at 50% each. The basil plants that we used had an average height of approximately 15cm high. So, 17.2cm was an appropriate height to take these measurements. In this manner, we ensured that the temperature was maintained in the interval 29ºC-35ºC while satisfying the PPFD requirement for basil.

In figure 6, we can see the image of the growth chamber from the potassium and heat stress experiment taken using the camera of one of the GrowBot. The image on the top is from the first day of the experiment. The image at the bottom is from the last day of the experiment. We can see interveinal chlorosis in a large number of the leaves and some necrotic spots in the older leaves.

One of the things noted in the experiments was that even though the plants were in experiencing heat and nutrient stress, the weights of the plants increased. This gave us an important insight into the `dilution effect' - an inverse relationship between yield and nutrient concentration in food [10]. This was observed across all the stress experiments. Table 2 shows the weights of the plants before and after the heat and potassium stress test.

Fig. 6. Plant canopy images at the start (top) and end (bottom) of the heat and potassium stress experiment

Table 2: Weight of the plants before and after the heat and potassium stress test

Before stress test		After stress test
Plant	Weight (g)	Weight (g)
P1 (Right)	51	60
P2 (Left)	59	82

5. Discussion and conclusions

The GrowBots allows their users to parameterize life growth conditions for plants through the recipes in terms of temperature, light cycles and intensities, colour of grow light, irrigation and fertilizing, etc. Different recipes will lead to different growth patterns. These experiments show how the GrowBot can be utilized to test hypotheses in order to understand different aspects about natural life. These types of stress experiments can be used on newly created genetically modified crops to test their resilience in the seemingly unpredictable climate scenario that we are facing.

Acknowledgements

We would like to thank the project partners - DTU Food, DTU Skylab, and Københavns Professionshøjskole. We also express our gratitude towards the funding bodies - Nordea Foundation and Novo Nordisk Foundation.

References

1. Mei He, Cheng-Qiang He, and Nai-Zheng Ding. "Abiotic stresses: general defenses of land plants and chances for engineering multistress tolerance". In: Frontiers in plant science 9 (2018), p. 1771.
2. Shah Fahad et al. "Crop production under drought and heat stress: plant responses and management options". In: Frontiers in plant science 8 (2017), p. 1147.
3. Ukrit Watchareeruetai et al. "Identification of plant nutrient deficiencies using convolutional neural networks". In: 2018 International Electrical Engineering Congress (iEECON). IEEE. 2018, pp. 1–4.
4. Neil Mattson. "Magnesium Deficiency of Hydroponic and Container Grown Basil". In: e-Gro Edible Alert 3 (2018).
5. Neil Mattson and Tanya Merrill. "Symptoms of Common Nutrient Deficiencies in Hydroponic Basil". In: e-Gro Research Update, Cornell University (2016).
6. Melis Mengutay et al. "Adequate magnesium nutrition mitigates adverse effects of heat stress on maize and wheat". In: Plant and soil 368.1 (2013), pp. 57–72.
7. Lund, H. H., Exner, M., Jensen, N. E., Leggieri, M., Outzen, M., Ravn-Haren, G., Sehested, M. v., Væring, A., Andersen R. (2022) Robotics for Growing Life. ibid
8. Kellie Jean Walters. "Quantifying the effects of hydroponic systems, nutrient solution, and air temperature on growth and development of basil (Ocimum L.) species". In: (2015).
9. Hardeep Singh and Dunn Bruce. "Electrial conductivity and pH guide for hydroponics". In: Oklahoma Cooperative Extension Fact Sheets, HLA-6722. Oklahoma State University, Division of Agricultural Sciences and Natural Resources (2016), p. 5.
10. Arielle J Johnson et al. "Flavor-cyber-agriculture: Optimization of plant metabolites in an open-source control environment through surrogate modeling". In: PLoS One 14.4 (2019), e0213918.
11. Angela R Beaman, Richard J Gladon, and James A Schrader. "Sweet basil requires an irradiance of 500 µ mol· m- 2· s- 1 for greatest edible biomass production". In: HortScience 44.1 (2009), pp. 64–67.

Authors Introduction

Mr. Rishav Bose

Rishav Bose is a Research Assistant at the Center for Playware, Technical University of Denmark (DTU). He completed his masters degree in Electrical Engineering from DTU in October 2021, where he wrote his master's thesis on Optimizing Plant Growth Recipes in Growbots with Machine Learning. His main research interests are food computing, cyber agriculture, and computer vision.

Prof. Henrik Hautop Lund

Professor, Ph.D. Henrik Hautop Lund, Center for Playware, Technical University of Denmark, is World Champion in RoboCup Humanoids Freestyle 2002, and has more than 200 scientific publications and 5000 citations. He was awarded the Most Outstanding Healthcare Innovator in the World in 2019. Over the last year, he received international awards in Tokyo, Singapore, and London. He has developed shape-shifting modular robots, presented to the emperor of Japan, and has collaborated closely on robotics and AI with companies like LEGO, Kompan, BandaiNamco, Microsoft, Mizuno. He is the inventor of the Moto Tiles (www.moto-tiles.com), which are used by seniors all around the world.

Convolutional Neural Network for Studying Plant Nutrient Deficiencies

Rishav Bose*
Center for Playware, Technical University of Denmark, Building 326,
2800 Kongens Lyngby, Denmark

Henrik Hautop Lund
Center for Playware, Technical University of Denmark, Building 326,
2800 Kongens Lyngby, Denmark
E-mail: boserishav@gmail.com, hhl@playware.dtu.dk
https://www.elektro.dtu.dk/

Abstract

We discuss the development of a vision-based plant phenotyping system based on a novel type of robotic system called a food computer. The food computer used in this project is called the GrowBot. It has a host of sensors to help analyse the growth chamber including a Raspberry Pi camera. The project revolved around developing a system to segment the plant canopy from its background and analyse nutrient deficiencies from the images taken by the camera. The pilot project investigated how a segmentation model called U-Net could be used to study the images. One of the drawbacks of many existing vision-based plant phenotyping systems is that their convolutional neural networks (CNNs) were trained to analyse very ideal images of individual leaves. This pilot project tried to address that issue, while at the same time explored how to train the neural networks to learn segmentation from a small image dataset.

Keywords: Convolutional Neural Networks, Cyber Agriculture, Food Computing, Image Segmentation, Plant Health Monitoring, U-Net

1. Introduction

During the process of plant growth, plants need to be provided with various nutrients. Proper nutrition is a vital factor that strongly determines many aspects of a plant's life cycle, such as growth rate, flowering, fruit development, and fertilization. A lack of one or more nutrients causes diseases in plants that affect crop yield. Thus, identifying a nutrient deficiency or disease correctly when it first appears is a crucial step for efficient disease management. Nutrient deficiencies manifest themselves as unusual appearances on a plant - especially on its leaves (see figure 1).

These symptoms can be identified by visual examination. However, such an analysis relies heavily on the presence of one or more individuals with domain expertise. It is also impractical in large-scale studies. Hence, a number of studies have made an effort to identify nutrient deficiencies and diseases by analysing images of the leaves using convolutional neural networks (CNNs).

A number of studies focused on performing disease classification using the PlantVillage dataset. In the PlantVillage dataset, the images were taken under ideal conditions with each image containing a single leaf, facing upwards, on a homogeneous background. The studies used both transfer learning and train-from-scratch approaches to learn how to predict the crop-disease pairs


Fig. 1. K defiency in Italian basil (left), Mg defiency in Italian basil (right)

in the images from the PlantVillage dataset. It was reported that the ideal conditions that the diseased leaves were photographed were a drawback for real-world applications. The requirement of a more diverse set of training examples was stressed upon [1][3].

Apart from disease classification, deep learning models have also been applied for the automatic diagnosis of plant disease severity. A fine-tuned VGG16 network was used to classify images of healthy apple leaves and apple leaf black rot images from the PlantVillage dataset that were further annotated with severity labels - early-stage, middle-stage, or end-stage [2].

Marni Tausen et al. [4] and Ke Lin et al. [5] used a U-net-based plant segmentation model to extricate clover plants (Trifolium repens) from their background and to segment powdery mildew on cucumber leaf images, respectively.

2. GrowBot Hardware and Software

The GrowBot [6] (see figure 2) is a robotic system that can control the growth of natural life in the form of edible food plants. It allows its users to parameterize growth conditions for plants in terms of temperature, light cycle, the colour of grow light, fertilization, etc. Such a parameterization is called a recipe. Different recipes bring about different growth patterns in plants in the GrowBot.

Inside the growth chamber, there is room for a water tray with a lid with 16 holes arranged in a 4x4 grid (see figure 3). The seeds of the plants can be sown into small rockwool cubes that are laid down into cups. The cups are then placed into the aforementioned holes. On the front side of the chamber is a door, which can be tightened when closed to ensure an airtight space inside the chamber.

Fig. 2. The GrowBot

The GrowBot is equipped with the OV5647 CMOS image sensor, which is connected to the Raspberry Pi 4 - the main processor of the GrowBot.

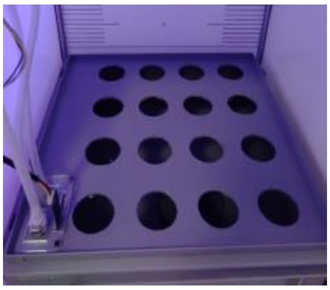


Fig. 3. Water tray in the growth chamber

3. Experiments

The first step of the experiments was to induce a combination of heat and nutrient stresses. For inducing heat stress, the growth chamber was maintained between 29ºC-35ºC. For inducing nutrient stress, nutrient solutions, each deficient in the individual elements, were created. The following stress experiments were performed:

- Heat and Magnesium (Mg) stress - Mg deficient solution was used
- Heat and Potassium (K) stress - K deficient solution was used
- Control I experiment - All nutrients were provided, and the temperature was maintained between 29ºC - 35ºC

- Control II experiment - All nutrients were provided, and the temperature was maintained between 24ºC - 29ºC

Images of the growth chamber were taken during all of these experiments. The training and test set split for the image data is shown in table 1. The train set had 89 images, while the test set had 42 images. The lighting combinations and intensities used while taking the images are shown in table 2. For each stress experiment, two plants were used. To record our dataset, we had placed the plants in unique locations for each stress test. In addition to the varying light intensities, changing the locations of the plants helped generate a lot of variety in the dataset. All the images used in the training, validation, and test sets were normalised, and rigid augmentations - horizontal flip, vertical flip, and random rotations were induced in the training images.

Table 1: Train and test set splits for the project

	Training set	Test set
Name of experiment	Number of images	Number of images
Heat and general nutrient stress	14	9
Heat and Mg stress	32	9
Heat and K stress	21	9
Control I	11	8
Control II	11	6

Table 2: Lighting setups used in the project

Type of coloured light	Intensities
Full spectrum	100%
Full spectrum	100%
Full spectrum	60%
Full spectrum and UV	50% (each)
Full spectrum and UV	30%, 45%

Our goal was to segment the plant canopy from its background and then study the variation of the hue values of the segmented image, as variation in hue has been linked to different nutrient deficiencies [4]. For performing the segmentation, we used a U-Net architecture which has an encoder-decoder structure (see figure 5). The encoder part of the U-net has a typical CNN-like architecture. This is the part of the U-Net that captures contextual information, while the decoder helps in localising the features [7].

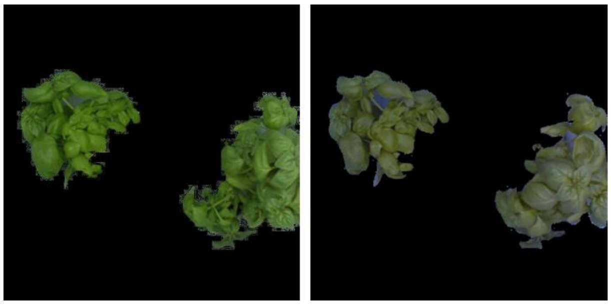


Fig. 4. Noisy images (left), denoised image (right)

The first step before training was to annotate the images, which was done using a tool called ImageJ [8]. One of the challenges with using ImageJ was that when the annotated masks were applied to the image, the resultant image in many cases was noisy around the edges of the plant canopy (see figure 4 (left)). So, the morphological opening operation was used to remove the noise along the edges of the plant canopy in the mask before applying it to the image. In figure 4 (right), we can see the results of applying morphological opening on a mask.

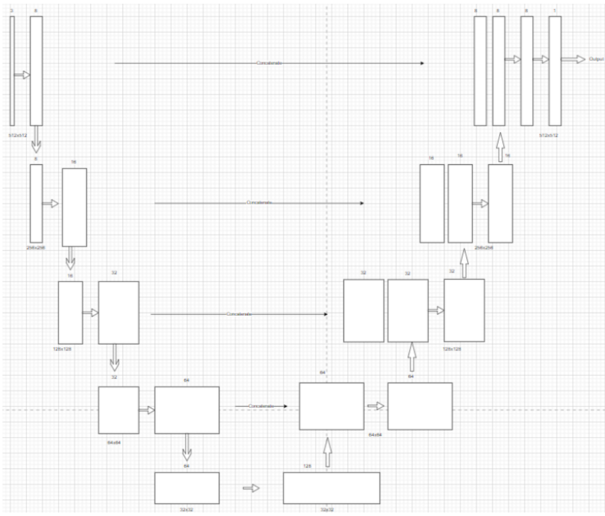


Fig. 5. Structure of the U-Net used in this project

4. Results and conclusion

We tested three models of the U-Net, see Table 3, and the best results on the training and validation sets were obtained with model 2. The training-validation split was

75% and 25%, respectively. Stochastic gradient descent (SGD) was used as the optimiser with a batch size of 1.

Table 3: Details of the three models used

	Model details
Model	Architecture
M1	8x-16x-32x-64x-128x
M2	16x-32x-64x-128x-256x
M3	64x-128x-256x-512x-1024x

The other parameters and the training and validation losses are shown in table 4. Dropout with a probability of 50% was used as the regularization technique to counter over-fitting. An epoch corresponds to the number of passes completed over the entire training set. Dice loss was used to compute the training and validation losses. Dice loss measures the relative overlap between the prediction and the ground truth. The best results on the test set were obtained on the heat and general nutrient test set. Table 5 shows the metrics of the results on the test set. The Jaccard index is a similarity coefficient to gauge the similarity and diversity of sample sets. F1score, recall, and precision take into consideration false positive and false negatives predicted by the algorithm. Lastly, pixel accuracy is the percentage of pixels the algorithm classified correctly.

Table 4: Parameters of the best performing model

Learning rate	Epochs	Dropout	Train loss	Validation loss
1e-4	20	50%	0.6298	0.6179

Table 5: Metrics on the best performing test set

Jaccard Index	F1 score	Recall	Precision	Pixel accuracy
0.4235	0.5914	0.5922	0.5983	0.7997

Figure 6 (left) shows the prediction on an image from the heat and general nutrient test set, and the ground truth is shown on the right. We can see that in the prediction, the algorithm predicts a part of the growth tray as the plant. This is because the ImageJ annotation tool used splines to enclose and annotate the region of interest. In many images, there were parts of the plant canopy that had gaps through which the water tray or the growth chamber walls were visible. Unfortunately, the tool did not have the resources to keep these parts out of the annotation (see figure 7). This hints towards the fact that the

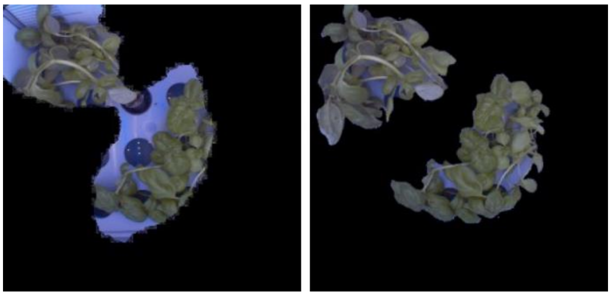


Fig. 6. Prediction (left), Ground truth (right)

Fig. 7. Drawback of ImageJ

annotation tool was not appropriate for our problem. Unfortunately, due to the fact that the segmentation was recognizing parts of the growth chamber walls and the growth tray as a part of the plant, the analysis of the hue values of the segmented regions were erroneous. However, this problem can be tackled by focusing on the development of a more appropriate annotation tool. Once a proper segmentation is achieved, the hue values of the plant canopies need to be computed, after which we can plot how the hue varies chronologically. This is because variation of hue has been linked to nutrition deficiencies [4].

A remarkable thing to be noted is the time efficiency of the U-Net. It took only 1.0812s to compute the predictions on 7 images on the CPU. This indicates a path towards plant phenotyping on smartphones or low-powered devices. In the future, transfer learning or knowledge distillation can be used to improve accuracy. Additionally, GANs (generative adversarial networks) can be used to generate more data points instead of relying on image augmentation alone.

Acknowledgements

We would like to thank the project partners at DTU Food, DTU Skylab, and Københavns Professionshøjskole. We also express our gratitude towards the funding bodies - Nordea Foundation and Novo Nordisk Foundation.

References

1. Sharada P Mohanty, David P Hughes, and Marcel Salathé. “Using deep learning for image-based plant disease detection”. In: Frontiers in plant science 7 (2016), p. 1419.
2. Guan Wang, Yu Sun, and Jianxin Wang. “Automatic image-based plant disease severity estimation using deep learning”. In: Computational intelligence and neuroscience 2017 (2017).
3. Muhammad Hammad Saleem, Johan Potgieter, and Khalid Mahmood Arif. “Plant Disease Classification: A Comparative Evaluation of Convolutional Neural Networks and Deep Learning Optimizers”. In: Plants 9.10 (2020), p. 1319.
4. Marni Tausen et al. “Greenotyper: Image-based plant phenotyping using distributed computing and deep learning”. In: Frontiers in plant science 11 (2020), p. 1181.
5. Ke Lin et al. “Deep learning-based segmentation and quantification of cucumber powdery mildew using convolutional neural network”. In: Frontiers in plant science 10 (2019), p. 155.
6. Lund, H. H., Exner, M., Jensen, N. E., Leggieri, M., Outzen, M., Ravn-Haren, G., Sehested, M. v., Væring, A., Andersen R. (2022) Robotics for Growing Life. ibid
7. Olaf Ronneberger, Philipp Fischer, and Thomas Brox. “U-net: Convolutional networks for biomedical image segmentation”. In: International Conference on Medical image computing and computer-assisted intervention. Springer. 2015, pp. 234–241.
8. Caroline A Schneider, Wayne S Rasband, and Kevin W Eliceiri. “NIH Image to ImageJ: 25 years of image analysis”. In: Nature methods 9.7 (2012), pp. 671–675.

Authors Introduction

Rishav Bose

Rishav Bose is a Research Assistant at the Center for Playware, Technical University of Denmark (DTU). He completed his masters degree in Electrical Engineering from DTU in October 2021, where he wrote his master’s thesis on Optimizing Plant Growth Recipes in Growbots with Machine Learning. His main research interests are food computing, cyber agriculture, and computer vision.

Henrik Hautop Lund

Professor, Ph.D. Henrik Hautop Lund, Center for Playware, Technical University of Denmark, is World Champion in RoboCup Humanoids Freestyle 2002, and has more than 200 scientific publications and 5000 citations. He was awarded the Most Outstanding Healthcare Innovator in the World in 2019. Over the last year, he received international awards in Tokyo, Singapore, and London. He has developed shape-shifting modular robots, presented to the emperor of Japan, and has collaborated closely on robotics and AI with companies like LEGO, Kompan, BandaiNamco, Microsoft, Mizuno. He is the inventor of the Moto Tiles (www.moto-tiles.com), which are used by seniors all around the world.

Generation of Checkered Pattern Images Using Prewitt Filter from RGB-D Images

Toru Hiraoka
Department of Information Systems, University of Nagasaki,
1-1-1, Manabino Nagayo, Nishisonogi, Nagasaki, 815-2195, Japan
E-mail: hiraoka@sun.ac.jp

Ryosuke Takaki
Division of Computer Science, Graduate School of Regional Design and Creation, University of Nagasaki
1-1-1, Manabino, Nagayo-chou, Nishisonogi-gun, Nagasaki-ken, 815-2195, Japan
E-mail: mc220002@sun.ac.jp

Abstract

A non-photorealistic rendering (NPR) method for generating checkered pattern images from gray-scale photographic images using Prewitt filter with an expanded window size has been proposed. In this paper, we propose an extension of the conventional method to apply to RGB-D images. Our method can change the size of the checkered patterns depending on the depth. To verify the effectiveness of our method, we conducted experiments that are visually confirmed the checkered patterns by changing the parameters in our method.

Keywords: non-photorealistic rendering, checkered pattern, RGB-D image, Prewitt filter

1. Introduction

NPR[1,2] combines computer graphics with artistic techniques, and generates non-photorealistic images from photographic images, videos and three dimensional data. An NPR method for generating checkered pattern images from gray-scale photographic images using Prewitt filter with an expanded window size has been proposed[3]. Checkered pattern images are expressed by superimposing checkered patterns on gray-scale photographic images. Since it is now possible to acquire RGB-D images with smartphones, the visual effect of checkered pattern images is expected to be improved by making the conventional method applicable to RGB-D images. RGB-D images have red, green and blue values and depth.

We in this paper propose an extension of the conventional method to apply to RGB-D images. The checkered patterns generated by our method dose not vary with RGB values, and the size of the checkered patterns varies with the depth. To verify the effectiveness of our method, we conducted experiments to visually confirm the checkered patterns generated by changing the values of the parameters in our method. The experimental result show that our method can generate the checkered pattern images without any shift in the checkered patterns depending on RGB values, and that our method can change the size of the checkered patterns by changing the values of the parameters.

2. Our method

Our method is implemented in two steps. Step 1 calculates the gradients of the gray-scale pixel values obtained the RGB pixel values using Prewitt filter. Step 2 converts the RGB videos using the gradients by changing the window size according to the depth. By repeating Steps 1 and 2, checkered pattern images of our method are generated.

The detailed procedure of our method is as follows.

Step 0: The input pixel values $(\mathrm{R}, \mathrm{G}, \mathrm{B})$ and the depth for spatial coordinates (i, j) of an RGB-D image are denoted by $f_{\mathrm{R},i,j}$, $f_{\mathrm{G},i,j}$, $f_{\mathrm{B},i,j}$ and $f_{\mathrm{D},i,j}$, qq

respectively. Subsequently, the pixel values of the image at the tth iteration number are denoted by $f^{(t)}_{\mathrm{R},i,j}$, $f^{(t)}_{\mathrm{G},i,j}$ and $f^{(t)}_{\mathrm{B},i,j}$, where $f^{(1)}_{\mathrm{R},i,j} = f_{\mathrm{R},i,j}$, $f^{(1)}_{\mathrm{G},i,j} = f_{\mathrm{G},i,j}$ and $f^{(1)}_{\mathrm{B},i,j} = f_{\mathrm{B},i,j}$. The pixel values $f^{(t)}_{\mathrm{R},i,j}$, $f^{(t)}_{\mathrm{G},i,j}$ and $f^{(t)}_{\mathrm{B},i,j}$ have value of U gradation from 0 to $U-1$. The depths $f_{\mathrm{D},i,j}$ are stored in cm and the unit of depths $f_{\mathrm{D},i,j}$ is meters. At each pixel, the window sizes $W_{i,j}$ that determine the size of the checkered patterns are calculated in the following equation:

$$W_{i,j} = W_{\max} - \frac{(W_{\max}-W_{\min})(f_{\mathrm{D},i,j}-f_{\mathrm{D,min}})}{f_{\mathrm{D,max}}-f_{\mathrm{D,min}}} \quad (1)$$

where $f_{\mathrm{D,min}}$ and $f_{\mathrm{D,max}}$ are respectively the minimum and maximum values in the depth $f_{\mathrm{D},i,j}$, and $W_{\min}$ and $W_{\max}$ are respectively the minimum and maximum window sizes set by the user. The smaller the values of the depth $f_{\mathrm{D},i,j}$, the larger the size of the checkered patterns.

Step 1: The gray-scale pixel values $f^{(t)}_{i,j}$ are calculated in the following equation:

$$f^{(t)}_{i,j} = \frac{f^{(t)}_{\mathrm{R},i,j}+f^{(t)}_{\mathrm{G},i,j}+f^{(t)}_{\mathrm{B},i,j}}{3} \quad (2)$$

The gradients of the pixel values $g^{(t)}_{\mathrm{x},i,j}$ and $g^{(t)}_{\mathrm{y},i,j}$ are calculated using Prewitt filter with the expanded window in the following equations:

$$g^{(t)}_{\mathrm{x}',i,j} = \sum_{l=j-W_{i,j}}^{j+W_{i,j}} \left(f^{(t)}_{i-W_{i,j},l} + f^{(t)}_{i+W_{i,j},l}\right) \quad (3)$$

$$g^{(t)}_{\mathrm{y}',i,j} = \sum_{k=i-W_{i,j}}^{i+W_{i,j}} \left(f^{(t)}_{k,j-W_{i,j}} + f^{(t)}_{k,j+W_{i,j}}\right) \quad (4)$$

$$g^{(t)}_{i,j} = \sqrt{{g^{(t)}_{\mathrm{x}',i,j}}^2 + {g^{(t)}_{\mathrm{y}',i,j}}^2} \quad (5)$$

$$g^{(t)}_{\mathrm{x},i,j} = \frac{g^{(t)}_{\mathrm{x}',i,j}}{g^{(t)}_{i,j}} \quad (6)$$

$$g^{(t)}_{\mathrm{y},i,j} = \frac{g^{(t)}_{\mathrm{y}',i,j}}{g^{(t)}_{i,j}} \quad (7)$$

where k and l are the positions in the window.

Step 2: The output pixel values $f^{(t+1)}_{\mathrm{R},i,j}$, $f^{(t+1)}_{\mathrm{G},i,j}$ and $f^{(t+1)}_{\mathrm{B},i,j}$ are calculated using the gradients of the pixel values $g^{(t)}_{\mathrm{x},i,j}$ and $g^{(t)}_{\mathrm{y},i,j}$ in the following equations:

$$f^{(t+1)}_{\mathrm{R},i,j} = \begin{cases} f_{\mathrm{R},i,j} + a g^{(t)}_{\mathrm{x},i,j} & (t \bmod 2 = 0) \\ f_{\mathrm{R},i,j} + a g^{(t)}_{\mathrm{y},i,j} & (t \bmod 2 = 1) \end{cases} \quad (8)$$

$$f^{(t+1)}_{\mathrm{G},i,j} = \begin{cases} f_{\mathrm{G},i,j} + a g^{(t)}_{\mathrm{x},i,j} & (t \bmod 2 = 0) \\ f_{\mathrm{G},i,j} + a g^{(t)}_{\mathrm{y},i,j} & (t \bmod 2 = 1) \end{cases} \quad (9)$$

$$f^{(t+1)}_{\mathrm{B},i,j} = \begin{cases} f_{\mathrm{B},i,j} + a g^{(t)}_{\mathrm{x},i,j} & (t \bmod 2 = 0) \\ f_{\mathrm{B},i,j} + a g^{(t)}_{\mathrm{y},i,j} & (t \bmod 2 = 1) \end{cases} \quad (10)$$

where a is a positive constant. If $f^{(t+1)}_{\mathrm{R},i,j}$, $f^{(t+1)}_{\mathrm{G},i,j}$ and $f^{(t+1)}_{\mathrm{B},i,j}$ are less than 0, then $f^{(t+1)}_{\mathrm{R},i,j}$, $f^{(t+1)}_{\mathrm{G},i,j}$ and $f^{(t+1)}_{\mathrm{B},i,j}$ must be set to 0, respectively. If $f^{(t+1)}_{\mathrm{R},i,j}$, $f^{(t+1)}_{\mathrm{G},i,j}$ and $f^{(t+1)}_{\mathrm{B},i,j}$ are greater than $U-1$, then $f^{(t+1)}_{\mathrm{R},i,j}$, $f^{(t+1)}_{\mathrm{G},i,j}$ and $f^{(t+1)}_{\mathrm{B},i,j}$ must be set to $U-1$, respectively.

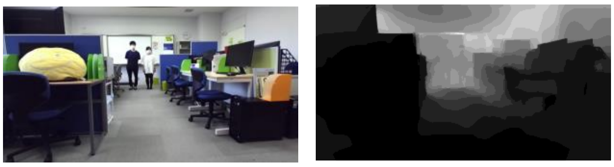
A checkered pattern image is obtained after Steps 1 and 2, which involves T iterations.

3. Experiments

An RGB-D image was obtained using ZED stereo camera. In the following experiments, the RGB-D image in Fig. 1 was used: the left and right sides of Fig. 1 are the RGB and depth images, respectively. In the depth image, a white area indicates a greater distance. The RGB-D image comprised 1280 * 720 pixels and 256 gradations.

Checkered pattern images generated by changing the iteration number T was set to 5, 10, 25 and 50. The other parameters a, $W_{\min}$ and $W_{\max}$ were set to 60, 2 and 7, respectively. The checkered pattern images generated under these conditions are shown in Fig. 2. As the iteration number T increased, the checkered patterns became clearer.

Checkered pattern images generated by changing the parameter a was set to 20, 40, 60 and 80. The other parameters T, $W_{\min}$ and $W_{\max}$ were set to 50, 2 and 7,



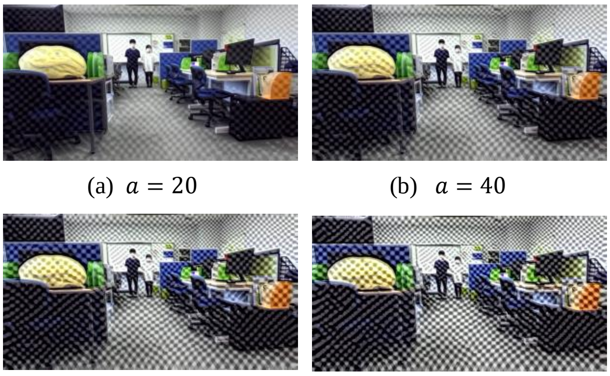
(a) RGB image (b) Depth image

Fig. 1. RGB-D image.

(a) $T = 5$ (b) $T = 10$

(c) $T = 25$ (d) $T = 50$

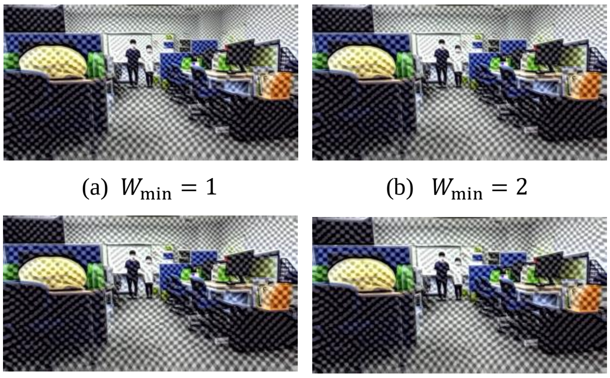
Fig. 2. Checkered pattern images generated by changing the iteration number T.



(a) $a = 20$ (b) $a = 40$

(c) $a = 60$ (d) $a = 80$

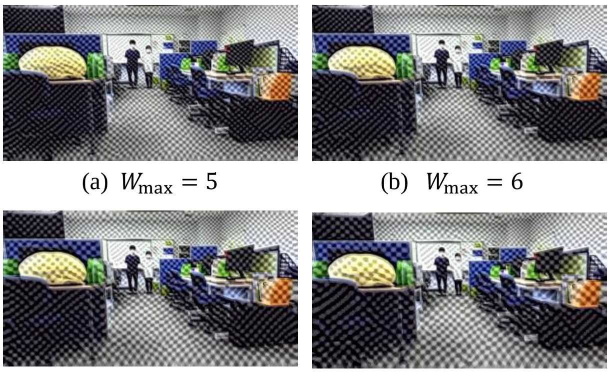
Fig. 3. Checkered pattern images generated by changing the parameter a.



(a) $W_{\min} = 1$ (b) $W_{\min} = 2$

(c) $W_{\min} = 3$ (d) $W_{\min} = 4$

Fig. 4. Checkered pattern images generated by changing the window size $W_{\min}$.



(a) $W_{\max} = 5$ (b) $W_{\max} = 6$

(c) $W_{\max} = 7$ (d) $W_{\max} = 8$

Fig. 5. Checkered pattern images generated by changing the window size $W_{\max}$.

respectively. The checkered pattern images generated under these conditions are shown in Fig. 3. The larger the parameter a, the deeper were the checkered patterns. In all checkered pattern images, the checkered patterns did not vary with RGB values.

Checkered pattern images generated by changing the window size $W_{\min}$ was set to 1, 2, 3 and 4. The other parameters T, a and $W_{\max}$ were set to 50, 60 and 7, respectively. The checkered pattern images generated under these conditions are shown in Fig. 4. The smaller the window size $W_{\min}$, the smaller were the size of the checked patterns in the distance. In all checkered pattern images, the checkered patterns did not vary with RGB values.

Checkered pattern images generated by changing the window size $W_{\max}$ was set to 5, 6, 7 and 8. The other parameters T, a and $W_{\min}$ were set to 50, 60 and 2, respectively. The checkered pattern images generated under these conditions are shown in Fig. 5. The larger the window size $W_{\max}$, the bigger were the size of the checked patterns nearby. In all checkered pattern images, the checkered patterns did not vary with RGB values.

4. Conclusion

We proposed an NPR method for generating checkered pattern images from RGB-D images. Our method was executed by an iterative process using Prewitt filter with an expanded window size. To verify the effectiveness of our method, the changes in the checkered pattern images by changing the values of the parameters were investigated. Using our method, the checkered pattern images could be generated without any shift in the checkered patterns depending on RGB values, and that the size of the checkered patterns could be changed by changing the values of the parameters.

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.* **24**(4), 1990, pp. 207-214.

qq

2. D. D. Seligmann and S. Feiner, Automated generation of intent-based 3D illustrations, *ACM SIGGRAPH Computer Graphics.* **25**(4), 1991, pp. 123-132.
3. T. Hiraoka, Generation of checkered pattern images by iterative calculation using Prewitt filter with expanded window size, *IEICE Transactions on Information and Systems.* **E103-D** (11), 2020, pp. 2407-2410.

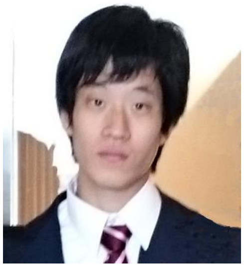
==

Authors Introduction

Dr. Toru Hiraoka

He graduated Doctor course at Design in Kyushu Institute of Design. He is a Professor in the Faculty of Information Systems in University of Nagasaki. His current research interests are non-photorealistic rendering and disaster prevention.

Mr. Ryosuke Takaki



He is a student in the Division of Computer Science, Graduate School of Regional Design and Creation in University of Nagasaki. His current research interests are non-photorealistic rendering and image processing.

==

An Approach of Analyzing Movement Patterns Using Word Embeddings from Geo-tagged Tweets

Masaharu Hirota
Department of Information Science, Okayama University of Science
1-1 Ridaicho, Kita-ku, Okayama-shi, 700-0005, Japan

Tetsuya Oda
Department of Information and Computer Engineering, Okayama University of Science
1-1 Ridaicho, Kita-ku, Okayama-shi, 700-0005, Japan
E-mail: hirota@mis.ous.ac.jp, oda@ice.ous.ac.jp

Abstract

Many people share content about daily events on social media sites. Understanding people's movement patterns using the contents benefits numerous applications, such as tourism recommendation, city planning, and geo-targeting. This study proposes a new approach for clustering user trajectories to discover movement patterns. Our approach generates feature vectors for movement patterns using the word embedding model to learn movements between a pair of areas quantized by their latitude and longitude. Then, our approach uses multiple models to learn each area of different sizes and integrates the generated embedding vectors. The vectors represent the relationships between the movements from one area to the next. We demonstrated that clustering results by our proposed method.

Keywords: trajectory, location information, clustering, Twitter.

1. Introduction

Because of the wide dispersion and expeditious improvement of various devices such as smartphones and tablets, diverse and vast data are generated on the web. Social media sites have become popular because users can quickly post data and various messages. Twitter[1], a social media site that provides a micro-blogging service, is used as a real-time communication tool. Numerous tweets have been posted daily by vast numbers of users. Therefore, Twitter is used to obtain, from a large amount of information posted by many users, real-time information corresponding to the real world.

Understanding human mobility patterns is important as basic information for various applications such as urban planning and transportation network analysis. To analyze these patterns, we need information on the vast movement trajectories. However, it is difficult to collect a large amount of data on the trajectory. Social media information, such as Twitter, can be used as an alternative source of information for such analysis. For example, some tweets have geo-tag with location information based on latitude and longitude. We can regard a sequence of consecutive geo-tagged tweets posted by a user as the movement trajectory. Here, information from social media such as Twitter is not a sufficient sampling of the population about age, race, ethnicity, and other attributes of people. However, the information has been used in various studies as a valuable information to analyze people's movement patterns.

Clustering of movement trajectories[2,3] is one of the methods to analyze the people's movement patterns from a large number of movement trajectories. By applying the

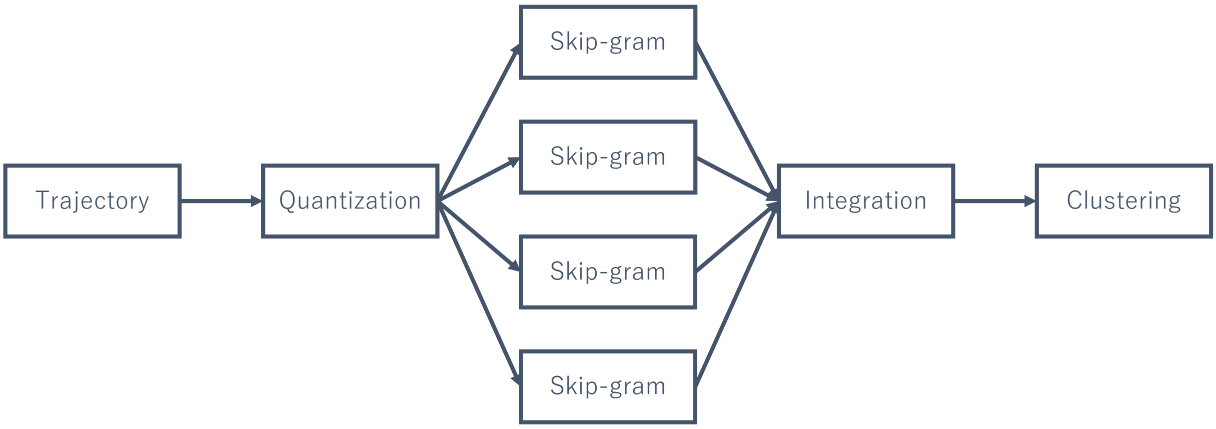


Fig. 1. Overview of our proposed method.

clustering method to the trajectories, similar trajectories are grouped called clusters. We can then regard each cluster as people's movement patterns.

This paper proposes a new approach to clustering trajectories represented by geo-tagged tweets to analyze movement patterns. Our proposed method uses a word embedding model by an improved skip-gram model[4], inspired by the skip-gram model[5], for learning user movements. The model learns movements between a pair of areas quantified by the latitude and the longitude. Then, by applying clustering to the trajectories based on the vectors generated by multiple models, we analyze the movement patterns of users between the quantized regions.

2. Literature Review

This section describes several studies that use the movement patterns obtained from social media sites such as Twitter.

Studies have been conducted to analyze behavior, destinations, and purposes of movement by analyzing the movement patterns using tweets. Jurdak et al.[6] showed that movement patterns differ according to their daily travel distance trends by analyzing tweets posted in Australia. Salas-Olmedo et al.[7] analyzed the spatial distribution of users of public spaces by analyzing the movement represented by geo-tagged tweets in a medium-sized city.

There is some study based on Twitter on the impact of the COVID-19 pandemic on the mobility of people. Huang et al.[8] proposed an index to evaluate the degree of change in people's movement and analyzed the impact of COVID-19. Jiang et al.[9] analyzed the impact of COVID-19 on changes in movement patterns for land use in New York City. These studies show that the change of movement patterns depends on the phase of the COVID-19 epidemic and the response to it by national.

3. Proposed Method

This section describes a method for clustering movement trajectories. Figure 1 shows the overview of the proposed method. The proposed method consists of the following steps.

(i). Quantize movement trajectories based on latitude and longitude.
(ii). Create multiple improved skip-gram models by learning the quantized trajectories.
(iii). Input the trajectories to each model and integrate obtained vectors.
(iv). Cluster the trajectories based on the vectors.

3.1. *Quantization of user-posted locations*

In this section, we describe how to quantize tweets. First, we sort the series of tweets of a user in post time by ascending order. Next, we regard the posting times of those tweets as different movement trajectories if the difference between the posting times of two consecutive tweets is more than 3 hours. This reason is that when a user posts tweets over a long time, we should regard them as different movements. We also delete the trajectory if the number of tweets is smaller than 5.

The next step is to convert the movement trajectories, which consist of latitude and longitude, into a continuous string. The improved skip-gram model is a method for learning continuous 1D features. Also, this process simplifies the learning while accounting for the differences due to GPS errors and the number of digits in the latitude and the longitude.

This study converts the latitude and longitude annotated to a tweet into quadkey[10]. The quadkey is a unique identifier represented by the string representation of a standard map tile on a particular zoom level. In addition, depending on the string length of the quadkey, it is possible to control the degree of accuracy loss. This study generated each quadkey with zoom levels from 14 to 17 from the tweets.

3.2. *Improved skip-gram*

Next, we describe improved skip-grams, a model for learning the movement trajectory. The conventional skip-gram learns $w(t-c)\cdots w(t-1)$ and $w(t+1)\cdots w(t+c)$ around a word $w(t)$ in a sentence. Next, the improved skip-gram learns only $w(t+1)\cdots w(t+c)$ for $w(t)$. The difference between these models is that the conventional skip-gram learns the movement to and from a region represented by w(t). In contrast, the improved skip-gram learns only the movement from one region to another represented by w(t). In this study, we create improved skip-gram models for each level of quadkey.

3.3. *Clustering trajectory*

We describe the procedure of movement trajectory clustering using the generated models. First, we convert the latitude and longitude of each trajectory into a sequence of multi-level quadkey codes. Then, each level of quadkey is input to the corresponding improved skip-gram model to generate multiple vectors. The trajectory vector is the average of the vectors obtained from the quadkeys included in the trajectory for each level of quadkey.

We apply k-means clustering, a popular clustering algorithm, to the vectors to cluster movement trajectories. In this study, we regard the result as the clustering result of the trajectory.

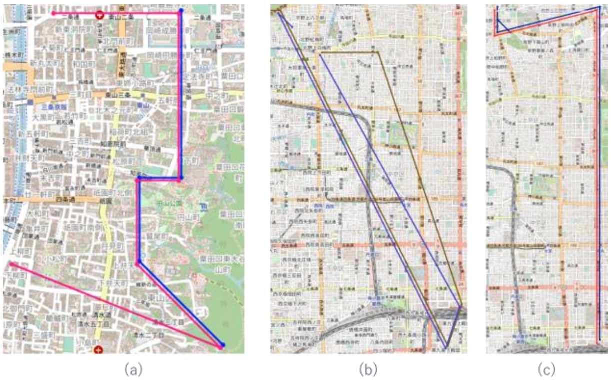


Fig. 2. Clustering results.

4. Visualization

4.1. *Dataset*

In this experiment, we used tweets, including latitude and longitude information. The tweets were posted in Kyoto, Japan, between May 21, 2015, and December 31, 2018. The number of tweets was 3,622,043, and the number of movement trajectories was 8,379.

The parameters of each improved skip-gram are that the vector dimension is 50 and that the window size is 3. We set the number of clusters k for the k-means clustering to 300.

4.2. *Result*

Figure 2(a) ~ (c) shows three clusters in the clustering result of our proposed method. Each line in those figures represents movement trajectories. A distinct color represents each trajectory. We used the OpenStreetMap[11].

In Figure 2(a), the cluster includes the trajectories for uses who mainly visited around Kiyomizu temple, Chion-in, and Okazaki Park. In Figure 2(b), the cluster includes the trajectories for uses that are mainly visited around Ritsumeikan University and Kyoto Station. In Figure2(c), the cluster includes the trajectories for uses that are mainly visited around Funaokayama Park and Kyoto Station. The trajectories included in each cluster differ mainly in the places visited. In those figures, the trajectories included in each cluster are different movement patterns. Therefore, we conclude that our proposed method has successfully clustered the trajectories.

5. Conclusion

This paper proposed a new approach for analyzing movement patterns by clustering movement trajectories using a word embedding algorithm. We presented three clusters of trajectories using our method. One of the limitations is that our method does not consider the direction of the trajectory. As a result, the proposed method regards the trajectories visited the same point as one cluster, even if they have the opposite movement direction. Therefore, future work will include the improvement of a method of word embedding.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number JP19K20418 and Grant for Promotion of OUS Research Project (OUS-RP-20-3).

References

1. "Twitter", https:/twitter.com/
2. S.J. Corrado, T.G. Puranik, O.J. Pinon, D.N. Mavris, "Trajectory Clustering within the Terminal Airspace Utilizing a Weighted Distance Function", Proceedings, Vol. 59, No. 1, 2020.
3. H. Li, J. Liu, K. Wu, Z. Yang, R. W. Liu and N. Xiong, "Spatio-Temporal Vessel Trajectory Clustering Based on Data Mapping and Density", Proceedings, IEEE Access, vol. 6, pp. 58939-58954, 2018.
4. M. Hirota, T. Oda, M. Endo, and H. Ishikawa, "Generating distributed representation of user movement for extracting detour spots" in Proceedings of the 11th International Conference on Management of Digital EcoSystems, pp. 250–255, 2019.
5. T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality", Advances in Neural Information Processing Systems, pp. 3111–3119, 2013.
6. R. Jurdak, K. Zhao, J. Liu, M. AbouJaoude, M. Cameron, D. Newth, "Understanding Human Mobility from Twitter", PLOS ONE, Vol. 10, No. 7, pp. e0131469, 2015.
7. M.H. Salas-Olmedo, C. Quezada, "The use of public spaces in a medium-sized city: from Twitter data to mobility patterns", Journal of Maps, Vol. 13, No. 1, pp. 40-45, 2017.
8. X. Huang, Z. Li, Y. Jiang, X. Li, D. Porter, "Twitter reveals human mobility dynamics during the COVID-19 pandemic", PLOS ONE, Vol. 15, pp. 1-21, 2020.
9. Y. Jiang, X. Huang, Z. Li, "Spatiotemporal Patterns of Human Mobility and Its Association with Land Use Types during COVID-19 in New York City", ISPRS International Journal of Geo-Information, Vo. 10, No. 5, pp. 334, 2020.
10. "Quadkey", https://docs.microsoft.com/en-us/azure/azure-maps/zoom-levels-and-tile-grid?tabs=csharp#quadkey-indices
11. "OpenStreetMap", https://www.openstreetmap.org

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.

Dr.Tetsuya Oda

He received a Ph.D. degree in 2016 from Fukuoka Institute of Technology (FIT). After working for the Faculty of Information Engineering, FIT, he has worked in the Faculty of Engineering, the Okayama University of Science, since April 2017. His research interests include wireless mesh networks, intelligent systems and robotics. He is a member of AAAI, ACM, DBSJ, IEEE and IPSJ.

A Proposal of a Software Defect Predication System Using SOM

Yoshihiro Kita
Faculty of Information Systems, University of Nagasaki,
1-1-1 Manabino, Nagayo, Nishisonogi , Nagasaki 851-2195, Japan

Kazuki Ueda
Nihon knowledge Co.,Ltd. 3-19-5 Kotobuki, Taito, Tokyo 111-0042, Japan

Kiyotaka Sakurai
Nihon knowledge Co.,Ltd. 3-19-5 Kotobuki, Taito, Tokyo 111-0042, Japan

E-mail: kita@sun.ac.jp, ueda@know-net.co.jp, sakurai@know-net.co.jp

Abstract

The goal of software testing is to detect all latent defects. However, it is difficult to know how many latent defects remain and where they are hidden. In this research, we propose a system that analyzes the characteristics and tendencies of already detected defects and predicts where the latent defects are likely to be. Specifically, the system inputs the data of detected defects into a Self-Organizing Map (SOM) and predicts the locations that contain many defects from this map. To confirm the validity of this proposal, we input past defect data into the SOM, analyze the trend of defects, and evaluate the predictability of the latent defects.

Keywords: software testing, exploratory testing, defect predication, self-organizing map (SOM).

1. Introduction

The goal of software testing is to detect all latent defects, but it is impossible to achieve that goal[1]. However, residual defects are causes of serious lost, so it is desirable to remove defects as much as possible.

Exhaustive or exploratory testing methods are needed to find latent defects in software. Especially in exploratory testing, it is difficult to the location of latent defects while deciding where to search next.

Mr. Ueda as the second author, proposed the testing method “FaRSeT (Flexible and Rapid Software Test)” [2] which is an exploratory testing method that utilizes test analysis by mind mapping to cope with projects have short delivery time, and frequent specification changes. This method reduces the time and effort required to prepare test cases in advance and allows us to prioritize important searching points for testing while obtaining the agreement of stakeholders. However, this method has two problems as follows.

- The searching points are not possible to determine as important, cause no defects occurred in the tests that are not executed.
- It is difficult to assume the remaining defects by the number of defects that are found only.

Function	Function completeness	Function correctness	Function appropriateness	Co-existence	Interoperability	Appropriateness recognizability	Learnability	Operability	User error protection	User interface aesthetics	Accessibility	Confidentiality
A	10	12	9	0	1	1	0	4	5	2	0	0
B	0	0	0	0	0	0	0	0	0	0	0	0
C	5	6	1	0	0	3	0	3	1	0	0	0
D	4	9	1	0	0	0	0	2	0	1	0	0
E	2	7	2	0	0	0	0	2	0	0	1	0
F	3	8	2	0	0	0	0	1	0	3	0	0
G	2	3	2	2	1	1	0	4	1	1	1	0
H	0	3	0	0	0	1	0	0	1	0	0	0
I	9	23	4	0	2	5	0	6	1	2	0	0
J	1	2	0	0	0	0	0	3	3	0	0	0
K	1	2	0	1	0	0	0	0	0	0	0	0

Fig.1. A sample of an exploratory testing matrix

In this paper, we propose a method “FaRSeT-#” which infer the important searching points, and a defect predication system that incorporates this method. In order to infer the important searching points, we use Self-Organizing Map (SOM) [3], and visualize them on a two-dimensional map.

2. Related Works

2.1. *FaRSeT*

FaRSeT (Flexible and Rapid Software Test) [2] is a flexible and rapid testing method that uses a mind-mapped job analysis and an exploratory testing matrix to determine the priority of searching points for testing in a current project. The process of FaRSeT is shown as follows.

(i) Conduct job analysis using a mind map.

(ii) Create the test charters which are broken down from the quality characteristics of software, based on the job analysis in step (i).

(iii) Create the table which is named “the exploratory testing matrix”, with the test charters as the horizontal axis, and the function as the vertical axis.

(iv) Define the intersection as a “session”, of each item in the exploratory testing matrix, and execute the exploratory testing while obtaining the agreement of stakeholders for high importance sessions.

(v) If defects are found in Step (iv), the number of defects should be listed in the session of the exploratory testing matrix.

(vi) Repeat Step (iv) to Step (v) while deciding the next session for the exploratory testing based on the number of defects listed in the session. For example, in Fig.1, the sessions colored in red, are the next session for the exploratory testing.

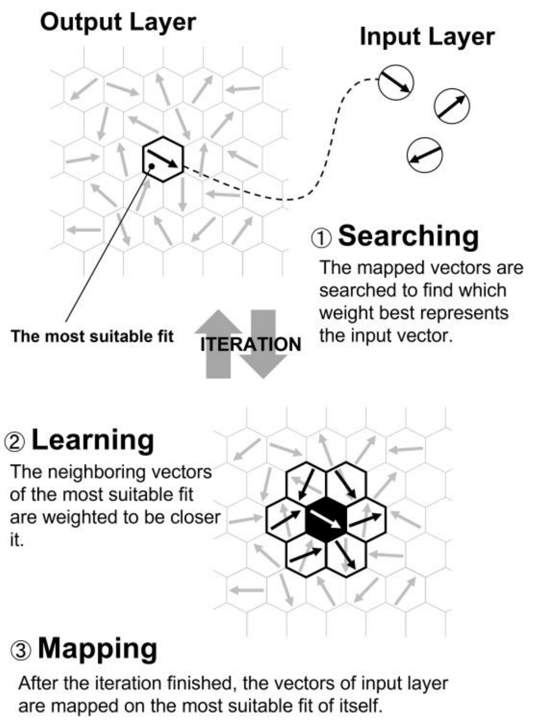


Fig.2. The learning process of SOM

2.2. *Self-Organizing Map*

Self-Organizing Map (SOM) [3] is an unsupervised competitive neural network. SOM is a machine learning model that represents the similarity of given multidimensional data on a two-dimensional map as shown as Fig. 2.

SOM consists of two layers, an input layer, and an output layer. The input layer contains nodes with input vectors. The number of their nodes is the same as the number of samples of data used as input. The output layer

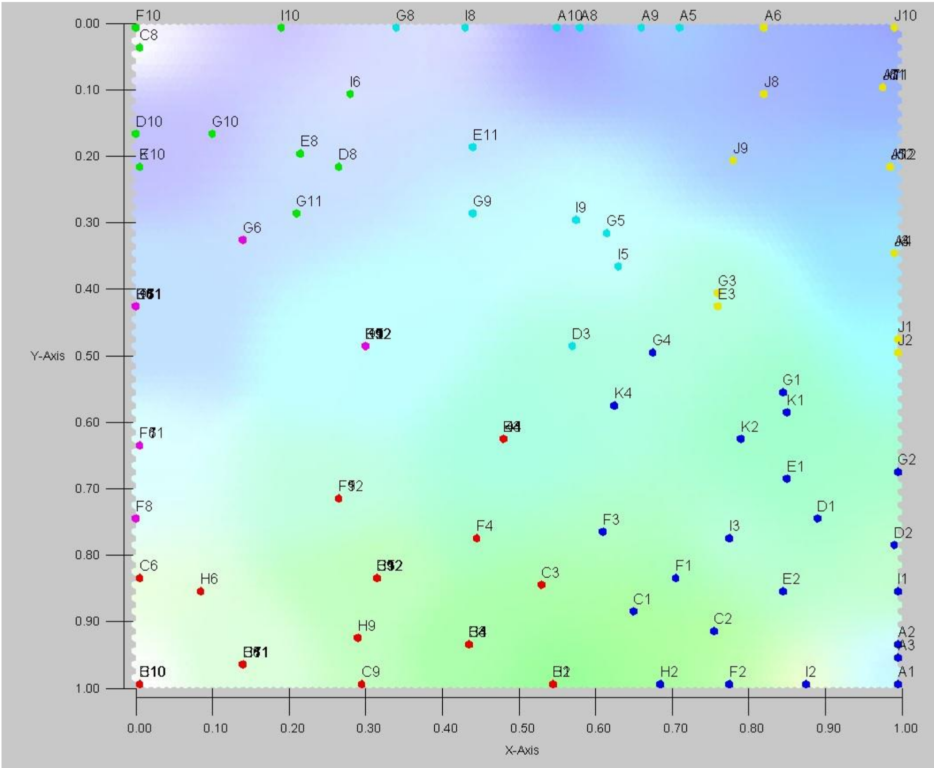


Fig.3. SOM with the session of the exploratory testing matrix in Fig. 1 as input

contains nodes which are arranged hexagonal honeycomb, and each node initially holds a random reference vector. The learning process of SOM is shown as follows.

(i) Compare an input vector to all the reference vectors in the output layer.

(ii) The node which has the most similar reference vector to the input vector is defined “The most suitable fitting node”.

(iii) The reference vector of the nodes which are adjoined the most suitable fitting node, is weighted to be close the input vector.

(iv) Iterate the learning process until the learning times is satisfied.

(v) After the iteration finished, the vectors of input layer are mapped on the most suitable fit of itself.

The nodes of input layer are placed close to the nodes that are similar to each other, and are placed far from the nodes that are not similar to each other. Therefore, the similarity between the samples is visualized in two-dimensional space as the distance between the nodes of input layer.

3. Software Defect Predication System

In this paper, we propose a method “FaRSeT-#” which infer the important searching points, and a defect predication system that incorporates this method. FaRSeT-# is an extended method of FaRSeT by incorporating SOM. The vectors of all session from the exploratory testing matrix input to SOM. The mapped nodes of session group into colored clusters in SOM. In this way, the sessions belonging to the same cluster are similar, which means that they have the commons: defect tendencies, target areas of search, and effective testing methods. The testers analyze the sessions of each cluster, and predict the priority of searching points, and the location of latent defects

Fig. 3 shows the results of training the exploratory test matrix in Fig. 1 using SOM, mapped to a two-dimensional map. For example, the label "A5" represents the row "Function A" and the column "Test Charter 5 (Interoperability in this case)" in the exploratory test matrix.

Function	Function completeness	Function correctness	Function appropriateness	Co-existence	Interoperability	Appropriateness recognizability	Learnability	Operability	User error protection	User interface aesthetics	Accessibility	Confidentiality
A	10	12	9	0	1	1	0	4	5	2	0	0
B	0	0	0	0	0	0	0	0	0	0	0	0
C	5	6	1	0	0	3	0	3	1	0	0	0
D	4	9	1	0	0	0	0	2	0	1	0	0
E	2	7	2	0	0	0	0	2	0	0	1	0
F	3	8	2	0	0	0	0	1	0	3	0	0
G	2	3	2	2	1	1	0	4	1	1	1	0
H	0	3	0	0	0	1	0	0	1	0	0	0
I	9	23	4	0	2	5	0	6	1	2	0	0
J	1	2	0	0	0	0	0	3	3	0	0	0
K	1	2	0	1	0	0	0	0	0	0	0	0

Fig.4. An exploratory testing matrix with sessions colored by cluster.

Function	Function completeness	Function correctness	Function appropriateness	Co-existence	Interoperability	Appropriateness recognizability	Learnability	Operability	User error protection	User interface aesthetics	Accessibility	Confidentiality
A	10	12	9	0	1	1	0	4	5	2	0	0
B	0	0	0	0	0	0	0	0	0	0	0	0
C	5	6	1	0	0	3	0	3	1	0	0	0
D	4	9	1	0	0	0	0	2	0	1	0	0
E	2	7	2	0	0	0	0	2	0	0	1	0
F	3	8	2	0	0	0	0	1	0	3	0	0
G	2	3	2	2	1	1	0	4	1	1	1	0
H	0	3	0	0	0	1	0	0	1	0	0	0
I	9	23	4	0	2	5	0	6	1	2	0	0
J	1	2	0	0	0	0	0	3	3	0	0	0
K	1	2	0	1	0	0	0	0	0	0	0	0

Fig.5. An exploratory testing matrix with a gray scale representing the priority of the searching points.

The data to be input to the SOM is the session information consisting of the following items:

- Metrics related to the function
  (e.g., size of the function, skills of the developer, etc.)
- Metrics related to test charter
  (e.g., importance of quality characteristics)
- Metrics related to sessions
  (e.g., number of defects, recent test results, test execution time, skills of testers, etc.)

The nodes on the SOM are divided into clusters by k-means method [4], and each cluster is color-coded with six colors (RED, BLUE, GREEN, YELLOW, CYAN, and MAGENTA). In Fig. 3, the blue area indicates that the scale of the function is large, and the green area indicates that the importance of the quality characteristic is high. In this map, the lower right and upper right areas are darker in color, and the center to upper left areas are lighter in color.

we analyze the priority of the clusters from the colors of this map, the priority order is determined as follows

(i) BLUE
(ii) YELLOW
(iii) RED
(iv) CYAN and MAGENTA
(v) GREEN

Next, all sessions colored in each cluster to the exploratory testing matrix in Fig.1, which is shown in Fig.4. Also, Fig. 5 shows the priority as a gray scale shading. The darker sessions are the higher the priority for the next test.

Comparing Fig. 1 and Fig. 5, it can be confirmed that this method enables a finer prediction of the session containing the defect and makes it easier to determine the next area to search.

4. Evaluation of validity

We applied this system to a development project in order to confirm the validity of the proposed method. First, we analyze the sessions that contain defects from the results of the exploratory testing by using our proposal system. Then, we perform the next exploratory testing, and analyze the number of defects for each session. Finally, these results are compared with the results of our proposal system.

Table.1 is shown the number of found defects for each testing and priority order to search in next testing. The priority predicted by our proposal was highest for GREEN, followed by MAGENTA and RED. CYAN had no defects found, but its

Table.1. The number of found defects for each testing and priority order to search in next testing.

Cluster	Number of found defects (Average of sessions)	Priority order to search in next testing	Number of found defect in next testing (Average of sessions)
RED	1.50	Third	0.00
BLUE	0.90	Fifth	0.06
GREEN	3.00	First	0.50
YELLOW	0.00	Fifth	0.00
CYAN	0.00	Fourth	0.09
MAGENTA	1.70	Second	0.11

priority was predicted to be close to MAGENTA, so it was ranked higher than BLUE and YELLOW.

The results of the next exploratory test showed that the defects were found in GREEN and MAGENTA as predicted. Defects were also found in CYAN, where no defects had been found. This indicates that the proposed method can predict the defects. However, no defects were found in RED, which is the third priority. This point needs to be verified in the future, including the selection of the data to be input to SOM.

5. Conclusion

In this paper, we proposed a method “FaRSeT-#” which infer the important searching points, and a defect predication system that incorporates this method.

We confirmed that it can objectively determine the search location and clearly state the basis for the decision mechanically mapping and clustering the data by the system using FaRSeT-#. In the future, it is necessary to verify the selection of data to be input to the SOM.

References

1. “Certified Tester Foundation Level Syllabus Version 2018 V3.1”, International Software Testing Qualifications Board, 2018.
2. K. Ueda, J. Tanba, and S. Kudo, “Effort of application of test method “FaRSeT (Flexible and Rapid Software Test)” for short-term delivery development project”, Software Quality Symposium 2018 (SQiP2018), pp.1-8, 2018 (in Japanese).
3. T. Kohonen, “Self-Organizing Maps”, Springer-Verlag, 2001.
4. D. Steinley, M.J. Brusco, “Initializing k-means Batch Clustering: A Critical Evaluation of Several Techniques”, Journal of Classification, Vol.24, No.1, pp.99-121, 2007.

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. Kazuki Ueda

He belongs Nihon Knowledge Co.,Ltd. His job is chief executive officer of verification technology. He research interest in development of quality assurance technology for test automation, exploratory testing, and AI products.

Mr. Kiyotaka Sakurai

He belongs the officer of verification technology, Nihon Knowledge Co.,Ltd. He verify the automatical tools, and research the operatable automatical tool by low-code.

Mapping the Motion of Highly-inclined Triple System into a Secular Perturbation Model

Masaya M. Saito
Department of Information Security, University of Nagasaki, Siebold, 1-1-1 Manabino, Nagayo-cho, Nishi-Sonogi-gun, Nagasaki 851-2195, Japan

Kiyotaka Tanikawa
National Astronomical Observatory of Japan,
2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan
E-mail: saitohm@sun.ac.jp, tanikawa.ky@nao.ac.jp

Abstract

The motions of three bodies like Sun-Asteroid-Jupiter system or triple star system are formalized as hierarchical three body problem. When the third body orbits around the rest in a highly inclined elliptic orbit, the system undergoes the oscillation, called the Kozai oscillation, where the eccentricity may increase with decrease of the inclination of the orbital plane.

While the Kozai oscillation seems to be a key process in orbital evolution, including disruption of triple system, its reflection into actual trajectories is quite complicated to analyze. For this reason, we try to map these trajectories into a secular perturbation model with data assimilation and demonstrate the extraction of state and its transition (libration to circulation and vice versa) as the Kozai oscillation.

Keywords: List four to six keywords which characterize the article.

1. Introduction

The motion of gravitationally interacting three bodies is called the three-body problem and known to be analytically unsolvable unlike to the two-body problem, in which the bodies draw elliptic orbits. There are two contrastive types in motion of three bodies. In one type, the bodies interact each other in an extremely complicated way and often end in the disruption into a pair of bodies and the rest one. This type of motion is generally called chaotic. The other type of motion is that bodies draw hierarchical elliptic orbits, each of which is similar to that of two-body problem, with gradual change in its shape. The latter may be realized when there is an enough contrast in the masses (e.g. the Sun occupies 99.8% of the entire solar system's mass) or in orbital radii of the inner and outer orbits.

Our interest is in between the two. If two elliptic orbits (a schematic illustration shown in see Fig. 1) are initially placed close to each other, the outer orbit evolve to a hyperbolic orbit and the third-body escapes on it. How close the system can be allocated without such a disruption of the system is called stability limits and has been studied since Harrington[1]. While the contribution of Kozai mechanism, explained later, to the instability of hierarchical triple systems has been pointed out[2], the process until the system are finally broken is still unclear. Partially, the difficulty comes from a necessity of a long numerical orbit to see the sign of instability as well as a complicated process being involved till the disintegration of the system. In this study, we construct a mapping from the motion of three bodies in a Cartesian flame into orbital elements by utilizing the most simple secular perturbation model as auxiliary dynamical system.

2. Method

2.1. *Equation of Motion*

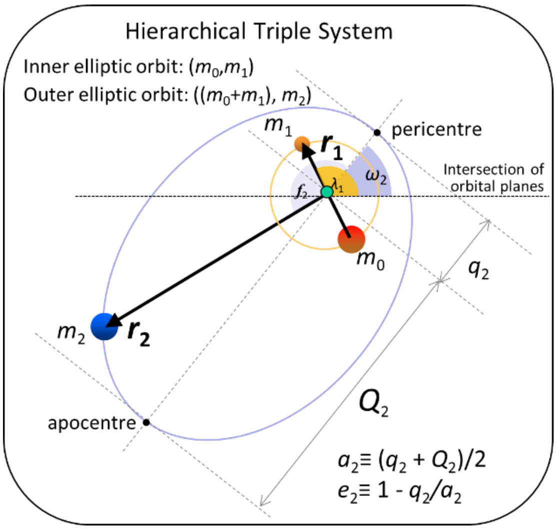
Let m_0, m_1 and m_2 be the masses of bodies gravitationally interacting three bodies. We introduce Jacobian coordinates to describe their motion, that is, $\boldsymbol{r}_1$ is the vector from m_0 to m_1, and $\boldsymbol{r}_2$ from their barycenter to m_2. Equations of motion of these bodies are given by

$$\mu_i^* \frac{d^2 \boldsymbol{r}_i}{dt} = -\frac{G\mu_i^* \sum_{j=0}^{i} m_j}{||\boldsymbol{r}||^2} + \frac{\partial R}{\partial \boldsymbol{r}_i} \qquad (1)$$

$(i = 1{,}2)$ with disturbing function

$$R = \frac{Gm_0 m_1}{||\boldsymbol{r}_{01}||} + \frac{Gm_2 m_0}{||\boldsymbol{r}_{02}||} - \frac{Gm_2(m_0+m_1)}{||\boldsymbol{r}_2||},$$

where vectors $\boldsymbol{r}_{ij}$ from m_0 to m_1 and reduced masses $\mu_1^* = \frac{m_0 m_1}{m_0+m_1}, \mu_2^* = \frac{m_2(m_0+m_1)}{m_0+m_1+m_2}$.



2.2. *Secular Perturbation and Kozai Oscillation*

A solution $\boldsymbol{r}_1(t)$ and $\boldsymbol{r}_2(t)$ of the equations of motion Eq. (1) defined above generally draw nearly elliptic orbits in short-term, changing gradually their shape change in long-term. When our interest is long-term evolution of the system, it is good to rewrite the equations of motion with respect to variables describing the orbital shapes, called orbital elements, under a certain approximation neglecting short-term variation.

Orbital elements consist of $a_i, e_i, q_i := a_i(1 - e_i), \omega_i$, corresponding to $\boldsymbol{r}_i$ and its derivative, as well as angle I between the orbital planes on which $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ are (see Fig.1 for geometrical definitions). With these variables, the disturbing function is rewritten as

$$R = \frac{Gm_2\alpha^2}{16a_2}[(3\cos^2 I - 1)(2 + 3e_1^2 + 3e_2^2) + 15e_1^2 \sin^2 I \cos 2\omega_1], \qquad (2)$$

where $\alpha \coloneqq a_1/a_2$ and e_i are kept up to their second order. Equation (1) is accordingly transformed to the first order differential equations w.r.t. these orbital elements, called planetary equations (Note that we have derived Eq.(2) using the algorithm[5] for computer algebra aiming at a higher order expansion for future study. For this reason, its exact form is slightly different than its traditional form[4]).

The equations derived by Eq. (2) has only solutions such that ω_1, indicating the direction of the pericenter measured from the intersection of two orbital planes, circulates (i.e. ω_1 takes all possible values from $0°$ to $360°$) if the inclination angle I of one plane against the other, is small, whereas another type of solutions exists for sufficiently high I, where ω_1 oscillates a limited range including $+90°$ or $-90°$. The former type of motion is called the circulation, and the latter the libration. When I is high so that the libration is possible, e_1 significantly rise up or down along with the entire period of ω_1, as well as I varies anti-correlatedly to e_1. This oscillation of e_1 and I is called Kozai mechanism[4], which is originally studied by Yoshihide Kozai for the cases of $m_1 \to 0$.

2.3. *Introduction of Stochastic change*

While we aim at extract such a process into the disintegration as a variation of orbital elements and consider mapping the outcome of the full model Eq. (1) to the perturbation model Eq. (2), the discrepancy is not small between them. For this reason, we introduce a stochastic process and allow the solutions of Eq. (2) to jump at each times step by adding the realization of random variables. Specifically, an extended system

$$(e_{1,n}, \omega_{1,n}) = \mathrm{RK4}(e_{1,n-1} + \delta e_{1,n}, \omega_{1,n} + \delta\omega_{1,n}) \qquad (3)$$
$$\delta e_{1,n} \sim N(0, \sigma_e^2), \delta\omega_{1,n} \sim N(0, \sigma_\omega^2),$$

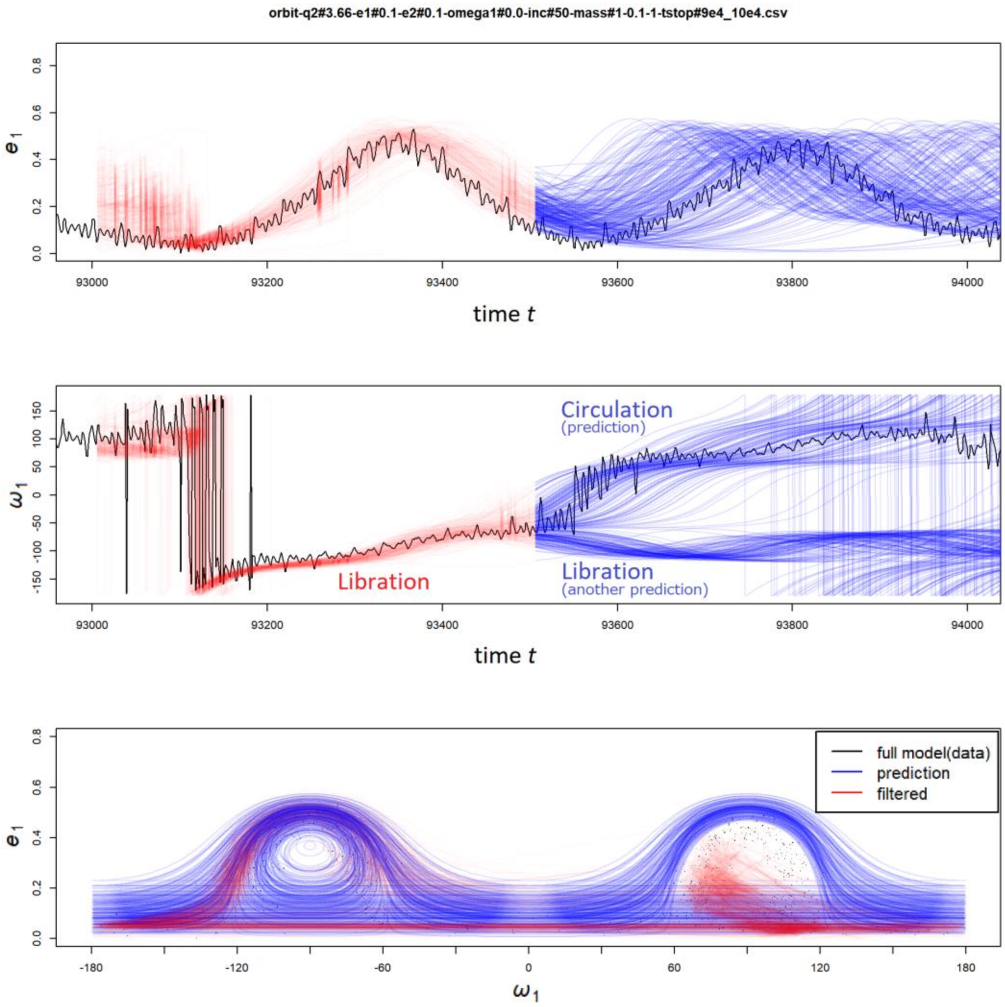


Fig. 2. Mapping of a solution of equations of motion to secular perturbation model. The solution of Eq (1) is shown in black (full model), corresponding mapped trajectories to secular perturbation model using particle filtering in red (filtered), and propagation from the last filtered time t = 93,500 in blue (prediction). Parameters are $m_1 = 0.1, m_0 = m_2 = 1,\ q_2/a_1 = 3.66,\ e_1 = e_2 = 0.1, I = 50°, \omega_1 = 0°, \omega_2 = 90°$.

is used instead of Eq.(2), where RK4(e_1, ω_1) is a propagator which advances the time by a given amount Δt, following to the planetary equations generated from Eq. (2), with parameters $\Delta t = 1$, $\sigma_e = \sigma_\omega/2\pi = 0.002$.

The discrepancy in (e_1, ω_1) between Eqs. (1) and(3) are measured by the likelihood for a single time point based on a normal,

$$\ln p\left(e_{1,n}^{\mathrm{obs}}, \omega_{1,n}^{\mathrm{obs}} | e_{1,n}, \omega_{1,n}\right) = \left(e_{1,n}^{\mathrm{obs}} - e_{1,n}\right)^2 + \left(\omega_{1,n}^{\mathrm{obs}} - \omega_{1,n}\right)^2/\pi^2 . \quad (4)$$

Here we regard the outcome of Eq. (1) as observation data (denoted by $e_{1,n}^{\mathrm{obs}}, \omega_{1,n}^{\mathrm{obs}}$), and that of Eq. (3) as the latent variables ($e_{1,n}, \omega_{1,n}$). Eqs. (3) and (4) form the state space model, to which sequential Bayesian estimation algorithms[6,7] are applicable. Of these algorithms, we implement the mapping from the "data" and the "latent" using Particle Filtering[6,7].

3. Results

We will demonstrate how an exact solution of the equations of motion Eq. (1) is mapped to the secular perturbation model, and that a transition between the libration and the circulation are identified.

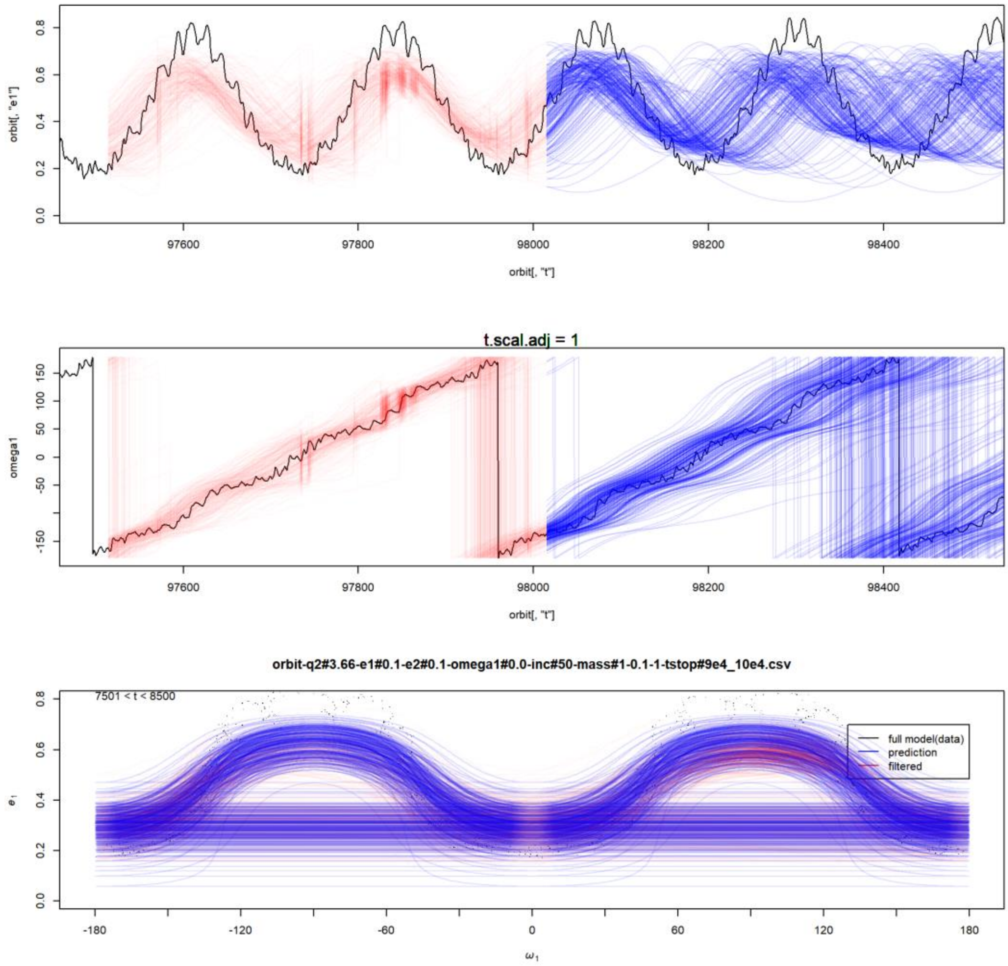


Fig. 3. The same orbit as in Fig. 2, but a later time range being covered.

An example shown in Fig. 2 includes a transition from the libration the circulation, followed by an elevation of e_1 in Fig. 3. Before going to a detailed inspection of these figures, we remark on the choice of the configuration. Masses and initial orbital parameters are chosen as $m_1 = 0.1$, $m_0 = m_2 = 1$, orbital separation $q_2/a_1 = 3.66$, eccentricities $e_1 = e_2 = 0.1$, and mutual inclination $I = 50°$ (the entire parameters are shown in the caption) so that the system undergoes the 1:6 mean motion resonance (MMR). MMRs may cause orbital instability and in fact we have observed a disruption of the system at the 1:6 MMR under more massive m_1 (specifically $m_1 = 1$). The process to the disruption is as follows: e_1 suddenly increases before increase of e_1 with e_2 after its long lasting quasi periodic variation, the increase of e_1 forces the increase of e_2, by which a_2 increases to go a hyperbolic orbit. Interestingly, the increase of e_1 often occurs when the system in a libration around $\omega_1 = \pm 90°$. This is the reason why we consider Kozai mechanism enhances the instability first realized by MMR. While the initial condition for Fig. 2 is chosen Kozai mechanism and a MMR coexists, we restricted ourselves rather less massive $m_1 = 0.1$ because the setting of $m_1 = 0.1$ provides too strong perturbation to adequately describe the motion with our simple perturbation model. Though under this lowered mass the system disruption does not occur, an elevation of e_1 is expected, which is a necessary step to the system disintegration.

First we confirm that the transition to the circulation is captured via mapping. In Fig. 2, a quasi-periodic sustains for t < 94,000, then range of e_1 gradually increases to reach 0.8 at maximum. There is a transition from libration around $\omega_1 = -90°$, when e_1 tends to raise up. We use particle filtering to learn the time course of (e_1, ω_1) provided by the solution of Eq. (1). The trajectories of respective particles forming filtered distribution are shown in red, followed by trajectories propagated from the last filtered state (t = 93,500) without stochastic jumps (i.e. just solving planetary equations), shown in blue. Time t = 93,500 is close to the transition to circulation. Some propagated trajectories keep in the libration, while others transit to circulation. Hence, we can say that our perturbation model with stochastic jumps can capture the transition between the two states. The ratio of the numbers of orbits in the two mode may be interpreted an probabilistic evaluation of which state the orbit is in, when a large number of orbits need to be classified in a systematic way.

Continuing the filtering and the prediction after the circulation, we see that the perturbation model catches up the increase of e_1 and shorted period of the Kozai oscillation. However, the maximum value of e_1 given by the model is that out full simulation outcome, which reaches to 0.8. Higher order terms neglected here may be necessary to improve the agreement.

References

1. R. S. Harrington, "Dynamical Evolution of Triple Stars", Astronomical Journal, Vol. 73, No.3, 1968.
2. N. Georgakarakos, "Stability Criteria for Hierarchical Triple System", Celestial Mechanics and Dynamical Astronomy, Vol. 100, pp. 151-168, March, 2008
3. M. M. Saito, K. Tanikawa, "Disintegration Process of Hierarchical Triple Systems II: Non−small Mass Third Body Orbiting Equal−mass Binary", Celestial Mechanics and Dynamical Astronomy, Vol. 116, pp. 1-10, March 2013.
4. H. Kinoshita, H. Nakai, "General Solution of the Kozai Mechanism", Celestial Mechanics and Dynamical Astronomy, Vol. 98, pp. 67-74, May 2007.
5. R. Broucke, G. Smith, "Expansion of The Planetary Disturbing Function", Celestial Mechanics, Vol. 4, pp. 490-499, December, 1971.
6. G. Kitagawa, "Monte carlo filter and smoother for non-gaussian nonlinear state space models", J. Comput. Gr. Statistics, Vol. 5, pp. 1-25, 1996.
7. T. Higuchi et al., "Data Douka Nyumon [Translated from Japanese: An Introduction to Data Assimilation]", Asakurashoten, pp. 101-115, 2011. (In Japanese)

Authors Introduction

Dr. Masaya M. Saito

He received his PhD degree in science from the graduate university "SOKENDAI" in 2005 for his research on the rectilinear three-body problem in Celestial Mechanics.

Dr. Kiyotaka Tanikawa

He received his PhD degree in Science in 1983 from Department of Astronomy, The University of Tokyo in Japan.

Measuring the entire degree centrality in Yokokai

Tsutomu Ito, Seigo Matsuno
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

Mei Ootani
System Development Department Manufacturing Technology Office, NOK Corporation,
1-12-15 Shibadaimon, Minato-ku, Tokyo, 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

Centrality is one of the most important indexes in network calculation. Based on the definition, more than 400 different centrality such as degree, betweenness have been developed. All centrality indexes are calculated using the number of connection line, and its position in a given network. In automotive industry, keiretsu is considered as one of typical networks. It is crucial to measure the centrality of transaction network in the keiretsu. It is widely well-known that different parts play different roles in assembly line. Thus, the importance of each connect line in a transaction network should be measured based upon the importance of the parts. A new parts-importance weighted centrality model is proposed in this paper.

Keywords: Keiretsu loosening, Degree centrality, Entire degree centrality, Diameter, Production cost rate.

1. Introduction

It is widely well known that keiretsu is considered as one of the successful factors in automotive industry. After bubble economy collapsed in the beginning of 90s, Toyota and Mazda continue to maintain its keiretsu network with their parts suppliers while most of the keiretsu in car makers such as Nissan and Honda have been dismantled. In a given network, all companies develop their business activities as part of corporate strategy. Thus, to calculate their position in a given network, and evaluate the structure of a whole network is becoming crucial issue for discovering rational interfirm relationship and forming effective business strategy. Based on technical connections and development of car assembly system, a new method to calculate the centrality, one of the indexes of network structure, called parts-importance weighted centrality model is proposed in this paper.

This paper is structured as follows: In Section 2, a plethora of typical literature of network research are reviewed. Section 3 introduced models and measurement including network diameter, degree centrality, entire centrality of degree, and a new method called parts-importance weighted centrality model. Section 4 show the results and discussed the managerial implication. The conclusions including limitations and future directions of this research are proffered in Section 5.

2. Background

Network analysis is a research method that expresses various objects as a network consisting of points and lines and explores their structural features. Network analysis has been used in the humanities and social sciences such as sociology, anthropology, and psychology that deal with human relationships and relationships between groups, as well as mathematics called graph theory and engineering fields such as information science and operations research [1]. There are social networks that calculate the centrality of the network to identify influential people, and interfirm networks that classify industries from business-to-business transaction data. Based on the definition, more than 400 different centrality such as degree, betweenness have been developed. All centrality indexes are calculated using the number of connection line, and its position in a given network.

Basically, most of the parts for assembly line is purchased from its part suppliers in Keiretsu while other parts are traded from market. The former is called keiretsu trading, and the latter is called market trading. A phenomenon called "the keiretsu loosening" had occurred when the percentage of market trading increased [2]. A great change happened due to technical connections, and the ties between parts suppliers and car maker [3].

It is crucial to measure the centrality of transaction network in the keiretsu. It is widely well-known that different parts play different roles in assembly line. Thus, the importance of each connect line in a transaction network should be measured based upon the importance of the parts. In this study, the centrality of Mazda's transaction relationships is calculated and evaluated based on graph theory. Furthermore, an original evaluation model has been developed using parts-importance weights and the effectiveness of the model is examined.

3. Models and Measurement

3.1. *Network Diameter*

The maximum distance from one vertex to another is called eccentricity. The maximum number of eccentricities of the vertices included in the graph is called the diameter of the network [4].

3.2. Degree Centrality

The value of degree centrality is high if the number of edges connected to other nodes is more. It means that the degree centrality of each node is highly valued for more relationships in a network. The degree centrality of each node is formulated by Nieminen's as below [5].

$$C_D(p_k) = \sum_{i=1}^{n} a(p_i - p_k). (1)$$

where

$a(p_i, p_k)$ = w if and only if p_i and p_k are connected by a line with weight w
= 0 otherwise;

$C_D(p_k)$ ……the degree centrality of the vertex p_k;

Table 1. Parts Classification and Weights.

Parts Classification	Edge Weight	Example
Engine (block parts, valve system, fuel system)	10	Engine block and crankshaft, etc.
Engine auxiliary equipment (intake, exhaust, lubrication, cooling)	9	Oil pump, etc.
Engine electrical components	8	Alternator, and spark plug, etc.
Powertrain related parts	7	Torque converter, and transmission parts
Brake related parts	6	Brake system
Suspension and steering related parts	5	Power steering system
Body exterior	4	Wipers, bumpers, and door mirrors
Electrical parts (including air conditioner)	3	Airbags, door locks, relays, and sensors
Interior parts	2	Upholstery, dash insulator, and handle
others	1	Oil seals, air pipes

n ……the number of vertices included in a network.

3.3. Entire Centrality of Degree

Entire centrality refers to the centrality index of the entire network and was proposed by Freeman [6].

The entire centrality of degree is a centrality index for calculating the degree centrality of the whole network based on the degree centrality of each node in the network obtained by equation (2). The entire centrality of degree is formulated as follows [6].

$$C_D = \frac{\sum_{i=1}^{n}[C_D(p^*)-C_D(p_i)]}{n^2-3n+2}. \qquad (2)$$

In equation (2), $C_D(p_i)$ indicates the degree centrality of each node existing in the network, and $C_D(p^*)$ indicates that the degree centrality of each node in the network is the maximum. In addition, n is the number of nodes.

3.4. Parts-importance weighted centrality model

In consideration of importance of parts based technological view, the evaluation standard of parts is designed, and ten categories of parts with different importance are divided based on interview results of experts, and a plethora of literature [7-12]. Parts classification and their weight are shown as in Table1.
Engine parts are of the utmost importance because they are the power source for automobiles, also known as the "heart of automobiles." The powertrain-related parts are less important parts compared with that of the engine parts because they transmit the driving force of the engine to the tires and determine the specification of the vehicles. Moreover, the car cannot turn or stop without these parts such as suspension, and brakes. Thus, these parts were evaluated and determined based on their technical importance. Furthermore, the engine body requires technology for casting and cutting, but the electrical components are the control devices for the fuel supply system, the ignition system for gasoline engines, and wiring technology such as wiring and circuits. The quality of the technical capabilities required is different. In addition, when the engine is considered as an internal combustion engine, it assists in the production of energy-producing parts such as the engine that produces power, the oil pump for supplying lubricating oil, and the water pump that cools the engine. Therefore, the engine-related parts are divided into three stages including engine body, engine accessories, and engine electrical components. Finally, based on technical importance, the weight of 10 different categories is proposed in this paper.

4. Analysis and Discussion

4.1. Entire degree centrality and network diameter

Using the free open-source software "Gephi", the visualization of the diameter, and the degree centrality of each node in Yokokai is illustrated as Figure1.

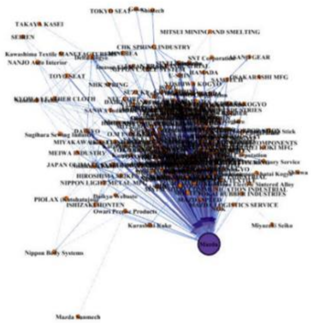


Figure 1. Entire degree centrality and network diameter.

Using an algorithm called ForceAtras2 in "Gephi", the network is visualized by reflecting the height of degree centrality and the weight of edge. In Figure 1, the higher the degree centrality value of each node, the darker the blue color, and the larger the size of the node circle is displayed. Even in the color of the line connecting each node, the heavier the weight of the edge, the darker the blue holds.

Figure 1 shows that the parts supply network of Mazda's cooperative organization "Yokokai" was affected by the loosening of the Keiretsu, the scale of the network expanded, and the overall degree centrality decreased holds. Thus, loosening of Keiretsu has led to the multi-centralization of technology networks, and the influence of the technology sharing among companies and the modularization of parts can be considered.

4.2. Mazda and entire degree centrality

To observe how the multi-centralization of the network due to the decrease in overall degree centrality affected Mazda, we compared Mazda's sales performance with overall degree centrality.

Since the affiliated transactions are premised on long-term transactions between car makers and parts suppliers, the parts suppliers were not exposed to competition. However, it is speculated that market principles have been introduced into the trading of parts and price competition among parts suppliers has started with Keiretsu loosening, which may have strong impacts on Mazda's cost reduction. Based on the speculation, we compared the changes in the overall degree centrality and Mazda's cost rate. The result is shown in Fig. 2.

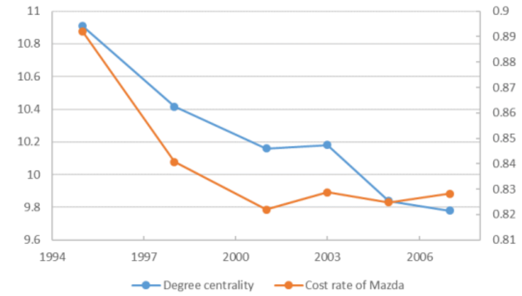


Figure 2. Time series changes in Mazda Group's entire centrality and Mazda's cost rate.

From Figure 2, it can be seen that Mazda's stand-alone cost ratio has declined as the overall degree centrality declines over the years. It was derived that the size of the parts supply network became wider and more centralized as the loosening of the affiliates progressed, thus the market trading of parts increased, the size of Yokokai expanded holds. This hypothesis is considered as the cause that make Mazda's cost rate down.

5. Conclusion

In this paper, a new parts-importance weighted centrality model is proposed. Using this model, the relations among scale of market trading, diameter of the network Yokokai, and the production cost of the core firm in Yokokai has been examined. Obviously, not only Mazda group, but also Toyota and other cars makers should be tested using this new model. In addition, only 6 fiscal years' data drawn from Yokokai is not enough. Much more data set should be gathered for our conclusion.

References

1. Tsutomu Suzuki, Network Analysis -Learning Data Science with R-, Second Edition (Japanese edition), Kyoritsu Shuppan Co., Ltd., 2017, Retrieved December 15, 2021.
2. Mori Hironori, Relations between companies in the Japanese automotive industry: evolving from maintaining domestic competitiveness to enhancing international competitiveness, Bulletin of Advanced Research Institute for the Science and Technology, Nihon University, 277-288, February 2007, Retrieved December 15, 2021.
3. Gouko Hiromichi, Transaction structure and changes of finished vehicle manufacturers and primary suppliers in the Japanese automobile industry, RIETI Discussion Paper Series 15-J-014, 2015, Retrieved December 15, 2021.
4. Stanley Wasserman, Katherine Faust, Social Network Analysis: Methods and Applications (Structural Analysis in the Social Sciences, Cambridge University Press, 1994, Retrieved December 15, 2021.
5. Nieminen J., On centrality in a graph, Scandinavian Journal of Psychology 15, 322-336, 1974, Retrieved December 15, 2021.
6. Linton C. Freema, Centrality in Social Networks Conceptual Clarification, Social Networks 1, 215-239, 1978/79, Retrieved December 15, 2021.
7. A total guidance of Mazda Group (Japanese edition), IRC, 1995, Retrieved December 15, 2021.
8. A total guidance of Mazda Group (Japanese edition), IRC, 1998, Retrieved December 15, 2021.
9. A total guidance of Mazda Group (Japanese edition), IRC, 2001, Retrieved December 15, 2021.
10. A total guidance of Mazda Group (Japanese edition), IRC, 2003, Retrieved December 15, 2021.
11. A total guidance of Mazda Group (Japanese edition), IRC, 2005, Retrieved December 15, 2021.
12. A total guidance of Mazda Group (Japanese edition), IRC, 2007, Retrieved December 15, 2021.

==

Authors Introduction

Dr. Takao Ito

Dr. Takao Ito is Professor of Management of Technology (MOT) in Graduate School of Engineering at Hiroshima University. He is serving concurrently as Professor of Harbin Institute of Technology (Weihai) China. He has published numerous papers in refereed journals and proceedings, particularly in the area of management science, and computer science. He has published more than eight academic books including a book on Network Organizations and Information (Japanese Edition). His current research interests include automata theory, artificial intelligence, systems control, quantitative analysis of inter-firm relationships using graph theory, and engineering approach of organizational structures using complex systems theory.

Dr. Tsutomu Ito

Dr. Tsutomu Ito is Assistant Professor of the Department of Business Administration at National Institute of Technology, Ube College, Japan. 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.

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. Matsuno has published several research papers and books in that area. He is a member of the Board of the Directors of Japan Society for Information and 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.

Ms. Mei Ootan

Ms. Mei Ootani is professional engineer working at NOK Corporation, Japan. She received her Foundation degree from the Department of Business Administration, Ube National College of Technology, Japan in 2021.

A comparative study on Michinoeki's efficiency in Japan

Tsutomu Ito, Seigo Matsuno
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

Michinoeki is considered as one of the most successful social experiments in Japan now. A plethora of literature of Michinoeki's have been published over past decades. The authors reviewed many important and typical literatures and found that the quantitative studies are still sparse. Obviously, it is a critical issue to measure the efficiency of Michinoeki's for its further development and revitalization of the local economy. Thus, the data of Michinoeki are gathered from all Michinoeki in Japan, and the relative efficiency are calculated using DEA model for comparison. Furthermore, the managerial implication of the results is discussed in this paper.

Keywords: Michinoeki, efficiency, DEA, revitalization, PRC.

1. Introduction

One of the unique systems in Japan called Michinoeki started from 1993 with 103 stations now reached 1,193 stations. Michinoeki are well known with its four functions of (1) providing free parking space, restrooms, (2) spreading information, (3) allying with regional society, and (4) preventing disaster [1]. Because of low birthrate and aging society, Japanese economy is facing a serious issue of deflation and economic recession. Thus, Michinoeki is considered as one of the most effective tools for economic recovery, and revitalization of local economy. To our best knowledge, the research papers of Michinoeki using scientific approach are still sparse. This paper attempts to shed light on calculating efficiency and comparing the performance of Michinoeki using DEA model.

This paper is structured as follows: Section 2 introduces the background of this research. In Section 3, the paper explicates data collection and the result using the DEA model. Section 4 shows the analysis and discussions on our findings. The conclusions and managerial implications are proffered in the final section.

2. Background

Most of the Michinoeki's research are based on four functions mentioned above. For instance, Yoshida et al.

studied the roles of Michinoeki when natural disaster happened and indicated the detailed reaction if the residents use the Michinoeki nearby their houses [2]. However, much more research focused on economic performance as the important index for their sustainability. Ozuka et al. conducted a questionaries survey and clarified some basic issues of revitalization of local economy [3]. Moreover, Michinoeki are studied from different perspectives such as regional economics and local tourist policies [4-6]. Recently, Ito et al. developed some mathematical models to calculate Michinoeki's efficiency in Yamaguchi prefecture and clarified the characteristics of different areas' Michinoeki [7]. This paper will calculate the efficiency of Michinoeki in Japan and compare the performance of different areas.

3. Data collection and Measurement

3.1. *Data Collection*

Data including basic information, management style, scale of free parking space, spreading information system, facilities for preventing disaster, population of local areas, and retail shop of local products from 2016 to 2017 were drawn from the internal databases of Michinoeki headquarters. Enot only many missing data, but also logical errors, omission, and errors in writing have been founded. All errors are corrected based on telephone survey and homepage checking. A selected part of collection is shown as in Table 1.

Before efficiency calculation, the variables should be determined. According to the advice from Michinoeki's experts, two variables of passengers of register count (PRC) and sales revenue of register count (SRRC) as description variables, and 20 variables such as square meters of land space (SMLS) and number of the registered farmers (NRF) as the explanatory variables of multiple regression model have been selected. Thus, this leads to the formation of the following regression equation:

$$y = a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + a_6x_6 + a_7x_7 + a_8x_8 + a_9x_9 + a_{10}x_{10} + a_{11}x_{11} + a_{12}x_{12} + a_{13}x_{13} + a_{14}x_{14} + a_{15}x_{15} + a_{16}x_{16} + a_{17}x_{17} + a_{18}x_{18} + a_{19}x_{19} + a_{20}x_{20} + c. \quad (1)$$

where

y … … Passengers of Register Count (PRC)and Sales Revenue of Register Count (SRRC) in H26 or H27;
x_1 … … Square Meters of Land Space (SMLS);
x_2 … … Number of Standard- Sized Car Parking Lot (NSSCPL);
x_3 … … Number of Large Vehicles Parking Lot (NLVPL);
x_4 … … Square Meters of Parking Area (SMPA);
x_5 … … Total Number of the Restroom (TNR);
x_6 … … Square Meters of Free Rest Place (SMFRP);
x_7 … … Total Seats of the Free Rest Place (TSFRP);
x_8 … … Weekdays' Traffic Near the Station (WTNS);
x_9 … … Holidays' Traffic Near the Station (HTNS);
x_{10} … … Population of the City Located (PCML);
x_{11} … … Operating Cost (OC);
x_{12} … … Total Customers (TC);
x_{13} … … Number of the Agriculture Products and Sales (NAPS);
x_{14} … … Number of the Agriculture and Marine Products (NAP);
x_{15} … … Number of the Local Products (NLP);
x_{16} … … Number of the Selling Items (NSI);
x_{17} … … Number of Original Products (NOP);
x_{18} … … Number of the Registered Farmers (NRF);
x_{19} … … Square Meters of the Facilities for Marine Products (SMFMP);
x_{20} … … Square Meters of Facilities Space (SMFS);
c … … constant.

Table 1. Selected part of the corrections.

ID	Name	Item	Correction
186	Jobonnosato	Sales revenue of register count (H27)	Replace "more than 1000" to "1000".
200	Kisakata	Sales revenue of register count (H26)	Replace "534713390" to "534.713390".
337	Agatumakyo	Passengers of register count (H27)	Replace "125093(from October to March)" to "250,186".
590	Utsunoyatouge	Passengers of Register Count (H27)	Replace "66056 (Shizuoka side), 71565 (Fujieda side)" to "137,621".
		Sales Revenue of Register Count (H27)	Replace "55 (Shizuoka side), 38(Fujieda side)" to "93".
		Total Customers	Replace "165140 (Shizuoka side), 178912 (Fujieda side)" to "344,052".

As result, 8 variables of SMLS, PCML, OC, TC, NLP, NSI, NRF, SMFMP are selected as key factors which have strong impact on PRA and SRRA based on trials and errors.

3.2. *Measurement*

Basically, efficiency is calculated as a ratio of output and input. For comparing Michinoeki' efficiency in different areas, a relative efficiency should be calculated. DEA (Data Envelopment Analysis) model is one of the effective tools to calculate relative efficiency. The DEA model (CCR) is defined by the following objective function to find Michinoeki's efficiency as:

Objective function

$$max \quad D_j = \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} = \frac{u_1 y_{1j} + u_2 y_{2j} + \cdots + u_s y_{sj}}{v_1 x_{1j} + v_2 x_{2j} + \cdots + v_m x_{mj}}. \qquad (2)$$

where

the Michinoeki's known r outputs $y_{r1},\ldots\ldots y_{rs}$ are multiplied by their respective weights $u_1,\ldots\ldots u_s$ and divided by the i inputs $x_{i1},\ldots\ldots x_{im}$ multiplied by their respective weights $v_1,\ldots\ldots v_m$.

Subject to

$$\frac{\sum_{r=1}^{s} u_r y_{rj}}{v_i x_{ij}} \leq 1 \quad (j = 1,2, \ldots, n). \qquad (3)$$

$$u_r \geq 0 \quad (r = 1,2, \ldots, s). \qquad (4)$$

$$v_i \geq 0 \quad (i = 1,,2, \ldots, m). \qquad (5)$$

Using 3 variables selected from above 8 variables: SMLS, NRF, SMFMP, as input, and PRA as output of equation (2), efficiency of 554 stations in H26, and 561 stations in H27 from total 1,107 stations are calculated because of lack of data. The result of 561 Michinoeki' efficiency of 10 areas in H27 is illustrated in Figure 1.

4. Analysis and Discussion

Figure 1 shows that the efficiency of Chugoku, Chubu, Hokkaido, Tohoku, and Okinawa is better compared with Hokuriku, Kinki, Shikoku, and Kyushu. For further detailed analysis, mean and coefficient of variation of different areas' efficiency are required. They are shown as in Figure 2 and 3 respectively.

Generally, higher efficiency and lower coefficient of variation, higher performance holds. The efficiency of Okinawa in H27 is higher than other areas, but the coefficient of variation of Okinawa is higher than other areas. Thus, the performance of Okinawa could be considered as one of the best among those 10 areas. Moreover, to clarify the transition of Michinoeki', rank of efficiency and coefficient of variation is illustrated in Figure 4.

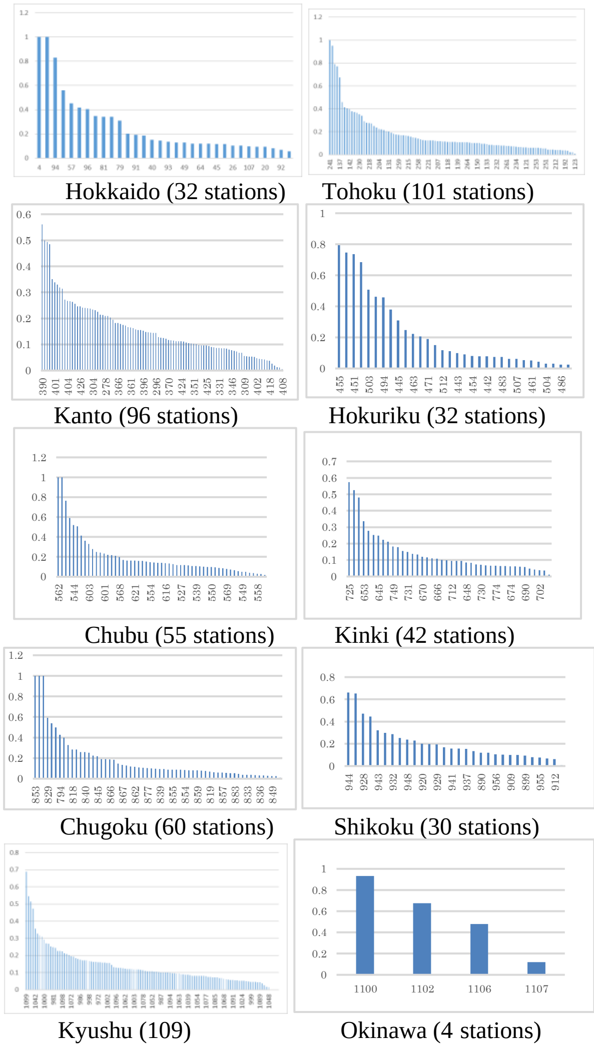


Figure 1. Michinoeki' Efficiency of 10 Areas (H27, 561 stations).

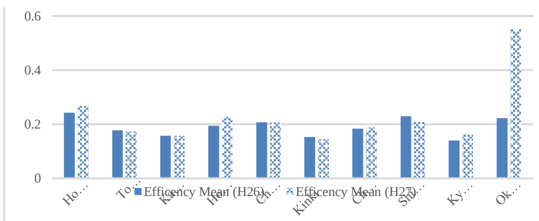


Figure 2. Mean of Michinoeki' Efficiency of 10 Areas.

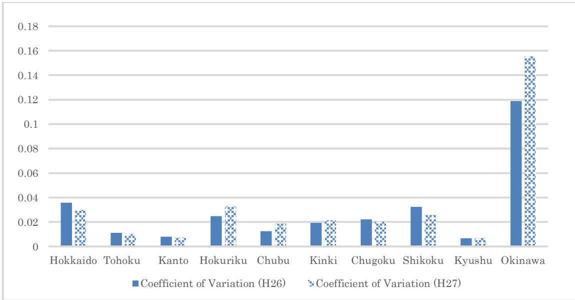


Figure 3. Coefficient of Variation of Michinoeki' of 10 Areas.

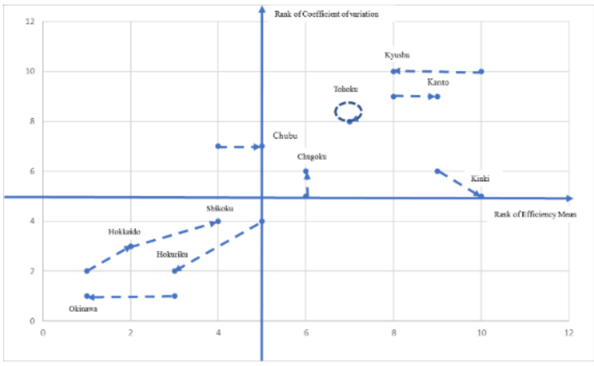


Figure 4. Transition of Michinoeki' Rank of 10 Areas in H26 and H27.

Figure 4 shows that the best Michinoeki is Chubu because of high efficiency and low coefficient of variation. Michinoeki of Shikoku, Hokkaido, Hokuriku, and Okinawa are high efficiency with low coefficient of variation. The rank of Okinawa and Hokuriku in H27 is higher than that in H27, thus Michinoeki in these areas are unstable with high performance. The rank of Michinoeki of Kyushu, Kanto, Chugoku, Tohoku, and Kinki are stable with low efficiency, thus there is room for further improvement for these areas.

5. Conclusion

The efficiency of All Michinoeki in H26 and H27 have been calculated in this paper. Tohoku and Chubu are considered as the Michinoeki with best performance. However, it is not enough to analyze transition of Michinoeki only using two year's data. In addition, more variables such as NLP and NSI should be tested as determinants of input for efficiency calculation in future.

Acknowledgements

This study was supported partially by JICHIRO grant.

References

1. Tsutomu Ito, Minoru Kumano, Takao Ito, Measuring Efficiencies of Michinoeki using DEA Model in Yamaguchi Area, Journal of Robotics, Networking and Artificial Life, Volume 7, Issue 3, 204–207, December 2020, Retrieved December 15, 2021.
2. Satoshi Yoshida, Yasuaki Matsuda, Naoto Takada, A consideration on functions and roles of Michinoeki' for preventing disaster (Japanese edition), Retrieved November 30, 2021.
3. Misuzu Ozuka, A study on finding the relationship between Managerial issues and effects as expected of Michinoeki -resulted from questionaries survey of the Michinoeki in Kinki area- (Japanese edition), Retrieved November 30, 2021.
4. Ryusaku Matsuo and Satoshi Yamaguchi, Michinoeki Economics -Economic revitalization and development of social society- (Japanese edition), Keiso Shobo, August 2019, Retrieved December 15, 2021.
5. Kyosei Sakamoto, kaori Toda, tourist information spreading and revitalization of regional economy from Michinoeki, Retrieved November 30, 2021.
6. T. Hiraoka, H. Nonaka, M. Kumano, S. Nishimura, T. Ito, A Correlation Analysis between Information Related to the Number of Customers and the Agriculture, Forestry, and Fisheries of Michi-no-eki in Japan, Journal of Institute of Industrial Applications Engineers, Vol. 6, No. 4, pp. 173-178, October 25, 2018, Retrieved December 15, 2021.
7. Takao Ito, Tsutomu Ito, Seigo Matsuno, Rajiv Mehta, Makoto Sakamoto, Identifying the Determinants of Michinoeki Performance, Journal of Robotics, Networking and Artificial Life, Vol. 8(3); 165–169, December 2021, Retrieved December 15, 2021.

===

Authors Introduction

Dr. Takao Ito

Dr. Takao Ito is Professor of Management of Technology (MOT) in Graduate School of Engineering at Hiroshima University. He is serving concurrently as Professor of Harbin Institute of Technology (Weihai) China. He has published numerous papers in refereed journals and proceedings, particularly in the area of management science, and computer science. He has published more than eight academic books including a book on Network Organizations and Information (Japanese Edition). His current research interests include automata theory, artificial intelligence, systems control, quantitative analysis of inter-firm relationships using graph theory, and engineering approach of organizational structures using complex systems theory.

Dr. Tsutomu Ito

Dr. Tsutomu Ito is Assistant Professor of the Department of Business Administration at National Institute of Technology, Ube College, Japan. 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.

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. Matsuno has published several research papers and books in that area. He is a member of the Board of the Directors of Japan Society for Information and 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.

The R&D Direction and Business Strategy: The case study on the cooperation of EV and battery makers

Yousin Park
Faculty of Regional Development, Prefectural University of Hiroshima
Hiroshima City, Hiroshima Prefecture, 732-0821, JAPAN

Iori Nakaoka
Faculty of Business Administration, Seijoh University
Tokai City, Prefecture, 476-8588, JAPAN

Yunju Chen
Faculty of Economics, Shiga University
1-1-1 Banba Hikone City, Shiga Prefecture, 522-8522, JAPAN

E-mail: ecventure@pu-hiroshima.ac.jp, nakaoka-i@seijoh-u.ac.jp
yun-chen@biwako.shiga-u.ac.jp

Abstract

This paper focuses on the R&D direction and the business strategy of EV firms and battery makers with reference to Porter's productive frontier. M. E. Porter (1996) claimed that the productivity frontier represents the maximum value that the organization can deliver at any a given cost, using technologies, skills and purchased inputs. He argued that strategic decisions are ones that are aimed at differentiating an organization from its competitors in a sustainable way in the future. We use the patent information of EV firms (Toyota, Tesla, Volkswagen) and battery makers (Panasonic, CATL, LG Chem) as the cases. We examine our propositions by social network analysis and text mining. The analysis in this paper includes: 1) trying to distinguish between differentiation and cost leadership strategy from R&D direction, and visualizing productivity frontier, 2) making discussing on the inter-organizational relation of EV firms and battery makers. In this paper, we clarify that patterns of cooperation EV firms and battery makers R&D.

Keywords: R&D direction, EV makers, Battery makers, Patent analysis

1. Introduction

EV (Electric Vehicle) is expected to spread as it has the potential to bring about major changes in global energy problems such as global warming countermeasures. The research and development of battery, which is the main component of EV, will be important for the spread of EV.

This paper will focus on the R&D direction and the business strategy of EV firms and battery makers. We use Porter's productivity frontier to discuss these issues. The analysis in this paper includes: 1) visualizing the productivity frontier of battery industry, and distinguishing between differentiation and cost leadership strategy from R&D direction 2) discussing on the inter-organizational relationships of EV firms and battery makers. In this paper, we clarify the patterns of strategic alliance between EV and battery makers.

2. Background

2.1. *Productivity frontier*

M. E. Porter defines the productivity frontier as: the sum of all existing best practices at a given time. And he explains the difference between operational effectiveness and strategic positioning with productivity frontier map [1]. Operational effectiveness (OE) means performing similar activities better than rivals performing them. In contrast, strategic positioning means performing different activities from rivals' or performing similar activities in different ways. (See Figure 1) He points out that constant improvement in operational effectiveness is necessary to achieve superior profitability. However, it is not usually sufficient. Competitive strategy is about being different. It means deliberately choosing a different set of activities such as non-price buyer value delivered or relative cost position to deliver a unique mix of value.

In the case of EV's battery, we focus on the R&D direction of battery makers that choosing from LFP (LiFePo4), NCM (Nickel, Cobalt, Manganese), and Solid-state batteries. First, characteristics of LFP are lower energy density than NMC, quite robust rather than economical (abundant materials, no need of Nickel and particularly Cobalt).Second, characteristics of NMC are higher energy density, shorter lifetime, lower safety margins, and higher price. Finally, solid state batteries where liquid electrolyte and separator are replaced by a solid material, have to solve many issues and the R&D are still on progress and mass production have not started yet.

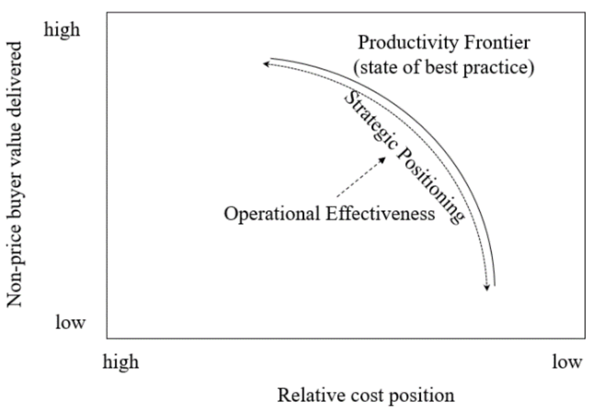


Fig. 1. Operational Effectiveness Versus Strategic Positioning modified by the authors

Table 1. World Plugin Vehicle Sales (2020)

Model	Brands	Battery Maker	2020H1Sales	2021H1Sales	Y-O-Y
Tesla Model 3	Tesla	CATL, LG, Panasonic	142,346	243,753	71.20%
Wuling HongGuang Mini EV	SAIC	CATL, Gotion High-tech	- - -	181,810	- - -
Tesla Model Y	Tesla	LG, Panasonic	13,415	138,401	931.70%
BYD Han EV	BYD	BYD	- - -	38,667	- - -
Volkswagen ID.4	Volkswagen	CATL, LG, Samsung SDI, Gotion High-tech	- - -	38,499	- - -
GW ORA Black Cat	GWM	SVOLT,CATL	- - -	32,013	- - -
Renault Zoe	Renault	LG,AESC	37,154	31,426	-15.40%
Hyundai Kona EV	Hyundai	SK Innovation	19,286	31,233	61.90%
Volkswagen ID.3	Volkswagen	CATL, LG, Samsung SDI, Gotion High-tech	- - -	31,079	- - -
GAC Aion S	GAC	CALB, CATL	14,516	30,456	109.80%
Li Xiang One EREV	Li Auto	CATL	- - -	30,154	- - -
Nissan Leaf	Nissan	AESC	23,867	29,372	23.10%
Changan Benni EV		Gotion High-tech, CATL, CALB, BYD	- - -	29,178	- - -
Kia Niro EV	Kia	SK Innovation	12,157	27,395	125.30%
Chery eQ	Chery Auto	CATL, Gotion High-tech, Farasis Energy	- - -	27,136	- - -
Volvo XC40 PHEV	Volvo Cars	CATL, LG	- - -	26,839	- - -
Audi e-tron	Audi	LG, BYD	17,592	25,758	46.40%
Toyota RAV4 PHEV	Toyota	Panasonic, CATL, BYD	- - -	25,279	- - -
BMW 530e/Le	BMW	CATL, Samsung SDI	20,586	24,985	21.40%
Ford Escape/Kuga PHEV	Ford	Samsung SDI,BYD,SK Innovation	- - -	24,763	- - -

Source: Inside of EVs Website

Table 2. Global EV Battery Deployment

Source: SNE Research

2.2. *The overview of EV and battery industry*

After a decade of rapid growth, in 2020 the global electric car stock hit the 10 million mark, a 43% increase over 2019, and representing a 1% stock share. Battery electric vehicles (BEVs) accounted for two-thirds of new electric car registrations and two-thirds of the stock in 2020 [2].

As they are key players in this market, this paper will focus on the EV firms such as Tesla, Volkswagen and Toyota, and battery makers such as Panasonic, CATL and LG Chem. (See Table 1 and Table 2)

3. Methodology and data

In the following sections, the R&D direction and business strategy of each EV firms and battery makers is analyzed with social network analysis, which can visualize the features of R&D patterns using archived patent information.

We selected patent documents archived in patent database service by Patent Integration Co.. All patents are classified according to the worldwide standard classification codes IPC (International Patent Classification). This paper utilizes all patents of EV firms and battery makers since 2000, including USA and WIPO (World Intellectual Property Organization). We extract related patents such as "Battery (H1), Charging (H2)", and then collect patents which are applied by each company. (See Table 3)

Table 3. IPC code on battery

H01M50/00	Constructional details or processes of manufacture of the non-active parts of electrochemical cells other than fuel cells, e.g. hybrid cells
H01M6/00	Primary cells; Manufacture thereof; In this group, primary cells are electrochemical generators in which the cell energy is present in chemical form and is not regenerated.
H01M8/00	Fuel cells; Manufacture thereof; In this group, the following expression is used with the meaning indicated: "Fuel cell" means an electrochemical generator wherein the reactants are supplied from outside.
H01M10/00	Secondary cells; Manufacture thereof; In this group, secondary cells are accumulators receiving and supplying electrical energy by means of reversible electrochemical reactions.
H01M12/00	Hybrid cells; Manufacture there of (hybrid capacitors H01G11/00); Note. This group does not cover hybrid cells comprising capacitor electrodes and battery electrodes, which are covered by group H01G11/00. In this group, hybrid cells are electrochemical generators having two different types of half-cells, the half-cell being an electrode-electrolyte combination of either a primary, a secondary or a fuel cell.
H02J7/00	Circuit arrangements for charging or depolarising batteries or for supplying loads from batteries

Source: Japan Patent Office

3.1. *An approach based on the number of patent publications*

In the case of EV firms, Toyota has more control over its rivals such as Tesla and Volkswagen, lead in patents, but its sales are unmatched by Tesla. In the case of Battery makers, Panasonic and LG Chem have more control over CATL, lead in Patent, but the sale of CATL is the highest in this industry. (See Table 2, Fig. 2 and Fig. 3)

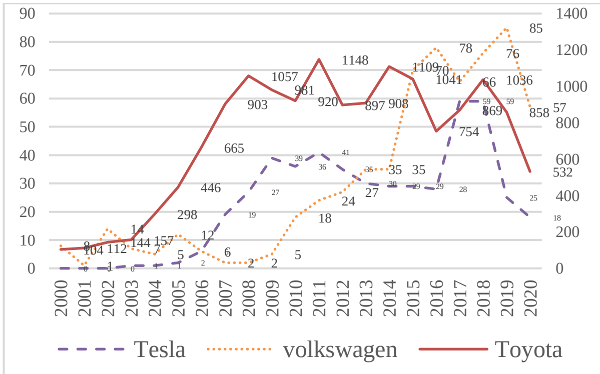


Fig. 2. Status of battery patents (by EV firms)

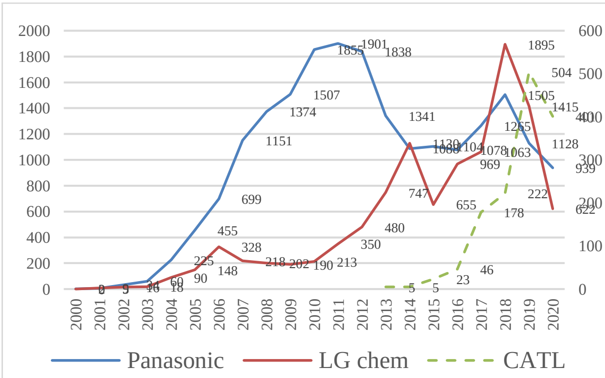


Fig. 3. Status of battery patents (by Battery Makers)

3.2. *An approach by social network analysis*

Fig. 4 shows the R&D direction and business strategy of each EV firms and battery makers with social network analysis with transaction and patent data. (See Table 2, Fig. 2 and Fig. 3) Toyota, Panasonic and LG Chem have more power than Volkswagen, Tesla and CATL in patent, but Tesla and CATL have better position than their rivals

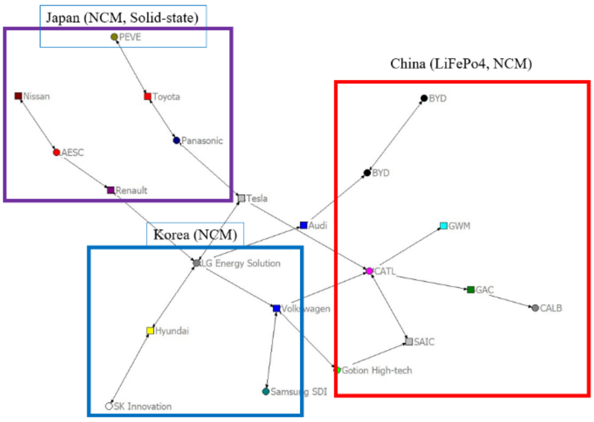


Figure 3. The relation between EV firms and battery makers by social network analysis

at the business level. Because they are center position in value network.

4. Discussion and conclusions

We have discussed the patterns of cooperative relationship between EV firms and battery makers under the impact of radical technological changes [3]. In our analysis, Tesla has good position which can use both low cost (LFP) and difference (NCM) in productive frontier. In Japanese firm's case, Panasonic have continued to grow in importance. Because they have cooperative relationship with Tesla and Toyota. Moreover, we find the similarities and differences in the R&D direction from EV firms and battery makers. In the future study, we will discuss how these companies have taken different R&D strategies?

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number JP 20K01941.

References

1. M. E. Porter, *On Competition, Updated and Expanded Edition*, Harvard Business Review Press (2008)
2. IEA Global EV outlook 2021: Trends and developments in electric vehicle markets (2021).
3. Y. Park and I. Nakaoka, Y. Chen,, "The R&D Strategy of Automobile Companies in Radical Innovation" *Journal of Robotics, Networking and Artificial Life*, Vol.7, No.3, pp. 184-189 (2020)

Authors Introduction

Prof. Yousin Park

He received Ph.D. degree in Economics, from Graduate School of Economics, the Kyoto University, Japan in 2009. He is now a professor with in the Faculty of Regional Development, Prefectural University of Hiroshima, in Japan. His research focus on the Management of Technology, Strategic Management, also the industry of Information Communication Technology. .

Associate Prof. Iori Nakaoka

He received Ph.D. degree in Engineering, from Graduate School of Science and Engineering, Ritsumeikan University, Japan in 2006 and now he is working as an Associate Professor in the Faculty of Business Administration, Seijoh University in Japan. His research fields are management information system and computational social science.

Associate Prof. Yunju Chen

She received a Ph.D. degree in Economics, from Graduate School of Economics, Kyoto University. She is now an associate professor in the Faculty of Economics, Shiga University, Japan. Her research's focus on the inter-organizational strategies, and the management processes of technology development, also the international comparative studies in these fields.

Discovering the relationship between tourists and tourist spots in Japan

Tsutomu Ito, Seigo Matsuno
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

Abstract

It is widely well-known that Japan currently faces serious problems with the declining birthrate and aging society. As actions to counteract our aging and shrinking population, revitalization of local economy is becoming the most effective economic development strategy. Obviously, tourism is considered as one of the effective policies to develop local economy today. To provide more practical evidence of tourism development, this paper is focusing on discovering the relationship between tourists and tourist spots using conventional regression model and 4-cell model developed on tourists and tourist spots. Moreover, as one of the successful factors, the importance of the development of public traffic among different tourist spots has been confirmed.

Keywords: Regional revitalization, tourist spot, tourism resource, travel time, four-cell model.

1. Introduction

Currently, Japan is facing the problems of declining birthrate and aging population. As population decline, low birthrate, and aging of the labor force, the economic activities will be discontinued which will cause many difficult issues for national economy and society. Thus, it is an important issue to revitalize local communities today obviously. In recent years, tourism has become one of the important policies for revitalization of regional cities and makes contribution to economic growth strategies if tourism is actively promoted. It will be easy to accurately provide tourism resources for regional revitalization if the relationship between tourism and the number of tourists clarified. This paper makes contribution to develop a new 4-cell model to measure the efforts to increase the number of tourists. And the validity of 4-cell model has been tested. Furthermore, we confirmed that to develop public traffic among different tourist spots could be considered as one of the important success factors for tourists increasing.

This paper is structured as follows: Section 2 reviewed typical literature of this research. In Section 3, the paper explicates four-cell model and associated hypotheses. Section 4 introduced the determinants of successful tourism. Analysis and discussion of the relationship between the four-cell cell and the distance between tourist spots is carried out in Section 5. The conclusions and managerial implications are proffered in the final section.

2. Background

Obviously, tourism is considered as one of the effective policies to develop local economy today. A plethora of literature have been published over past decades. Kudo et al. once proposed a new method of event collection for tourist information distribution to foreign tourists using SNS data [1]. Aihara analyzed tourist's behavior and thoughts using big data [2]. Moreover, Okamura and Fukushige stressed the importance of repeater tourists, and conducted an empirical analysis on how to promote repeater tourists in Kansai area [3]. Butler presented a concept of a recognizable cycle in the evolution of tourist areas and explained the issue of continuing decline in the environment quality and the attractiveness of many tourist areas using a basic s curve [4]. Accordingly, as actions to counteract our aging and shrinking population, tourism is becoming the most effective economic development strategy, especially revitalization of regional areas today. However, the research of analyzing the determinants of tourism, and evaluating the results of tourists' promotion policy are still sparse. To provide more practical evidence of tourism development, this paper is focusing on discovering the relationship between tourists and tourist spots using conventional regression model and 4-cell model based on the relationship between tourists and tourism spots. Furthermore, as one of the successful factors, the importance of the development of public traffic among different tourist spots has been confirmed.

3. Four-cell model and 4 Hypotheses

It is crucial to observe and evaluate changes in the number of tourists and tourist spots in chronological order. Accordingly, the relationship between the change of the number of tourists and the number of tourist elements are calculated. A 4-cell model proposed in this paper is targeted at 13 prefectures in Japan, and data on customers increasing for 7 types of tourism elements is collected for 7 years, while the number of 7 types of tourism elements maintained in each of the 13 prefectures is used [5-6]. By using this, it becomes possible to collectively express and compare fluctuations in the number of tourists and the number of tourist facilities. An example is shown in Figure 1.

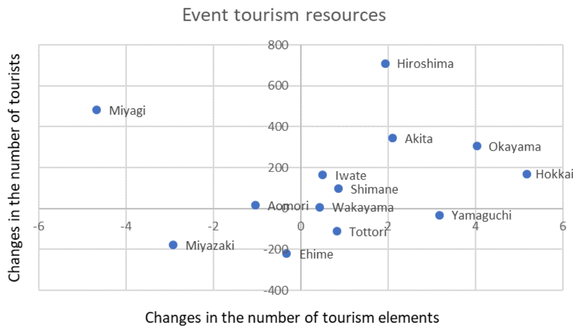


Figure 1. Calculation result of 4-cell model for event.

The horizontal axis and the vertical axis are designed for indicating the change of the number of tourist spots, and the increase or decrease of the number of tourists respectively.

Based on the 4-cell model, positioning is performed for each quadrant with building hypotheses for each quadrant and giving weighting factors, which is used as an index showing the tourism policy of each prefecture. The hypotheses are shown below.

H1. If it belongs to the first quadrant, it is successful as a tourism industry.

H2. If it belongs to the second quadrant, you can call it repeater.

H3. If it belongs to the third quadrant, it is declining industry.

H4. If it belongs to the fourth quadrant, it failed as a tourism industry.

Table 1. The planning and control components.

	Hokkaido	Aomori	Iwate	Miyagi	Akita	Wakayama	Tottori	Shimane	Okayama	Hiroshima	Yamaguchi	Ehime	Miyazaki
Nature	1	2	1	1	3	1	2	4	1	2	1	4	1
History and culture	1	2	4	2	2	1	1	1	2	1	1	4	1
Hot spring and healthcare	1	3	1	2	3	2	2	3	3	1	4	4	4
Sports	1	1	3	2	3	3	3	3	2	1	1	1	2
Urban tourism	1	1	4	1	1	1	3	3	4	2	1	4	1
Event	1	2	1	2	1	1	4	1	1	1	4	3	3
Others	1	1	1	2	1	1	1	1	1	1	1	1	2

Regarding Hypothesis 1, both the number of tourists and the number of tourist spots are increasing, and it is considered that the most desirable tourist attraction has been achieved, thus a weight of 4 is given to the first quadrant. For Hypothesis 2, the number of tourist spots is decreasing despite of the increasing in the number of tourists. It is suggested that from the references that the number of repeaters is involved in the number of tourists, and it is considered that the efforts of tourists who do not seek new tourist spots have been successful. Therefore, a weight of 3 is given to the second quadrant. Regarding Hypothesis 3, both number of tourist spots and the number of tourists is decreasing, and it is possible that they are entering a period of decline, thus a weight of 2 is given to the third quadrant. Regarding Hypothesis 4, the number of tourists is decreasing despite of the increase in the number of tourist spots, indicating that supply and demand do not match with each other. This gives a weight of 1 to the fourth quadrant because it is considered that the desired tourism attraction has not been achieved. The result is shown in Table1.

Figure 2 shows the evaluation index for calculating the tourism attraction status for each prefecture by summarizing the tourism elements belonging to each quadrant for each prefecture and considering the weight of each quadrant.

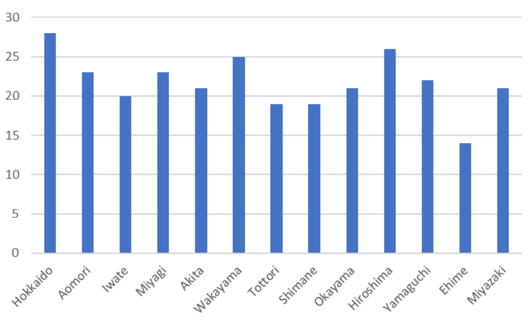


Figure 2. Evaluation index of the selected prefecture.

4. Determinants of successful tourism

We will examine the successful factors by representing the best two prefectures and the worst two prefectures that are considered to have succeeded in attracting tourism with the 4-cell model. To verify what factors are involved in attracting tourists, we focused on the time required to move between tourist spots. It is unlikely that tourists will visit their spots only for one factor. It is more likely that they will visit several tourist spots in the prefecture. Therefore, based upon the accessibility index proposed by the Ministry of Land, Infrastructure, Transport and Tourism [7-8], we calculated the evaluation travel time based on the travel time between tourist spots and compared it with the tourism trends of typical prefectures.

We selected famous tourist spots from Jaran Net [9] to check whether it is possible to access tourist spots from tourist spots, and when it is possible for travel time among tourist spots. If you use public transportation, use public transportation as a railroad. If not, the travel time as a car was measured respectively. The travel time between tourist spots is summarized in a matrix format.

The evaluation travel time is calculated by the following formula.

$$\text{Evaluation travel time} = \frac{\text{Travel time required}}{\text{Area of each prefecture}} \qquad (1)$$

Tables 2 and 3 show a list of evaluation travel times for railways and highways in the four prefectures, respectively.

Table 2. Railroad evaluation Travel time average (min/km^2).

Hokkaido	Hiroshima	Shimane	Ehime
0.0027	0.0104	0.0146	0.0268

Table 3. Highway evaluation Travel time average (min/ km^2).

Hokkaido	Hiroshima	Shimane	Ehime
0.0021	0.0074	0.0107	0.0104

5. Analysis of the relationship between the 4-cell model and the distance between tourist spots

Comparing these calculation results with the tourist attraction evaluation index evaluated by the 4-cell model, the travel time per unit area of kilometers tends to be short in Hokkaido and Hiroshima prefectures, and Shimane and Ehime are moving. It was confirmed that it is time consuming. On the other hand, it is confirmed that the travel time per unit area of Hokkaido and Hiroshima was short even when using a car, but in Shimane and Ehime, the result was reversed in the case of railroads. The results are shown in Figure 3.

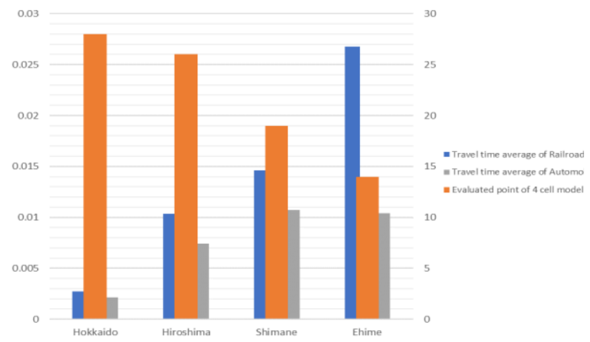


Figure 3. Relationship between evaluation value and distance between tourist spots.

In Figure 3, blue represents the value obtained by dividing the moving average time when using a railway and gray represents the value obtained by dividing the moving average time by the area. Orange represents the value of evaluation in the 4-cell model, and it is desirable that the value is high. From Fig. 3, there is an inversely proportional relationship between the evaluation value and the distance between tourist spots, and the correlation coefficient is -0.959, which is highly correlated with the number of tourists attracted between tourist spots. It was able to obtain a relationship with travel time.

6. Conclusion

This paper aimed at clarifying the factors of tourism and customers increasing to promote the accurate provision of tourism resources that will lead to regional revitalization in the age of declining population and aging population. We proposed a method to clarify the above purpose and verified its effectiveness. In the process of verification, it is confirmed that it is difficult to find out the factors which affect the number of tourists by a simple comparison between the number of tourists and the number of restaurants in the vicinity. Therefore, we proposed a new method to evaluate the tendency of attracting tourists by summarizing the factors of attracting tourists in Japan with the 4-cell model. Furthermore, we investigated the development status of the transportation network of the 4-cell model and the local indication itself and confirmed that there was a high correlation between the development status of public transportation and the index obtained by the proposed method. This result shows that the development of public transportation that connects tourist spots is one of the factors for the success of developing local tourism, and it is considered as development direction for future tourism policy planning.

References

[1] Ruriko Kudo, Miki Enoki, Akihiro Nakao, Shu Yamamoto, Saneatsu Yamaguchi, and Masato Oguchi, Event name collection method for information distribution corresponding to place and time using SNS data (Japanese edition), Proceedings of the 81st National Convention, 411-412, 2019, retrieved 2021/4/20.

[2] Kenro Aihara , Grasping dynamism of tourist behavior and thoughts through big data analysis -Case studies on analyzing trajectories and microblogs of foreign visitors (Japanese edition), Information Management Vol.59, 743-754, February 2017, retrieved 2021/12/15.

[3] Kaoru Okamura, Mototsugu Fukushige, How to Promote Repeaters? Empirical Analysis of Tourist Survey Data in Kansai Region (Japanese edition), Discussion Papers in Economics And Business , Vol. 07-42 , November 20007, retrieved 2021/4/20.

[4] Richard W. Butler, The concept of a tourist area cycle of evolution: implications for management of resources, Canadian Geographer 24, 5-12, June 2008. retrieved 2021/4/20.

[5] Japan tourism agency, Statistics of the number visitors, Statistics/White paper (Japanese edition), retrieved 2021/4/20.

[6] Regional economy society analyzing system, A regional economic analysis system (Japanese edition), retrieved 2021/5/26.

[7] Ministry of Land, Infrastructure, Transportation, and Tourism, An evaluation method of "accessibility of service" for local public traffic (Japanese edition), retrieved 2021/4/24.

[8] Urban Research Department, National Institute for Land and Infrastructure Management, A guideline for usage of accessibility indexes (Draft) (Japanese edition), retrieved 2021/4/20.

[9] Recruit Co., Ltd., Tourism Guidance, Jalan Net, retrieved 2021/4/20.

Authors Introduction

Dr. Takao Ito

Dr. Takao Ito is Professor of Management of Technology (MOT) in Graduate School of Engineering at Hiroshima University. He is serving concurrently as Professor of Harbin Institute of Technology (Weihai) China. He has published numerous papers in refereed journals and proceedings, particularly in the area of management science, and computer science. He has published more than eight academic books including a book on Network Organizations and Information (Japanese Edition). His current research interests include automata theory, artificial intelligence, systems control, quantitative analysis of inter-firm relationships using graph theory, and engineering approach of organizational structures using complex systems theory.

Dr. Tsutomu Ito

Dr. Tsutomu Ito is Assistant Professor of the Department of Business Administration at National Institute of Technology, Ube College, Japan. 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.

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. Matsuno has published several research papers and books in that area. He is a member of the Board of the Directors of Japan Society for Information and Management.

Dr. 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.

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.

A Research on Image Dehazing Technology for Image Enhancement

Haokang Wen *, Chang Sheng

College of Electronic Information and Automation, Tianjin University of Science and Technology, 300222，China

*E-mail: * 504912967@qq.com*

Abstract

Image defogging is to study the method of image enhancement in foggy weather with low definition and lighter color. Image defogging technology aims to improve image contrast and scene clarity, and has broad application prospects in the fields of target recognition, traffic navigation, and remote sensing. Three defogging algorithms in image defogging technology are introduced based on the current research status: global histogram equalization, local histogram equalization, and Retinex algorithm. This article introduces the main steps of the three algorithms and discusses the advantages and disadvantages of each algorithm. Finally, the processing results of these three algorithms are compared and comprehensively evaluated.

Keywords: digital image processing, image enhancement, foggy degraded image, dehazing algorithm

1. Introduction

In recent years, with the development of computer vision systems in the fields of military, transportation, and security monitoring, as well as the impact of haze on people’s real lives in previous years, image defogging has become an important research direction of computer vision and is currently widely used in the satellite remote sensing system, aerial photography system, target recognition system and outdoor monitoring system, etc. The visibility of outdoor images is greatly degraded due to the presence of fog[1], the color of the image is gray and white, the contrast is reduced, the object characteristics are difficult to recognize, the visual effect is deteriorated, and the image is very low in viewability. This will seriously affect the post-processing of the image, so image dehazing is required.

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 global histogram equalization algorithm performs global processing on the histogram of an image to realize the contrast enhancement of the entire image. The global histogram equalization method is simple and easy to implement, and the processing speed is fast. The disadvantage is that the contrast enhancement effect is not obvious, and the depth of the scene in the same image is diverse, and the degree of degradation is also different. So this method of global histogram equalization does not work well.

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

In order to ensure that the local area of interest obtains the required enhancement effect, it is necessary to perform local histogram equalization. The local histogram equalization algorithm disperses the histogram equalization operation to all local areas of the image, and adaptively enhances the local information of the image through the superposition of the local operations, so that the contrast of each area of the image can be greatly improved. But this method is computationally intensive.

3. Retinex Algorithm

Retinex is a commonly used image enhancement method based on scientific experiments and scientific analysis. It was proposed by Edwin H. land[3].The basic theory of Retinex theory is that the color of an object is determined by its ability to reflect long wave (red), medium wave (green) and short wave (blue) light. Color uniformity is the basis of color uniformity. Different from the traditional linear and nonlinear methods, Retinex can achieve a balance in dynamic range compression, edge enhancement and color constancy. Therefore, Retinex can enhance various types of images adaptively.

The main steps are as followed:

According to the image formation model, an image can be expressed as:

$$\mathrm{I}(x, y) = \mathrm{S}(x, y) \cdot \mathrm{R}(x, y) \tag{1}$$

(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 $\mathrm{I}(x,y)$.

$$\begin{aligned}\mathrm{I}'(x, y) &= \log[\mathrm{S}(x, y) \cdot \mathrm{R}(x, y)] \\ &= \log\big(\mathrm{S}(x, y)\big) + \log\big(\mathrm{R}(x, y)\big)\end{aligned} \tag{2}$$

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).

$$\mathrm{R}(x, y) = \exp\big(\mathrm{G}(x, y)\big) \tag{5}$$

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:

$$\mathrm{R_i}(x, y) = \log \mathrm{I_i}(x, y) - \log[\mathrm{I_i}(x, y) * \mathrm{F}(x, y)] \tag{6}$$

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).

$$\mathrm{F}(x, y) = \mathrm{K} \cdot e^{\frac{-(x^2+y^2)}{\sigma^2}} \tag{7}$$

Among them, σ is the wrapping scale, K is the normalization constant. The wrapping function satisfies:

$$\iint \mathrm{F}(x, y)dxdy = 1 \tag{8}$$

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]. The specific formula of MSR is as followed:

$$R_i(x,y) = \sum_{k=1}^{K} w_k \{\log I_i(x,y) - \log[I_i(x,y) * F(x,y)]\},$$
$$i = 1, \dots, N \qquad (9)$$

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, $i \in (R,G,B)$. $I_i(x,y)$ 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 the Fig.1, it is the flow chart of MSR algorithm.

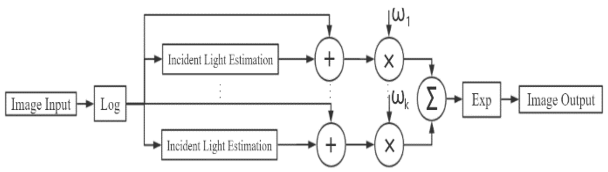


Fig.1 MSR algorithm flow chart

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.

Fig.2 Three algorithm processing results

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.

5. Conclusion

This article introduces the traditional image enhancement algorithm of global and local histogram equalization and the Retinex algorithm. Regardless of the specific reasons for image degradation, only from the perspective of image processing, based on improving the contrast of the image, the commonly used method is histogram equalization, and its algorithm is simple and easy to implement. It can be seen from the experimental results that the histogram equalization algorithm greatly reduces the influence of impurities on the image contrast enhancement, so this method can achieve better clarity when processing foggy scenes. However, this method is only suitable for degraded images with a single depth of field. For degraded images with complex depth of field, a certain degree of blocky response will appear while the image is sharpened, especially the detailed information of the distant scene in the image appears blurred.

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.

Authors Introduction

Mr. Haokang Wen

He is the third-year graduate student of 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 is the third-year graduate student of Tianjin University of Science and Technology. His research interests are machine learning and deep learning, and he has published several papers during the school.

Design of Intelligent Daylily Picking Robot

Jiaxin Li

College of Electronic Information and Automation, Tianjin University of Science and Technology, 300222，China

E-mail: L21807939 @tust.edu.cn

www.tust.edu.cn

Abstract

Daylily as a daily food crop, can significantly reduce serum cholesterol, root, leaves, flowers after treatment can be used as medicine, has a high nutritional value. But the picking conditions of daylily are bad, and long-term picking is likely to cause extremely serious harm to human body, especially to hands. Through the study and summary of the biological characteristics of daylilies, this paper aims to develop a kind of intelligent picking mechanism with strong applicability, The main work includes: designing the binocular recognition system of the intelligent picking mechanism of daylily; Determine the structure type of picking mechanism, and complete the design of the intelligent picking mechanism of daylily.

Keywords: Picking Robot, Daylily, Agricultural, Robot Design, Machine Vision

1. Introduction

As daylily flower planting scale increasing, the traditional manual picking gradually appeared the bad side, one is daylily flower picking very pay attention to detail, early harvest can make the production, because of too much water too late picking and affect mouthfeel, and artificial picking long time consuming and inefficient, easily lead to miss the best harvest time. Second, the cost of manual picking has been rising in recent years. With a great deal of workers pouring into high-earning positions, there is a shortage of labor force for such boring and repetitive work as daylily picking. Third, because the stamen part of daylily contains toxicity, a long time, high frequency of picking will be likely to touch the stamen at the hand part of damage, even skilled farmers are difficult to ensure that the whole body can withdraw.

At present, there is no intelligent daylily picking robot on the market. Based on this, we need an intelligent daylily picking robot that can operate autonomously to realize the automatic operation of daylily picking.

2. Hardware and Software of Control System

The hardware of intelligent daylily picking system consists of three main modules: upper and lower machine, picking arm and sensor. The corresponding functions can be realized by controlling different modules. Overview of control system structure is shown in Fig.1.

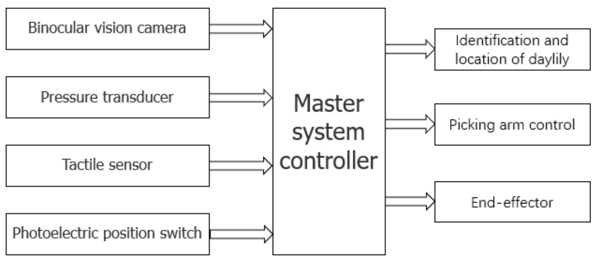


Fig.1 Control system structure

2.1. Control system hardware

The hardware part of the control system needs to have three basic functions: real-time information acquisition, analysis and processing, feedback control. Real-time information acquisition is accomplished by data acquisition and image analysis and processing.

Real-time information data is mainly collected by sensors, and the different signals fed back by pressure sensors and tactile sensors enable the system to locate its position and posture in real time.

Image analysis and processing is mainly through the binocular vision camera to take pictures, and then sent to the master controller of the system, through the algorithm to identify and locate the shape, position and surrounding environment of mature daylily and other useful information.

The feedback control part is realized by the upper and lower computers and internal algorithms and circuits. For example, in motion control, the obtained pose information is compared with the route planned by the master controller to realize real-time motion control.

The output part of the system is composed of controllable steering gear and end-effector. After receiving relevant signals, the manipulator is controlled to move to the specified position and the end-effector is closed to complete the picking. Hardware structure of control system is shown in the Fig.2.

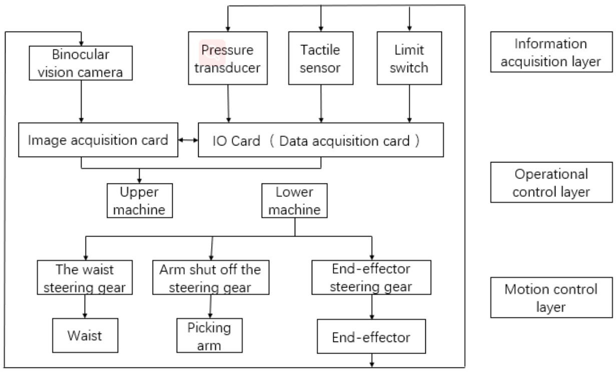


Fig.2. Control system hardware diagram

2.2. Control system software

In the real-time control system of picking mechanism, the control software is divided into host computer and servo controller.

Main functions of host computer: information processing, human-computer interaction, communication, etc.

- Information processing module: process the data transmitted by the sensor, segment and identify the daylily image, locate the spatial coordinates of mature daylily, and plan the picking path.
- Human computer interaction: display the real-time data transmitted by the sensor, display the temperature, humidity, and other states in the current environment, and have a debugging interface.
- Communication module: correct the received pose signal and then transmit it to the lower machine, receive the pose signal of picking arm and end effector and the opening and closing signal of end effector.

Screenshot of some program block diagrams is shown in the Fig.3.

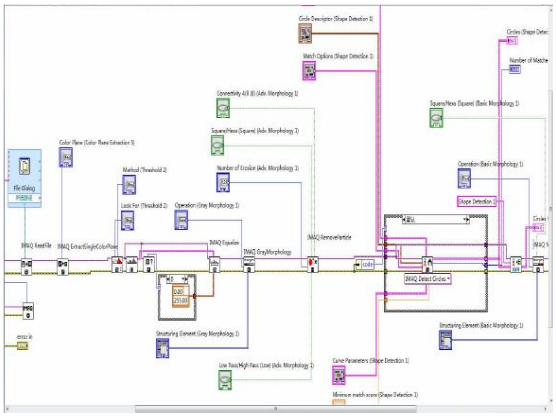


Fig.3. Main program block diagram of binocular vision system

2.3. Work flow of picking mechanism

After the system is started, all equipment will be initialized to ensure the normal operation of the picking arm. Then, the picking robot will complete the three-dimensional coordinate calculation of the daylily on the left side of the robot through the binocular camera and the upper computer, and plan the motion path of the picking arm, control the picking arm to complete the specified action, transport the end effector to the positioned cauliflower [1], close the end effector, realize the clamping and cutting of cauliflower, and complete the picking. Operate continuously until there is no mature daylily within the left visual range. After resetting the picking arm, the robot rotates 180 ° counterclockwise and repeats the previous operation within the right visual range. When there is no mature daylily in the left and right visual range, the picking arm is reset. Workflow of daylily intelligent picking mechanism is shown in the Fig.4.

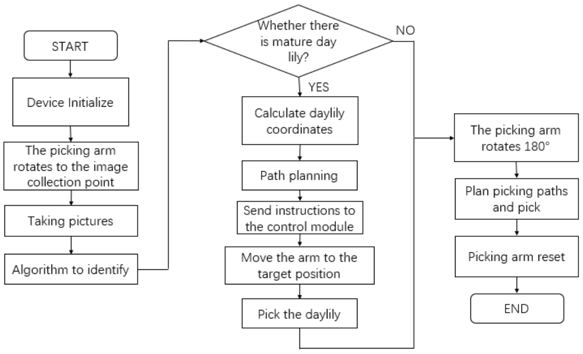


Fig.4. Workflow of daylily intelligent picking mechanism

3. Binocular Vision System Calibration

3.1. Visual system design scheme

The hardware of binocular vision system equipped with intelligent daylily picking robot in this paper is mainly as follows：

(1) A binocular camera that captures images

(2) Image transmission equipment

(3) Industrial computer in charge of image analysis and processing

Fig.3. Binocular vision camera

Binocular stereo camera is shown in Fig.3. The communication mode is TCP/IP protocol or SDK function call interface. The specific parameters are shown in Table1.

Table 1. Main parameters of binocular stereo camera

Parameters of binocular stereo camera	
Model	ZED STERE CAMERA zed2
Camera sensor	1/3“Backside lighting sensor”
Depth frame rate	100FPS
Detection range	20m(3D)；40m(2D)
Output resolution	3840×1080（1080P）
Three high	175mm×30mm×33mm

In this paper, the ZED second-generation binocular camera is selected as the image acquisition equipment. The left and right binocular cameras are coplanar, which fundamentally solves the offset error generated by the left and right cameras when they are installed separately. Binocular vision system construction scheme is shown in Fig.4.

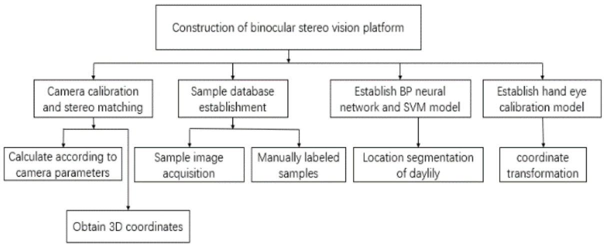


Fig.4. Vision system scheme of picking robot

The ZED2 has a 2μm pixel sensor and a backside light sensor. It has low low-light sensitivity and can collect super clear pictures of daylily for easy software processing [2]. Segmentation localization image of daylily is shown in Fig.5.

Fig.5. Segmentation and location image of daylily

3.2. Binocular camera calibration

For the intelligent cauliflower picking robot, after the threshold segmentation of the cauliflower image, the left and right cameras must be calibrated before calculating the spatial three-dimensional coordinates of the mature cauliflower. Because the binocular camera is used, binocular stereo calibration is also required.

- Stereo calibration of binocular camera

In order to calculate the spatial coordinates of mature daylily, it is necessary to establish the geometric model of camera imaging for the left and right cameras, which is to calibrate the left and right cameras [3]. The transformation matrix of two camera coordinate systems consists of a 3×3 rotation matrix R and a translation vector T. Located in the world coordinate system, the P coordinate of a point in space is (X_w, Y_w, Z_w), The abscissa of the imaging points of the left and right cameras is (X_l, X_r). The relationship between the rotation matrix from the world coordinate system to the camera coordinate system $(R_l,\ R_r)$, Translation vector from world coordinate system to camera coordinate system $(T_l,\ T_r)$ and Spatial coordinates(X_w, Y_w, Z_w) can be expressed as:

$$\begin{cases} X_l = R_l X_w + T_l \\ X_r = R_r X_w + T_r \end{cases} \quad (1)$$

Order $R = R_r R_l^{-1}$, $T = T_r - RT_l$. The simplified formula (1) can be obtained:

$$X_r = RX_l + T \quad (2)$$

- Remote AR registration method

There is a remote AR registration method in which the virtual robot coordinate system coincides with the physical robot coordinate system, which is an efficient and cheap human-computer interaction method based on RGB-D imaging and attitude teaching device. Through human-computer interaction, the local operator can effectively complete the trajectory planning of the remote robot (including the path planning and attitude planning of the robot end effector) [4].

4. Conclusion

In this paper, the rationality and practicability of the structure of cauliflower intelligent picking mechanism still need to be further verified. At present, the cauliflower picking robot is still in the design stage. The follow-up research work is to make the first generation of picking robot, and then carry out model training and simulated picking of cauliflower in the laboratory. After the success rate reaches a certain degree, select the flat area for natural picking in the field, and make version changes according to the actual situation, Finally, an intelligent picking robot that can participate in the actual production will be formed.

References

1. Feng Ma et al. (2016). A novel approach of collision assessment for coastal radar surveillance. *Reliability Engineering and System Safety*, 155pp. 179-195.
2. E.J. van Henten et al. (2002). An Autonomous Robot for Harvesting Cucumbers in Greenhouses. *Auton. Robots*, 13(3), pp. 241-258.
3. Zhengyou Zhang. (1999). Flexible camera calibration by viewing a plane from unknown orientations. *Proceedings of the Seventh IEEE International Conference on Computer Vision*. doi:10.1109/iccv.1999.791289
4. Pan Yong et al. (2021). Augmented reality-based robot teleoperation system using RGB-D imaging and attitude teaching device. *Robotics and Computer-Integrated Manufacturing*, 71.

Authors Introduction

Mr. Jiaxin Li

He received his Bachelor of Engineering degree from the school of Electronic Information and Automation of Tianjin University of Science and Technology in 2021. He is acquiring a master's degree in engineering from Tianjin University of Science and Technology.

50KN Compression Spring Fatigue Testing Machine Design

Peng Lu *, Peng Jia

Tianjin University of Science and Technology, China

*E-mail: * lupeng970504@163.com*

www.tust.edu.cn

Abstract

Spring fatigue test is the key process of spring performance testing, especially for automotive suspension springs, train damping springs, engine valves and other critical parts of the spring, must do the reliability assessment of spring fatigue performance. Because the requirements of different springs are different, the frequency and amplitude used in the test also have different requirements. In this case, a compression spring fatigue tester was developed and designed to test the maximum number of cycles of a spring under a given failure condition by applying a cyclic variable load to the spring. Through the design of the compression spring fatigue testing machine, the test prototype is finally developed.

Keywords: spring, fatigue test, tester, structural design

1. Introduction

The spring is an elastic element that can transform mechanical work or kinetic energy into deformation energy, or deformation energy into mechanical work or kinetic energy when it deforms and returns to its original state.

In actual engineering, most springs work under variable loads, and their working stress is often lower than the yield strength of the material. Spring in the role of this variable load, after a longer period of operation and the phenomenon of operational failure is called the spring fatigue damage[1,2]. Fatigue damage is the main form of spring failure, according to statistics, about 80% or more of the spring failure is caused by fatigue damage [3]. With the development of modern machinery in the direction of large-scale, many springs in high temperature, high pressure, heavy load and corrosion and other harsh operating conditions, fatigue damage accidents are numerous [4].

Therefore, experimental verification of the fatigue strength design of springs, correct evaluation of their true fatigue characteristics, and verification of the true effect of the fatigue design are of great importance to improve the reliability and service life of mechanical products [5]. The compression spring fatigue tester designed in this paper can effectively determine the mechanical properties, process properties and fatigue strength of coil springs. It can help spring manufacturers to improve the quality of their products and ensure the safety and reliability of their products.

2. Main Technical Indexes

The main technical indicators of the spring fatigue tester are.

- Maximum load: ±50KN.
- Maximum loading displacement: ±15mm.
- Test frequency: 0.1~10Hz.
- Load displacement deviation amount: ≦3%.
- Test length space: 60-300mm.
- Counting range: 1~100000 times.

3. System Composition and Working Principle

The spring fatigue tester designed in this paper is realized by plane mechanical type, and the tester is composed of motor, transmission mechanism, frame, fixture and other parts. The structure of the spring fatigue tester is shown in Fig.1.

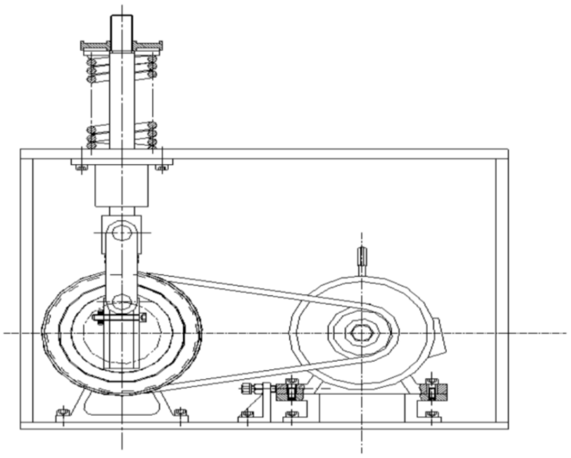


Fig.1 Test machine structure

At the beginning of the test, the motor drives the eccentric wheel to rotate, and at the same time drives the driving rod to do round-trip motion to apply a certain frequency of dynamic load to the spring under test. The motion system uses a crank slider mechanism to allow the driving rod to move in the guide rail, and the motor frequency control is used to control the loading frequency of the test, and sensors are used to count the number of tests. At the same time, in order to ensure the safety of the test, the tester can achieve automatic alarm and stop when the specimen fatigue fracture.

4. Drive Train Design

The scheme design of the mechanical drive system is an important part of the overall design of the fatigue testing machine. Its fundamental task is to convert and transmit the power generated by the motor to the actuating part according to the needs of the system. The design of the transmission scheme is to complete the power distribution ratio, calculate the motion and power parameters of the transmission device, and the selection of the motor.

4.1. Motor selection

The total efficiency of the system is assumed to be $\eta = 0.8$ in the design because of the power loss in the transmission process. Calculate the work required for the tester to make one stroke.

$$\omega = \frac{1}{2} sf = \frac{1}{2} \times 15 \times 50 = 375j \quad (1)$$

If the transmission ratio is $i = 2.9$ and the motor speed is $n_1 = 1400\mathrm{r/min}$, the spindle speed is $n_1 = 480\mathrm{r/min}$. The time required for one rotation of the spindle is 0.125s. The spindle power can be calculated.

$$\mathrm{P} = \frac{\omega}{\mathrm{t}} \approx 3000\mathrm{W} \quad (2)$$

motor power:

$$P_E = \frac{P}{\eta} \approx 3750W \quad (3)$$

Therefore, we can temporarily choose the servo motor model Y2-112M2-4, whose technical parameters is shown in Table 1

Table 1. Motor parameters

Model	Y2-112M2-4
Power Rating(kw)	4
Rotational Speed(r/min)	1400
Stall torque/rated torque	2.2
Maximum torque/rated torque	2.3

4.2. Drive train design

Because the machine overload easily damage the motor, in order to protect the motor, so the mechanical drive form of belt drive is used.

The belt drive is an extremely widely used flexible drive. Advantages of belt drive: smooth operation without noise; can moderate shock and absorb vibration; can prevent damage to other parts when the machine is overloaded; suitable for transmission with large center distance; simple structure and low cost.

According to the smooth load of V-belt, 16-hour two-shift working system, take $\mathrm{K}_A = 1.3$. The calculated power can be obtained.

$$P_{ca} = K_A P = 5.2KW \quad (4)$$

According to the calculated power $P_{ca} = 5.2KW$ and the small pulley speed $n_1 = 1400\mathrm{r/min}$, by checking the belt type chart, we can know that we should choose the A type V belt. The belt selection diagram is shown in Fig.2.

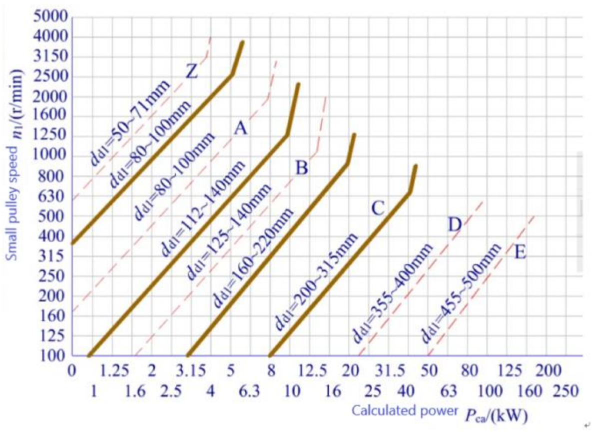


Fig.2. Belt selection diagram

A type V-belt small pulley diameter range is $d_{d1} = 80\sim100\text{mm}$. We take the small pulley diameter as $d_{d1} = 100\text{mm}$, and since the transmission ratio is $i = 2.9$, the large pulley diameter is derived as $d_{d2} = 290\text{mm}$, and finally the large pulley diameter is selected as $d_{d2} = 280\text{mm}$ according to the actual situation.

To check the correctness of the drive belt selection, a ratio error check and a belt speed check are performed.

(i) Transmission ratio error check.

$$i_e = \frac{d_{d2}}{d_{d1}\times(1-\varepsilon)} = 2.86 \tag{5}$$

ε is the elastic sliding rate. The percentage error of the transmission ratio is.

$$i = \frac{i - i_e}{i} \times 100\% = 1.3\% \tag{6}$$

The error percentage is less than the allowable error percentage and meets the requirements.

(ii) Band speed check.

$$v = \frac{\pi\times d_{d1}\times n_1}{60\times1000} = 7.33\ m/s \tag{7}$$

The belt speed satisfies $v > 5\ m/s$, so the belt speed is appropriate.

5. Mechanical Component Design

In the design of the mechanical components of the spring fatigue tester, the design of the spindle and eccentric wheel mechanism is the most important.

5.1. Main shaft

The spindle is one of the important parts that make up the machine. The main function of the spindle is to support rotating parts and transmit torque and motion. The working condition of the spindle directly affects the performance and quality of the whole machine.

When designing spindles generally the main focus is on designing the structure, diameter and strength of the spindle.

(i) Structure of the spindle.

The structure of the spindle depends mainly on the type, number, location and method of mounting and positioning of the fixtures, transmission parts, bearings and seals mounted on the spindle, as well as on the processability of spindle machining and assembly. In order to meet the stiffness requirements and to obtain sufficient thrust surface as well as to facilitate assembly, the spindle is often designed as a stepped shaft, i.e., the shaft diameter decreases from the front shaft diameter to the back. The spindle of this design, also designed as a stepped shape, was also designed as a solid shaft while meeting the stiffness requirements. The structure of the spindle is shown in Fig.3.

Fig.3. Main shaft structure

(ii) Spindle diameter.

In this design, the spindle is made of 45# steel. Its tensile strength limit is $\sigma_b = 600MPa$, the allowable bending stress is $[\sigma_{-1b}] = 55MPa$, the allowable shear stress is $\tau_T = 30\sim40MPa$, and the coefficient is $C = 106\sim108$. From this, the diameter of the spindle can be calculated.

$$d = C\sqrt[3]{\frac{P}{n_2}} = 19.7\sim21.7\ mm \tag{8}$$

Considering that there is an eccentric wheel installed on the shaft and a keyway at the end of the shaft, in order to ensure sufficient stiffness. Therefore, the estimated diameter range is increased by 3% to 5% and taken as 20.3~22.8mm. Final take $d = 22\text{mm}$.

(iii) Spindle strength check.

In this design, the spindle only transmits torque. Therefore, the strength check of the spindle can be calculated according to the torsional strength.

$$\tau = \frac{T}{W_T} = \frac{9.55\times10^6 P}{0.2d^3 n_2} \le \tau_T \tag{9}$$

In the above equation: τ is the torsional shear stress of the spindle; T is the torque; W_T is the torsional cross-sectional coefficient, for a circular section shaft $W_T \approx 0.2d^3$.

The final calculation yields $\tau = 37.37\text{MPa}$, which satisfies the strength requirement.

5.2. Eccentric wheel mechanism

When the stroke of the crank slider mechanism is small and the crank is short, the crank slider mechanism is usually designed as an eccentric wheel structure, using the eccentric pitch as the crank length. The spring fatigue tester in this design requires amplitude adjustment, i.e., the travel of the slider can be adjusted as needed, so the eccentric wheel is designed is shown in Fig.4.

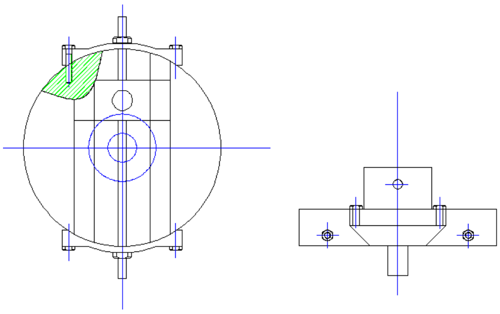


Fig.4. Eccentric wheel mechanism

In order to ensure the stability of the eccentric wheel mechanism, strength calibration of the eccentric wheel is required. The eccentric wheel is made of 45# steel and the slider is modulated. The circular shaft at the upper end of the slider is connected to the slider by welding, and the eccentric wheel is hinged between the circular shaft and the connecting rod. The maximum load transmitted to the drive rod by the eccentric wheel is $\text{F}_{max} = 250KN$. The articulated part of the slider and the connecting rod is a short circular shaft with a length of $\ell = 200\text{mm}$ and a diameter of $\text{d} = 20\text{mm}$. Calculate the maximum bending moment stress on the slider section.

$$\sigma_{max} = \frac{M_{max}}{W} \tag{10}$$

$$M_{max} = \text{F}_{max} \times \frac{\ell}{2} \tag{11}$$

$$W = \frac{\pi d^3}{32} \tag{12}$$

The maximum bending moment stress on the slider section is calculated to be $\sigma_{max} = 64MPa$. After the modulation treatment, the flexural fatigue limit of 45# steel is $\sigma_{-1} = 270MPa$. The comparison of the results shows that the strength of the eccentric slider meets the design requirements.

6. Conclusion

This design of spring fatigue tester realizes the fatigue test of compression spring. The machine has a simple structure, accurate and rapid running action, stable and reliable performance, easy to operate, etc. At the same time, the machine has good reliability, stability and versatility, and is suitable for use in small and medium-sized spring manufacturing plants and professional laboratories.

References

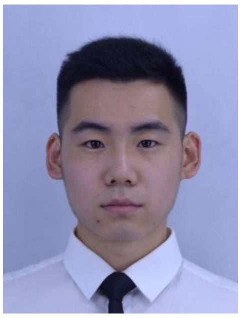
1. Akiniwa Y, Stanzl-Tschegg S, Mayer H, et al. Fatigue strength of spring steel under axial and torsional loading in the very high cycle regime, *International Journal of Fatigue*, 2008, 30(12): pp. 2057-2063.
2. El-Sheikh M M, Godfrey K, Manosudprasit M, et al. Force-deflection characteristics of the fatigue-resistant device spring: an in vitro study, *World Journal of Orthodontics*, 2007, 8(1): pp. 30-36.
3. Maciejewski J, Akyuz B. Spring Fatigue Fractures Due to Microstructural Changes in Service, *Journal of Failure Analysis & Prevention*, 2014, 14(2): pp. 148-151.
4. Mao Q, Jianghua H, Zhao K. Design for Air Spring Fatigue Testing Machine by Simulation of Actual Working Conditions. *Machinery & Electronics*, 2016, 34(09): pp. 50-54.
5. Wang R Z, Yang W M. An optimized method of loading for the helical spring fast fatigue test, *Metal Products*, 2013, 39(1): pp. 23-29.

Authors Introduction

Mr. Peng Lu

He is a second year master student at Tianjin University of Science and Technology. His research is about multi-sensor fusion, artificial intelligence, and robotics.

Mr. Peng Jia



He is a first year master student in Tianjin University of Science and Technology. His research is about embedded hardware design, mechanical and electrical control.

New Intelligent Unmanned Retail Shopping Container Design

Peng Lu*, Yiting Gao

Tianjin University of Science and Technology, China

*E-mail: * lupeng970504@163.com*

www.tust.edu.cn

Abstract

During the COVID-19 epidemic, unmanned retail has seen new opportunities for growth. The common unmanned retail methods are vending machines and unmanned supermarkets, which have the disadvantages of low product variety and high store costs, respectively. In this paper, a new intelligent unmanned retail shopping container is designed, integrating unattended technology with the container. Through the two-dimensional code, electronic anti-theft, RFID and other technologies, designed to achieve the shopping process of the goods that are taken away, the whole data. At the same time, the designed shopping containers can be deployed closer to consumers, providing a safer and more convenient shopping experience for people.

Keywords: intelligent, unmanned retail, unattended, shopping container

1. Introduction

Since the beginning of 2020, the COVID-19 virus has been wreaking havoc around the world and countries around the world have taken a variety of measures to promote social distancing to slow the spread of the virus. Against this backdrop, traditional forms of retailing have been severely impacted, while the unmanned retail sector has been presented with new opportunities for growth.

The overall unmanned retail industry is growing rapidly, with the unmanned retail market in China alone set to exceed US$10 billion by 2020. In the unmanned retail sector, unmanned kiosks are still the dominant format. In addition, unmanned stores have gradually emerged in recent years, with companies such as Amazon, Jingdong and Alibaba launching their own brands of unmanned stores [1,2].

The aim of unmanned stores is to combine the benefits of physical and online retail using IoT and AI technologies [3]. However, each of these two forms of unmanned retail has its own distinct disadvantages. Unmanned kiosks are limited by their own capacity and can offer customers a very limited variety and number of items. And because of the small size of the vending machine, the amount of goods stored is not much, requiring staff to frequently replenish the goods, which invariably increases the cost of labor.

Most of the unmanned convenience stores operated by e-commerce companies are laid out in stores along the streets around large communities and office buildings. Unmanned convenience stores require a large operating area and cannot really reach consumers. At the same time, the high rents of street-level stores around large communities and office buildings also bring significant operating costs for businesses.

Based on the above discussion, a new unmanned retail model is designed in this paper, which is the new intelligent unmanned retail shopping container. The new intelligent unmanned retail shopping container area and ordinary small convenience store the same but does not require a fixed storefront, companies can change the shopping container operation location at any time, in the convenience of a greater increase [4].

The types and quantities of goods in the shopping containers are also the same as in a normal convenience store, and can achieve a high degree of matching between the container arrangement and the location of the consumer's shopping needs. At the same time, unattended technology is used in the new intelligent unmanned retail shopping containers to truly realize the full self-service and digitalization of the shopping process.

2. Overall Design

The new intelligent unmanned retail shopping container is a 6 meters long, 3 meters wide, 2.9 meters high, weighing within 4 tons of a shopping container, which covers an area of about 18 square meters. The new intelligent unmanned retail shopping container mainly adopts four key technologies: access control system, RFID IOT, automatic settlement and intelligent monitoring. In terms of structure, the shopping container adopts a design that separates the entrance from the exit and the shopping area from the settlement area. The overall design effect of the new intelligent unmanned retail shopping container is shown in Fig.1.

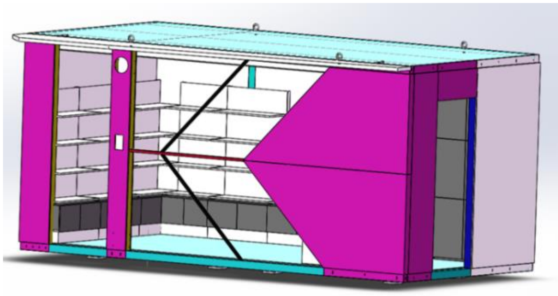


Fig.1 Shopping container modeling drawing

3. Hardware Design

3.1. Industrial control computer

Industrial control computer, is a general term for a tool that uses a bus structure to detect and control production processes and electromechanical equipment and process equipment. It has important computer attributes and features, such as having a CPU, hard disk, memory, peripherals, interfaces and touch screen with operating system, control network and protocols, computing power, and friendly human-machine interface.

The model number QY-P070-HM industrial control computer is used in this design. It has a highly scalable multi-core series application processor, 2GB DDRIII high-speed memory on board, and 8GB FLASH storage. The industrial control computer also provides on-board WiFi, 100 Gigabit Ethernet, and 3G/4G network connectivity, allowing for a flexible network application environment.

3.2. Access control

The access controller used in this design is a four-door unidirectional network controller, the model selected WG2054. WG2054 has a series of functions such as real-time monitoring, illegal sweeping alarm, number of people in the door inquiry and limited number of people. WG2054 adopts high-speed operation circuit, fast computing speed, even after full storage swipe card can open the door instantly, without any delay. At the same time, the access control system uses TCP/IP communication protocol, which effectively optimizes the communication quality, communication speed and real-time communication of the system and can provide faster services to consumers.

3.3. Alarm host

The alarm host of the new intelligent unmanned retail shopping container adopts the model 968C-NET anti-theft alarm. When the front-end detector is triggered, the detector will send a signal to the alarm host, and the alarm host will immediately drive the tweeter to sound the alarm on site after receiving the signal, and at the same time call the administrator's cell phone alarm. The alarm host can not only link wireless cameras for regular deployment, but also remote monitoring, remote listening and fire alarm through cell phone APP. It can effectively face the sudden fire and power failure of the shopping container and deter customers from stealing in the shopping container.

4. Software Design

4.1. General system control

The overall control of the system is the fundamental guarantee that the new intelligent unmanned retail shopping containers can operate properly. The core of the shopping container control system design is to use the embedded system to complete the control of the whole program by adding several task blocks to achieve the required functions.

Through the powerful multi-process and multi-task real-time management function of the embedded system, thus greatly simplifying the control of the whole system. The most important task modules in the whole shopping container control system are: intelligent access control, RFID IOT, automatic settlement and intelligent anti-theft.

The control system is run by first testing the network connection part to test whether the Wi-Fi communication is smooth and whether the console is properly connected to the cloud server. After ensuring the network communication is smooth, enter the access control task, when the customer is ready to enter the cashier room with goods, close the exit door of the cashier area at this time and wait for the code sweeping module to complete the task of code sweeping and upload to the cloud server side.

After the data comparison on the server side completes the calculation of the total amount of goods, the payment QR code is generated to the customer. After the customer finishes the payment on the cell phone, the server side will transmit the information to the main control system, and the controller will open the exit door to let the customer leave and then open the entrance door to welcome the next customer again. Throughout the operation process, there are voice prompts to ensure that customers are familiar with the operation process.

4.2. Intelligent access control

As the new intelligent unmanned retail shopping container is unattended during daily operation, an intelligent access control system is designed at the entrance of the shopping container to ensure customers' shopping safety and reduce the rate of commodity theft and loss. The smart access control system uses attribute-based access control.

The smart access control system uses the Attribute-Based Access Control (ABAC). ABAC implements an attribute-based access policy that provides greater security and flexibility while accomplishing fine-grained access control [5], with the following four main types of attributes.

(1) Subject attributes: Each attribute that can determine the identity of the accessing user.

(2) Object properties: the door lock itself has the properties.

(3) Environmental attributes: time, network rate, etc. are environmental attributes.

(4) permission attributes: different access rights attributes of intelligent access control.

The ABAC model submits unlocking requests based on the user's attributes and needs. The request is uploaded to the Policy Enforcement Point (PEP), which sends the door lock request to the Policy Decision Point (PDP). The PDP is authorized by application and sent down to the PEP. PEP returns the authorization to the user who applied for it, and the user gets the authorization to open the door. The implementation flow of ABAC is shown in Fig.2.

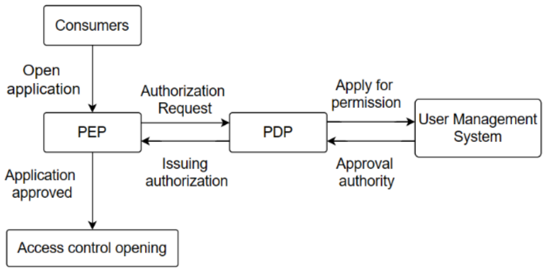


Fig.2 ABAC Process

4.3. RFID identification system

In the design of new intelligent unmanned retail shopping containers, RFID technology is one of the most central technologies. RFID recognition system mainly includes reader, wireless radio frequency electronic label and network information data exchange equipment, etc. The composition of RFID identification system is shown in Fig.3.

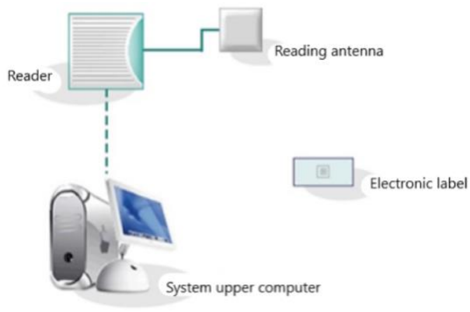


Fig.3 RFID identification system

The new intelligent unmanned retail shopping containers use passive RFID technology, and its basic working principle is not complicated. Passive RFID in

RFID tags without built-in power, when the electronic tag into the reader magnetic field range, through the electromagnetic induction coil induction RFID reader issued by the microwave signal, the chip to achieve a short supply of energy and complete data transmission [6]. The reader reads the information and decodes it, and then sends it to the upper computer system for relevant data processing. Passive RFID is characterized by battery-free, no contact, no card swipe so not afraid of dirt, and the chip code for the world's only can't be copied, high security, long life.

4.4. Automatic settlement

The automatic settlement system of the new intelligent unmanned retail shopping container is a settlement system based on the above RFID tags, and the two are closely related and inseparable. The settlement process is shown in Fig.4.

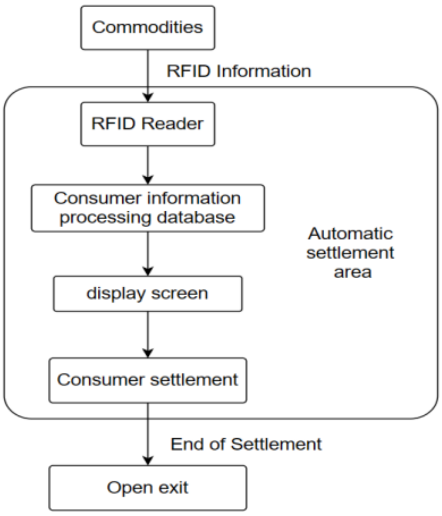


Fig.4 Settlement process

The automatic settlement system consists of three parts: commodities, automatic settlement desk and upper computer. When the customer selects the product and enters the automatic checkout area, the RFID reader will automatically identify the RFID electronic tag attached to the outer package of the product. After receiving the commodity information, the consumer information processing upper system settles the commodity and transmits the corresponding consumption information to the display through the serial port. The consumer's purchase information and payment QR code are eventually fed back on the display, and the customer checks out by paying online.

At the same time, this design links the automatic billing system to the access control system for linkage control. When the customer enters the settlement area for payment, the entrance and exit of the settlement area will be in the closed state, at this time the system will send instructions through the serial interrupt to open the RFID reader for tag reading. Wait for the customer to finish paying before the exit door of the checkout area opens and allows the customer to leave. When the infrared radar in the checkout area does not detect a customer, it opens the entrance door to receive the next customer who is ready to check out from the shopping area.

4.5. Intelligent anti-theft

In order to effectively reduce the theft rate of the unmanned retail shopping containers, this design uses the " Electronic Article Surveillance System " (EAS). The EAS system consists of transmitters, transmitting antennas, receivers and receiving antennas installed at the entrances and exits of shopping containers. RFID tags on goods with special printed circuitry inside, when the customer normally sweeps the code to settle the bill, and then carries the goods through the detection door equipped with detectors will not send an alarm, while the consumer will trigger an alarm if he carries unpaid goods through the detection door.

5. Conclusion

The new intelligent unmanned retail shopping container designed in this paper provides consumers with a variety of services such as identity verification, automatic settlement, and self-service payment by using unattended technology, allowing consumers to enjoy a high-quality and fast consumption experience without going through the merchant's manual service.

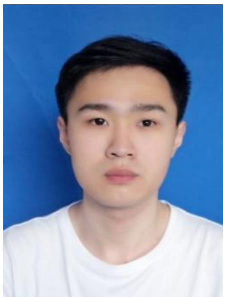
At the same time the new intelligent unmanned retail shopping containers can be deployed closer to consumers, which can greatly improve the current status of community shopping, significantly improve the business efficiency of unmanned retail and effectively increase the repurchase rate of goods.

References

1. Lee S H, Lee D W. Changes in Distribution Industry Using Information and Communication Technology, Journal of Advanced Information Technology and Convergence, 2018, 8(1): pp. 23-33.
2. Wu H C, Ai C H, Cheng C C. Experiential quality, experiential psychological states and experiential outcomes in an unmanned convenience store, *Journal of retailing and consumer services*, 2019, 51: pp. 409-420.
3. Guo B, Wang Z, Wang P, et al. DeepStore: Understanding Customer Behaviors in Unmanned Stores, *IT Professional*, 2020, 22(3): pp. 55-63.
4. Yao L, Shuai Y, Chen X, et al. A two-stage EBM-based approach to evaluate operational performance of unattended convenience store, *International Journal of Retail & Distribution Management*, 2020, 48(6): pp. 609-627.
5. Hu V C, Kuhn D R, Ferraiolo D F. Attribute-Based Access Control, *Computer*, 2015, 48(2): PP. 85-88.
6. Want R. An introduction to RFID technology, I*EEE Pervasive Computing*, 2006, 5(1): pp. 25-33.

Authors Introduction

Mr. Peng Lu



He is a second year master student at Tianjin University of Science and Technology. His research is about multi-sensor fusion, artificial intelligence, and robotics.

Ms. Yiting Gao

She is a first year master student in Tianjin University of Science and Technology. Her research is about deep learning, neural networks.

.

Java - based Dream Cloud ERP System - Inventory Management Subsystem Design and Implementation

Yiting Gao*, Peng Lu
Tianjin University of Science and Technology, Tianjin, China

*E-mail: *2801110578@qq.com*
www.tust.edu.cn

Abstract

With the continuous development of economy, large and medium-sized enterprises of heavy production tasks,more and more high to the requirement of inventory management. The traditional manual management will cause low efficiency of goods supply. The problems such as material management and data collection have become increasingly unable to meet the requirements of enterprise inventory management.The traditional inventory after making the decision need human convey,but the inventory management system as new way of management, inventory management system can provide managers with a convenient platform directly communicate need to complete the operation, task allocation directly down, people can be faster to implement.

Keywords: ERP, B\S, Management system, Autonomous

1. Introduction

Enterprise inventory management [1] is very complex and cumbersome. Due to the large variety of materials, different ordering, management and distribution channels, different management systems and various statistical reports among enterprises, a set of inventory management information system must be prepared for warehouse inventory management to realize computerized operation.

According to the current enterprise management system, the general inventory management system is always divided into several departments according to the mastered material categories, which are respectively responsible for planning, ordering, write off, delegated collection, acceptance and warehousing of materials.

At the same time, these departments will send materials and equipment according to the needs of each department of the enterprise, and make inventory and bookkeeping at any time. According to the needs of enterprise leaders and their own management, carry out quarterly and annual statistical analysis, and generate corresponding reports at the same time.

Through careful investigation of these situations, the following warehouse inventory management system is developed.

2. Operating Environment

The system is modeled on the structure of layui. The front end adopts bootstrap framework and the back end adopts SSM framework to realize the function of front and back end. In addition, it also applies other related technologies, such as Baidu echarts diagram. At the same time, a good configuration environment is also required.Configuration environment is shown in Table 1.

Table 1. configuration environment server

Server name	Server
Operating system name	Window10
Back end server name	Tomcat
Back end server model	Apache Tomcat/9.0.12

The database server is shown in Table 2, and other softwares are shown in Table 3.

Table 2. configuration environment server

Database server type	Server
Database name	MySQL
Database version	Mysql/8.0.15

Table 3. configuration environment server

Other software type	Server
Development platform	Myeclipse2015
Software name	Myeclipse
Software version number	Myeclipse5.0

3. Software Design

The importance of software architecture [2] is that it determines the main structure, macro characteristics, basic functions and characteristics of a system. For example, the key to the success of large-scale building design lies in the main structure, and the success of complex software design lies in the correctness and rationality of structural design at the macro level. Therefore, software architecture is the key to the success of the whole software design.

3.1. System structure design

System structure design is to disperse the structure of the whole system into four parts.

(1) The administrator can input the quantity and price of items in the initial inventory.

(2) The administrator can perform business logic operations such as transfer, disassembly and counting.

(3) The administrator can view the issue / receipt documents and generate and print the documents corresponding to the above operations.

(4) Administrators can view inventory history and print out.

3.2. Architecture design

The system adopts the system design style of browsing / server. The presentation layer is responsible for user input and output to customers. The function layer is responsible for establishing the connection to the database, generating SQL statements to access the database according to the user's request and returning the results to the client. The data layer is responsible for the actual database storage and retrieval and responding to the data processing request of the function layer, And return the result to the function layer.

Overall architecture design is shown in Fig.1.

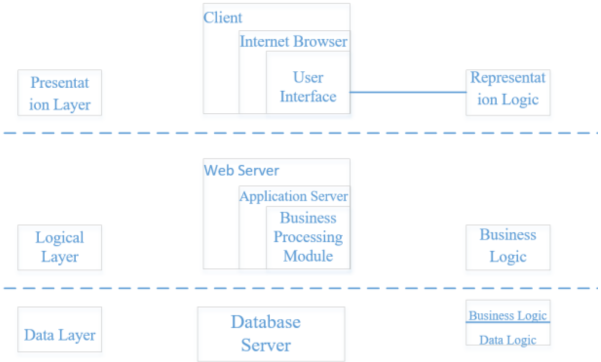


Fig.1. overall architecture design

3.3. System structure design

In the design, the system structure design is particularly important. The system is mainly designed for the administrator. It is divided into 9 modules: setting module, transfer document module, disassembly document module, count document module, miscellaneous issue document module, miscellaneous receipt document module, cost adjustment document module, inventory history module and statistical analysis module.The users in this program are inventory managers, which has nothing to do with customers.System structure design is shown in Fig.2.

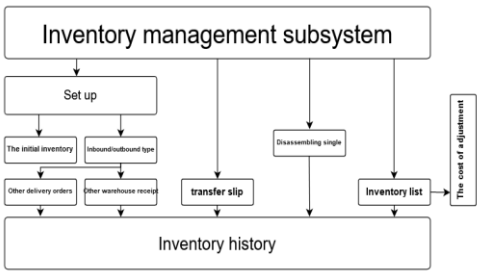


Fig.2. system structure design

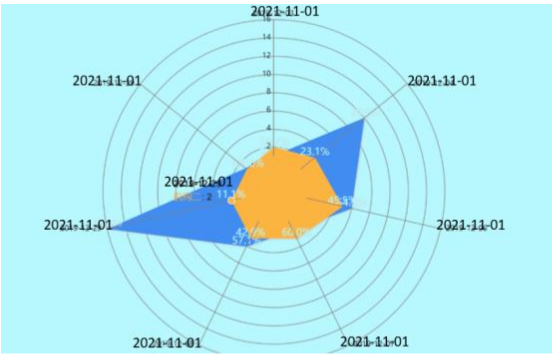
4. Project Testing and Statistic

The purpose of the test is to find out the possible problems in the system before the system goes online, check the system performance [3,4], and test whether the network request can be sent normally, whether the function can be realized stably and whether the operation is stable under the condition of cross domain. The project test is shown in Table 4.

Table 4. Project test

Test unit	Function description	test result
Issue / receipt category	New，modify，delete issue / receipt type	success
opening inventory	Manually enter (add), view and output (document) inventory goods	success
Transfer order	Add, modify, delete, query and output transfer documents	success
Disassembly order	Add, modify, delete, query and output transfer documents	success
Inventory sheet	Add, modify, delete, query and output transfer documents	success
Miscellaneous issue doc	Query and output issue doc	success
Miscellaneous receipt doc	Query and output issue doc	success
	Transfer doc operation generation record	success
Inventory history	Disassembly order operation generation record	success
	Count sheet operation generation record	success

Statistical analysis is the image display of the data to be displayed, so as to more intuitively feel the changes in inventory. Different graph edges have corresponding display data to help managers better manage and make judgments. The statistical analysis is shown in Fig.3.



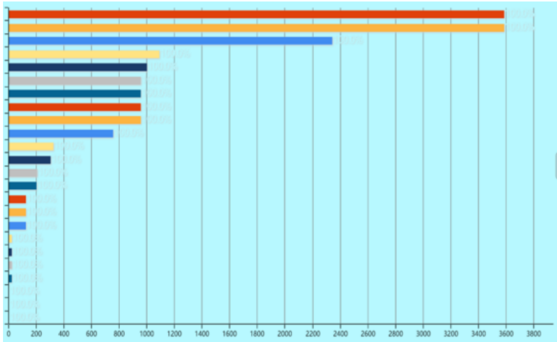


Fig.3.Statistical analysis

5. Conclusion

Dream cloud ERP system based on Java-inventory management subsystem v10. It is a management software based on inventory management in many parts of the enterprise. The software has powerful functions. It is mainly aimed at the related inventory management functions of the enterprise background administrator to manage inventory, such as inventory setting, opening inventory, transfer doc, disassembly doc, count doc, other receipt doc and other issue doc, In addition, there are functions such as inventory history and statistical analysis. The software interface is clean and simple to operate. Users can intuitively view the real-time results of system management. The system has a good user experience effect.

At present, the inventory management level of China's enterprises, especially small and medium-sized production enterprises, still stays on the basis of paper media. Such a mechanism can no longer adapt to the development of the times, because it wastes a lot of human and material resources. In the information age, this traditional management method is bound to be replaced by computer-based information management.

Therefore, in this case, it is very necessary to replace the manual inventory management system with such a collection, which has the advantages of fast retrieval, convenient search, high reliability, good confidentiality and low cost. Dream cloud ERP system based on Java - inventory management subsystem v10 provides great convenience for enterprises to process the data statistics of inventory information management, make the information clearer, make the data query easier, and promote better communication between enterprises. The system runs stably and is convenient for daily maintenance.

References

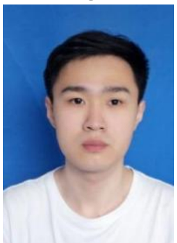
1. Chuangming Cao, Shaohua Dong, Yuhang Duan. Analysis of enterprise storage system management mode under the Internet plus background, *Logistics engineering and management*, 2021, 43(11): pp. 44-48+21.
2. Ying Li. Design of inventory management software for lightweight Supermarket Based on CS structure, *Procedia Engineering*, 2018, 21(05): pp.38-41.
3. Qile Qi, Research on Key Technologies of embedded computer software testing, Electronic production, 2021, 21:pp. 99-100+78.
4. Bo Weng, Yanfeng Yan, Research on software testing in big data environment, Internet Weekly, 2021, 22:pp. 42-44.
5. Lina Ma. Discussion on teaching reform of applied multivariate statistical analysis under the background of big data, *Mapping and spatial geographic information*, 2021,44(11): pp.87-88+92.

Authors Introduction

Ms. Yiting Gao

She is a first-year master candidate in Tianjin University of Science and Technology, majoring in neural network, deep learning.

Mr. Peng Lu



He is a second-year master candidate in Tianjin University of Science and Technology. His research is about deep learning, pattern recognition.

A Design of Micromouse Control System

Hucheng Wang

Tianjin University of Science and Technology, China

E-mail: 1776315920@qq.com

Abstract

This paper designs micromouse. The micromouse collects external information through six groups of infrared sensors. The STM32 single chip microcomputer adjusts the posture of the micromouse according to the environmental signal detected by the sensor, adjusts the PWM signal of the driving circuit, searches the unknown maze according to the software search algorithm, finds the shortest path in the strange environment, and on the premise of ensuring stability, Sprint from the starting position to the end of the maze at the fastest speed.

Keywords:micromouse,STM32,path search, sensor

1. Introduction

With the continuous development of science and technology in China, it has become a social development trend that intelligent robots replace human labor. Intelligent robots are widely used in many frontier innovation fields, such as smart home system, aerospace system, biomedicine and so on.

As one of the intelligent robots, maze walking computer mouse has the basic functions of detecting strange environment, automatically controlling movement, completing path search and planning.

Since the concept of micromouse was put forward in the 1970s, the design and manufacture of micromouse have been continuously optimized and improved. However, so far, the micromouse still has some disadvantages, such as heavy weight, inflexible steering, slow path search and so on. We optimized these existing problems and redesigned its mechanical structure. Creatively use the new infrared sensor layout to enable the computer mouse to collect external information more sensitively and quickly. Combined with the potential value search algorithm, the micro mouse can complete the sprint task in the maze more efficiently and quickly.

2. Mechanical Structure

2.1 Chassis design

The chassis should be as light as possible to reduce the impact of turning inertia caused by excessive weight on the smooth operation of the computer mouse, and have the strength to withstand the welding of components on its surface. Considering such problems, the printed circuit board is selected as the chassis of the micromouse.

2.2 Mobile mechanism design

The computer mouse moving mechanism is composed of a driving wheel and a driving wheel. The driving wheel is driven by a motor, and the driven wheel has no power. In order to prevent slipping in motion, it is required that there should be a certain friction between the tire and the ground. In order to make the computer mouse have high flexibility in sprint and turning, this design uses a low-cost plastic wheel hub with a diameter of 30 cm as the driving wheel of the computer mouse.

2.3 Transmission part design

This design adopts belt indirect transmission. The driving force output by this transmission mode is more appropriate, and it is relatively simple and convenient in actual installation, and the noise generated during operation is relatively small. The mechanical design drawing is shown in Fig.1.

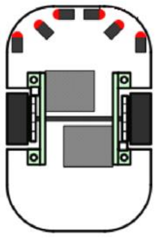


Fig.1. The mechanical design drawing

3. Hardware Design

3.1 Minimum system board circuit

1. Power supply circuit

The main power supply of maze computer mouse adopts 7.4v Mini lithium battery. Since the reference voltage of this design is 3.3V, ams1117 is adopted_ 3.3 the chip is used as a step-down module. The schematic diagram of power supply module is shown in Fig.2.

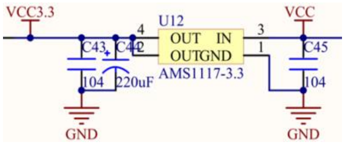


Fig.2. The schematic diagram of power supply module

2. System download and debugging circuit

The system adopts JTAG interface, also known as joint action working group, which is used to complete the burning, debugging, online debugging and other operations of the program [1]. The schematic diagram of JTAG download circuit is shown in Fig.3.

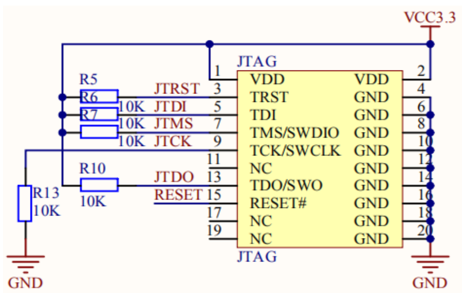
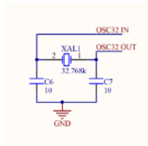


Fig.3. The schematic diagram of JTAG download circuit

3.System clock circuit

The frequency of low-speed clock source in this design is 32.768KHz (as shown in Fig.4), and the frequency of high-speed clock source is 8MHz (as shown in Fig.5).



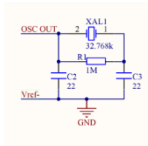


Fig 4.Low speed clock source Fig.5.High speed clock source

4.Reset circuit

The function is to avoid accidental contact. The reset circuit diagram is shown in Fig.6.

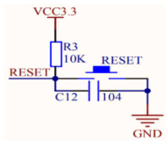


Fig.6. The reset circuit diagram

3.2 Sensor module

In general, the micromouse is composed of five groups of sensors in the front, left, right and left 45 degree oblique direction [2]. A forward sensor is added in this design to prevent the computer mouse from tilting during its travel.

3.3 Motor and encoder circuit

This design adopts a DC motor with encoder which is easy to control and low power consumption. The encoder acts as a speed measurement feedback device. The purpose is to make the microcontroller perceive the walking distance of the micromouse and the accurate steering angle of the micromouse calculated according to the pulse. This design adopts photoelectric weight device, and the schematic diagram of encoder is shown in Figure 7.

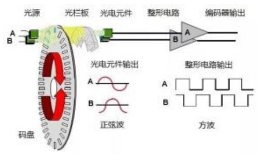


Fig.7.The schematic diagram of encoder

3.4 Voltage detection circuit

This design uses the analog-digital module of STM32 microcontroller to detect the battery voltage. The voltage detection circuit is shown in Fig.8.R9 has a resistance value of 10kΩ, one end is connected to the 3 pin of ams1117, one section is connected to the ADC acquisition pin of microcontroller, and R15 has a resistance value of 5kΩ One section is pulled down to the ground to form a voltage dividing circuit, and the ADC port is connected to the ADC acquisition port of STM32 microcontroller.

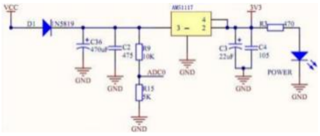


Fig.8. The voltage detection circuit

4. Sotware design

The potential value search algorithm calibrates different potential values into the maze when searching the maze. The closer the cell is to the center of the maze, the smaller the potential value is[3]. Therefore, the potential value of the destination is zero, and the potential value calibration of each unit is shown in Fig.9.

14	13	12	11	10	9	8	7	7	8	9	10	11	12	13	14
13	12	11	10	9	8	7	6	6	7	8	9	10	11	12	13
12	11	10	9	8	7	6	5	5	6	7	8	9	10	11	12
11	10	9	8	7	6	5	4	4	5	6	7	8	9	10	11
10	9	8	7	6	5	4	3	3	4	5	6	7	8	9	10
9	8	7	6	5	4	3	2	2	3	4	5	6	7	8	9
8	7	6	5	4	3	2	1	1	2	3	4	5	6	7	8
7	6	5	4	3	2	1	0	0	1	2	3	4	5	6	7
7	6	5	4	3	2	1	0	0	1	2	3	4	5	6	7
8	7	6	5	4	3	2	1	1	2	3	4	5	6	7	8
9	8	7	6	5	4	3	2	2	3	4	5	6	7	8	9
10	9	8	7	6	5	4	3	3	4	5	6	7	8	9	10
11	10	9	8	7	6	5	4	4	5	6	7	8	9	10	11
12	11	10	9	8	7	6	5	5	6	7	8	9	10	11	12
13	12	11	10	9	8	7	6	6	7	8	9	10	11	12	13
14	13	12	11	10	9	8	7	7	8	9	10	11	12	13	14

Fig.9. the potential value calibration of each unit

The course coordinates of the micromouse direction in which the micromouse moves in the maze. It is mainly defined by the following four directions, namely West, South, East and North. Other variable algorithms are defined to indicate whether the maze has been searched by the micromouse. False if the maze is not searched, otherwise true. Variables marked as prohibited indicate that the maze here is impassable, that is, the maze here is a dead end or part of a dead end or a closed-loop path. The micromouse also identifies a search times variable to mark the number of times the computer mouse searches here. After the micromouse reaches a maze, the count variable of the maze increases by one. The operation strategy (i.e. the strategy adopted by the micromouse to reach a certain position) is independently selected by the micromouse. Possible situations are: there is no access (dead end) ahead, the front can perform left turn, the front can perform right turn, or a combination of several situations. When the micromouse faces a variety of selection strategies in the maze, it will judge in combination with the cell potential value and the number of searches here, and take different execution strategies independently. The micromouse strategy is shown in Fig.10.

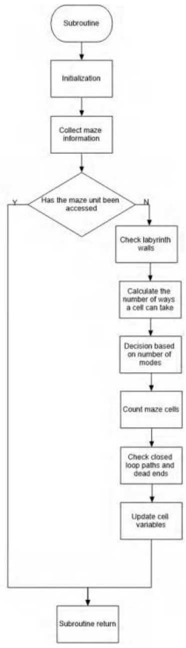


Fig.10. The micromouse strategy

5. Conclusion

When the left and right sides of the micromouse leave 5mm away from the wall (or when it is at an angle of 45 degrees from the horizontal and 20mm away from the wall), the motor can react quickly to prevent the micromouse from accidentally touching the wall during high-speed operation. The distance between the micromouse and the wall is shown in Fig.11.

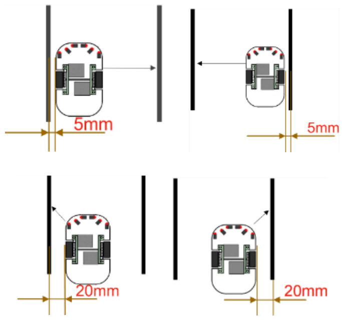


Fig.11. The distance between the micromouse and the wall

When the left and right sensors leave the wall, you can judge whether the micromouse body is inclined according to the two infrared sensors in front. If the micromouse tilts during operation, the data collected by the first two sensors will be larger and smaller. At this time, the main controller of the micromouse will adjust the posture in time, so that the micromouse is always in the center of the maze.

The schematic diagram of intersection detection is shown in Fig.12.

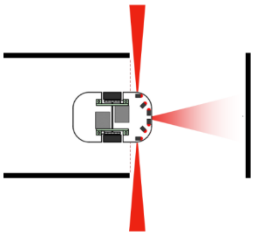


Fig.12. The schematic diagram of intersection detection

References

1. Cao Jingming, Dai Shenghua, Na Lan. Research on Development-oriented IEEE MicroMouse Simulation System. Journal of Physics: Conference Series, 2021, 1802(4). Microsoft Word - CDMMS120.docx (iop.org)
2. Zhang Haoming, PEH Lian Soon, Wang Yingha. Turnings and Its Calibration of Micromouse Based on ARM9 and FPGA. Sensors Transducers, 2014, 168(4).
3. Yining Wang,Han Jiang,Bo Wang. Research and optimization of computer mouse maze intelligent algorithm (in Chinese). Science and technology innovation guide, 2015,12(32):129-130+132.

Authors Introduction

Mr. Hucheng Wang

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.

A Driver Reaction Time Detection System Design

Yuhui Cheng[1*], Mochi Li[2]

[1] *Tianjin University of Science and Technology, 300222, China;*

[2]*Ocean University of China, OUC, Qingdao, China*

*E-mail: * 13642017570@163.com*

www.tust.edu.cn

Abstract

This paper proposes a design scheme of driver reaction time detection system, which collects electroencephalography (EEG) data and reaction time required by the driver to complete corresponding test actions by deploying sensors on the driving simulator. The scheme can be divided into control unit, data acquisition unit and interaction unit. The testee performs corresponding operations according to the instructions issued by the interaction unit, and the system sorts out and saves the collected data. This system is of great significance to study the reaction time of drivers in different states and scenarios.

Keywords: reaction time, driving safety, EGG, driving simulator

1. Introduction

With the rapid development of road traffic, not only has it brought a convenient way of travel, but also the importance of traffic safety has become increasingly prominent. The driver is an important part of the road system, and most of the safety accidents are closely related to the driver's behavior. Driver reaction time is an important indicator for studying driver behavior [1]. Since the reaction time of different drivers is not the same, using a fixed reaction time for research will cause certain errors. Collecting relevant data through hardware equipment and establishing an association model between it and reaction time can more accurately evaluate and predict driver behavior [2].

Driver reaction time refers to the time required for the driver to implement the corresponding decision after receiving the emergency stop signal. It is usually divided into perception time and execution time. At present, there are few devices for studying the driver's reaction time, and most of the data are collected independently.

Based on the above overview, this article proposes a driving assistance device for detecting driver reaction time. The equipment can be divided into three parts: control unit, acquisition unit and interaction unit. Through the EEG acquisition equipment and the hardware modules deployed on the driving simulator, the EEG signals of the driver and the time required to complete the corresponding actions are obtained. The system diagram is shown in Fig.1.

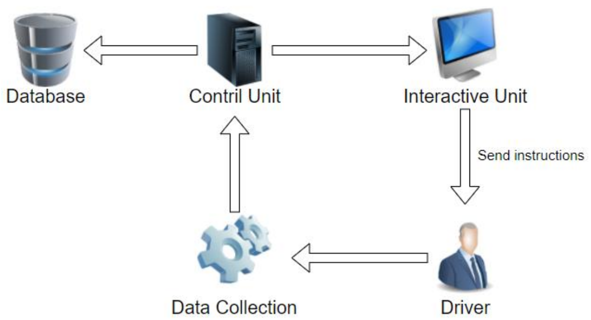


Fig.1 The system diagram

2. System Hardware Structure

In order to accurately obtain the driver's perception time, this design monitors the driver's EEG data. The hardware device deployed on the driving simulator can complete the execution time collection. Since the hardware needs to be deployed on the driving simulator, in order to facilitate installation, all functional modules use Bluetooth for communication. The system hardware structure diagram is shown as in Fig. 2.

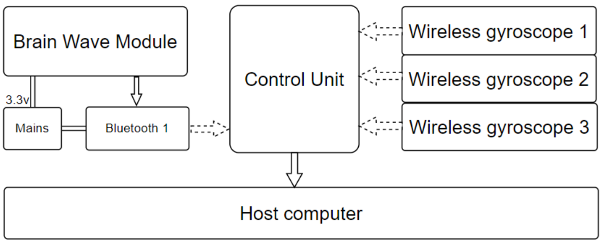


Fig.2. The system hardware structure diagram

2.1. EEG acquisition module

Perception time means that the driver does not act immediately after receiving the signal, but has a period of delay. This period of time is the driver's perception time. In order to accurately obtain the data at this stage, the system uses the TGAM brainwave module to identify and extract the driver's brainwave signal. The TGAM brainwave module is shown in Fig.3.

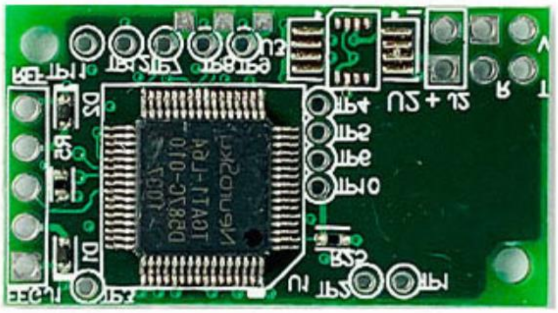


Fig.3. The TGAM brainwave module

The TGAM brainwave module can collect raw brainwave signals, including γ waves, β waves, and α waves. The high sampling rate of 512HZ ensures the timeliness of the collected data. The module collects EEG signals directly through dry electrodes, and sends the data to the control terminal through serial communication. The module has excellent anti-interference performance and has a hardware filter of 3-100HZ. Other hardware parameters are as follows:

- Operating voltage: 3 -3.6V
- Operating current: 15mA
- Electrostatic protection: 4KV contact discharge; 8KV air discharge

When the TGAM EEG module data acquisition module is used, in order to achieve the best effect, the input signal can be preprocessed. For example, to gain signal, consider the polarization voltage and physiological interference to the EEG signal coverage to reduce the previous value gain in the amplifier and increase the input impedance.

2.2. Motion detection module

The execution units detected by this system mainly include brakes, accelerators and steering wheels. Since the detection unit has a certain range of motion, this design will use a wireless nine-axis gyroscope sensor for data collection. The sensor is shown in Fig.4.

Fig.4. Motion detection module

The sensor integrates a high-precision three-axis gyroscope, a three-axis accelerometer, a three-axis Euler angle, a three-axis magnetic field, and a high-performance microprocessor. Through dynamic calculation and Kalman dynamic filtering algorithm, the module can quickly solve the current real-time motion posture. In terms of anti-jamming interference, the door module is integrated with a posture solver, which can accurately calculate the posture in the dynamic process with high stability.

The data protocol of this module is shown in Table 1:

Table 1. Nine-axis wireless gyroscope data format

Num	Data	Meaning
1	0x55	frame header
2	0x53	logo angle package
3	RollL	X-axis angle low byte
4	RollH	X-axis angle high byte
5	PitchL	Y-axis angle low byte
6	PitchH	Y-axis angle high byte
7	YawL	Z-axis angle low byte
8	YawH	Z-axis angle high byte
9	TL	temperature low byte
10	TH	temperature high byte
11	Sum	checksum

Through the above data, the roll angle, pitch angle and yaw angle of the measured object can be calculated.

The calculation formula of the roll angle is as follows:

$$ROLL = ((ROLLH<<8)\ |ROLLL)/32768*180 \quad (1)$$

The pitch angle calculation formula is as follows:

$$PITCH = ((PITCHH<<8)\ |PITCHL)/32768*180 \quad (2)$$

The yaw angle calculation formula is as follows:

$$YAW = ((YAWH<<8)\ |YAWL)/32768*180 \quad (3)$$

2.3. The control unit

The microprocessor model selected in this design is STM32F103VET6. This chip is a 32-bit ARM microcontroller with Cortex-M3 as the core, and the frequency is up to 72MHz, which can meet the real-time requirements of the system. The chip has 3 12-bit ADCs, 3 SPI interfaces, an IIC interface, 5 serial ports, 48KB of SRAM and 256KB of FLASH, which is convenient for wireless communication. It has the following characteristics:

- Operating Voltage 2.0V ~ 3.6V.
- Operating temperature range: -40 ºC ~ 105 ºC.

3. System Software Design

This part mainly includes system software requirement analysis, feasibility analysis and overall design plan introduction.

3.1. Software requirements analysis

In this design, the upper computer software needs to obtain the data collected by the lower computer through the serial port, and decode and analyze it. The processed data can be displayed through the interface software, mainly including EEG signals and action trigger time records. At the same time, the upper computer plays the role of commander and issues test instructions to the driver and the control unit.

3.2. Software feasibility analysis

In order to realize the above-mentioned functions, this design intends to use the C# (C Sharpe) programming language to complete the realization of the upper computer in the Windows 10 environment. C# is an object-oriented developed by Microsoft Corporation, which runs a high-level programming language on the NET Framework. It has the characteristics of speed block, powerful function and data type safety. The development tool to be selected this time is Visual Studio 2019, which is a free version developed by Microsoft for individual users, and its functions are not much different from the professional version.

3.3. Software design plan

According to the demand, the upper computer of this design mainly realizes the functions of setting, data processing and interactive information sending. The system software design structure diagram is shown in Fig.5.

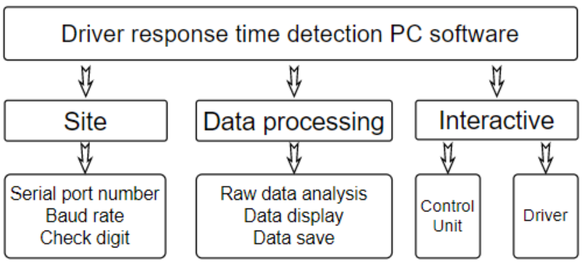


Fig.5. The system software design structure diagram

The setting unit mainly completes the setting and connection of the upper computer software to the communication serial port of the lower computer. The serial port number, baud rate, parity bit, data bit, and stop bit can be set through the software. After setting the parameters, choose to open the serial port to connect to the lower computer. All communication data are displayed in the data receiving and sending area, and two data formats, hexadecimal and decimal, can be selected. The software setting interface is shown in Fig.6.

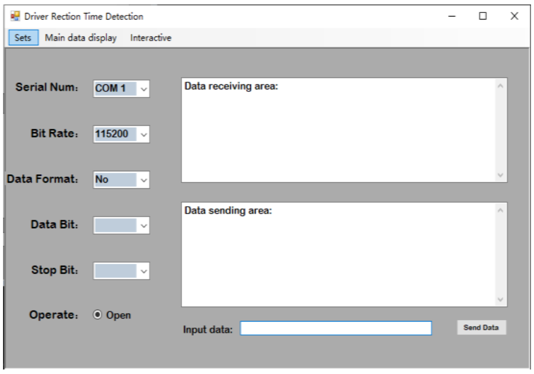


Fig.6. Setting interface

The data display interface is used to display real-time data, including EEG data and the data time collected by the action feedback. The data display interface is shown in Figure 7:

Fig.7. Data display interface

The interactive prompt is completed through the pop-up window of the host computer software, and different instructions, such as steering, braking, etc., are displayed through the pop-up text. When the pop-up window pops up, the software sends a timing instruction to the lower computer to start timing.

4. Conclusion

The driver's reaction time detection system accurately measures the driver's perception time and execution time. Researchers can obtain the reaction time of different populations, scenarios and conditions, which has a positive effect on follow-up research.

References

1. Droździel, Paweł, Tarkowski, et al,Drivers 'reaction time research in the conditions in the real traffic,*Open Engineering*,2020,1:pp.35-47.
2. R. Foong, K. K. Ang , C. Quek, Correlation of reaction time and EEG log bandpower from dry frontal electrodes in a passive fatigue driving simulation experiment, *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*,2017,11: pp.2482-2485
3. Haokang Wen, Qiang Zheng, Design of a High Precision Digital Clock Based on Single Chip Microcomputer, *Proceedings of the 2020 International Conference on Artificial Life and Robotics*, Japan, Jan. 2020, pp. 693-697.

Authors Introduction

Mr. Yuhui Cheng

He received his Bachelor's degree from the Internet of Things Engineering, Tianjin University of Science and Technology, China in 2019. He is currently studying for a master's degree in electronic information at Tianjin University of Science and Technology.

Ms. Mochi Li

She is a 2020 undergraduate's student at Ocean University of China. Her major is aquaculture. Her area is bioinformatics.

A PID Tracking Car Based on STM32

Yande Xiang, Teng Zhang, Shixiang Zhao*, Ling Zhou, Mingjuan Tian, Haoran Gong
School of Electronic Information and Automation, Tianjin University of Science and Technology,300222, China
*E-mail:*1422889849@qq.com*
www.tust.edu.cn

Abstract

Intelligent car is a high-order intelligent product integrating intelligent integrated control, visual terminal, automatic control output and cognitive computing. This paper introduces an innovative intelligent tracking car based on STM32 embedded chip. The vehicle adopts fuzzy PID algorithm to control the vehicle operation, and adopts the innovative grayscale card designed by the team hardware designer. This car fully caters to the global environmental protection trend, the new concept of green development. On the premise of low cost and low power consumption, it also ensures the smooth, smooth and high precision operation of the vehicle.

Keywords: Intelligent integrated control, Visual terminal, STM32 embedded chip, Innovative gray card, Green, Preciseness.

1. Introduction

The intelligent tracking car can detect the surrounding situation autonomously. Using image acquisition and processing technology and gray sensing technology to help themselves to identify the various road conditions in front of the feedback to the control center for judgment and processing and then automatic output instructions. A high degree of automatic control and intelligent.

The knowledge involved includes: the use principle of single chip microcomputer, digital circuit, analog electronic technology, innovative design and construction of hardware, digital signal, image acquisition and processing and artificial intelligence applications. This paper will be detailed in the following three aspects: 1) Parts and circuit design; 2) Body material and manufacturing; 3) Detailed algorithm explanation.

2. Components and Circuit Design

The parts are assembled by combining traditional parts with new parts. The traditional parts used are: STM32f103c8t6 microcontroller, L298n drive module, step-down module with digital tube, bread board, 18650 lithium battery, bull eye wheel and N20 motor. The new parts include self-designed grayscale card (shown in Fig.1).

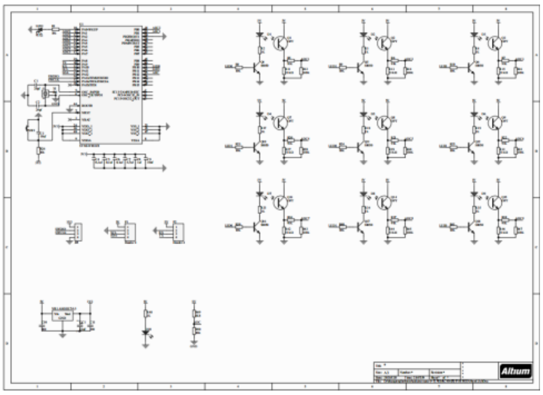


Fig.1. Gray card schematic diagram

This gray card uses 5V power supply, 5V is connected to the positive pole of power supply, GND is connected to the negative pole of power supply. Compared with the traditional gray scale measurement module, its advantages are stable, sensitive, not easy to be interfered by external stray light, easy to use, not easy to damage. As shown below, the two pictures are PCB with grayscale card (shown in Fig. 2 and Fig. 3).

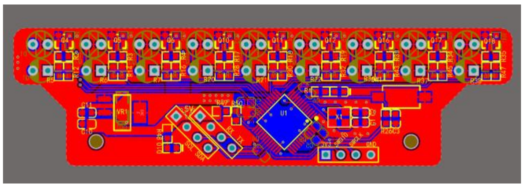


Fig. 2.Gray card PCB

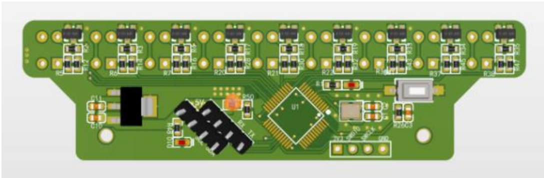


Fig. 3.Gray card PCB (3D)

In terms of circuit design, the car keeps the circuit connection as simple as possible and the power utilization is maximized. It gives full play to its own advantages and suits its actual situation.

The advantages of the STM32f103c8t6 microcontroller are: small package size, low price compared with other microcontroller chips, better performance than 8-bit microcontroller, if the STM32 series of chips will be used, it can be used as an alternative, high cost performance.

The five signal ports A0, A1, A2, A3 and A4 of STM32f103c8t6 are respectively connected to the five signal output ports of the gray card, and the signal output ports of B6, B7, B8 and B9 are connected to the signal receiving port of L298n.

The information recognized by the grayscale sensor[2] is converted into signal and transmitted to STM32f103c8t6 MCU, which is fed back to L298n after processing to realize motor speed regulation.

The power supply outputs two voltages, one of which supplies L298n, one of which reduces the voltage of the transformer module to 3V3 and then inputs it to the gray sensor, thus ensuring the safety of the infrared module and greatly prolonging the service life of the gray card.

3. Body Material and Manufacture

In order to follow the new concept of global green, cater to the global environmental trend. The car adopts green environmental protection material, the whole car material pollution-free. Lightweight and environmentally friendly materials and streamlined design make it even more portable.

Traditional car bodies generally use acrylic panels and other materials. The protective paper should not be broken during transportation, and it may be toxic in some special cases. The car body material using epoxy resin, full environmental protection, reliable.

When manufacturing the body, 3D printing technology is used to empty the reserved position to achieve the body and module tightness. Another point, the body manufacturing light, greatly reduce the vehicle load.

4. Detailed Dlgorithm Explanation

In order to realize the intelligent vehicle tracking highly automatic, high precision, the vehicle adopts fuzzy PID automatic control algorithm. Compared with the traditional PID algorithm, this algorithm has better performance. Fig 4 is the flow chart of traditional PID:

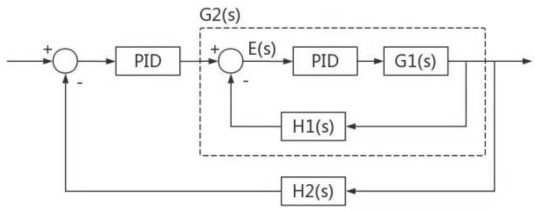


Fig.4.Flow chart of traditional PID

Traditional PID is a classical mathematical model whose principle can be given by mathematical formula[1], as shown below:

$$\mu(t) = K_P e(t) + K_I \int_0^t e(t)dt + K_D \frac{de(t)}{dt} \tag{1}$$

Where KP stands for proportional gain, KI stands for integral gain, and KP stands for differential gain.

But fuzzy control does not need to determine the exact mathematical model of the system, it is an automatic control principle to adapt to rules. She combines control strategy, application variables, fuzzy set theory and integrated algorithm theory. It has three main uses:

●Fuzzification: Transform the exact quantity into the fuzzy quantity (change E into EC).

● Fuzzy judgment: According to the existing linguistic basis (fuzzy control rule table, shown in Table 1), fuzzy judgment is made on the fuzzy quantity obtained [3].

Table 1. The fuzzy rule list

		e						
u		NB	NM	NS	ZE	PS	PM	PB
de	PB	ZE	PS	PS	PM	PM	PB	PB
	PM	NS	ZE	PS	PS	PM	PM	PB
	PS	NS	NS	ZE	PS	PS	PM	PM
	Z	NM	NS	NS	ZE	PS	PS	PM
	NS	NM	NM	NS	NS	ZE	PS	PS
	NM	NB	NM	NM	NS	NS	ZE	PS
	NB	NB	NB	NM	NM	NS	NS	ZE

● Fuzzy decision: The fuzzy quantity of fuzzy decision is converted into the actual precision quantity for output [4].

These three points will make the fuzzy PID run smoothly. The flow chart of fuzzy PID is as follows (shown in Fig. 5).

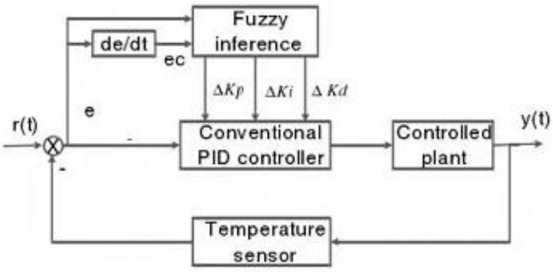


Fig.5 Fuzzy rule list

Compared with the traditional PID algorithm, fuzzy PID algorithm is the embodiment of high precision operation. Aided by visual terminals and cognitive computing, the car as a whole can be elevated to a new level.

5. Conclusion

Intelligent car is a product of high degree of self-control, combining software and hardware cleverly, and widely using a variety of knowledge and material principles.

At present, the car has achieved high precision automatic control, in the future, the car will use cloud data comprehensive computing, using the advanced processing of big data to make the intelligent car running more smooth and concise.

Based on the original knowledge, the vehicle has been successfully tested and achieved good experimental results. Next, we will add intelligent autonomous learning mode, using the advantages of intelligent autonomous learning mode can modify parameters independently, comprehensive improvement and optimization.

In order to ensure high precision operation, the power consumption is reduced as much as possible, so as to achieve power persistence. It is committed to bringing intelligent vehicle tracking to a new height in the field of intelligent vehicle tracking.

References

1. L. Eriksson, T. Oksanen, K. Mikkola, PID controller tuning rules for integrating processes with varying time-delays, *Journal of the Franklin Institute*, 2009, 346(5), pp.470-487.
2. Jing Tao Li, Zhen Feng Li, Xiao Fan Li, et al., A Simple and Efficient Algorithm Design for Improving the Infrared Tracking Accuracy of Smart Cars, *Procedia Computer Science*, 2020, 166, pp.339-343.
3. Juan Carlos Gámez, David García, Antonio González, et al., An approximation to solve regression problems with a genetic fuzzy rule ordinal algorithm, *Applied Soft Computing*, 2019, 78, pp.13-28.
4. Zhinan Hao, Zeshui Xu, HuaZhao, Optimized data manipulation methods for intensive hesitant fuzzy set with applications to decision making, *Information Sciences*, 2021, 580, pp.55-68.

Authors Introduction

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. Teng Zhang

He is currently pursuing his undergraduate degree at the School of Electronic Information and Automation, Tianjin University of Science and Technology

.

Mr. ShiXiang Zhao

He is currently studying for an undergraduate degree at Tianjin University of science and technology.

Mr. Ling Zhou

He is currently studying for a bachelor's degree at Tianjin University of science and technology. He has three software copyrights, one utility model appearance and one utility model patent.

Ms. Mingjuan Tian

She is currently pursuing a bachelor's degree at the School of Electronic Information and Automation, Tianjin University of Science and Technology.

Mr.HaoRan Gon

He is currently pursuing a bachelor's degree at the School of Electronic Information and Automation, Tianjin University of Science and Technology.

Matrix Approach to Current-state Detectability of Discrete-event Systems

Jinliang Wang, Jiawei Wei, Xiaoguang Han*

College of Electronic Information and Automation, Tianjin University of Science and Technology, 300222, China

E-mail: wangjltust@163.com

wjwtust@163.com

* *hxg-allen@163.com*

www.tust.edu.cn

Abstract

In our previous work, a matrix-based framework is proposed to tackle the problem of verifying strong detectability in the context of partially-observed nondeterministic discrete event systems (DESs). Two key concepts, namely, unobservable reach and detector, are redefined therein. Also, the dynamics of a detector, under the frameworks of the Boolean semi-tensor product of matrices, are converted equivalently into an algebraic representation. In this paper, we extend our previous work to other versions of detectability, including strong periodic detectability, weak detectability, and weak periodic detectability. Several necessary and sufficient conditions are derived for verifying aforementioned three types of detectability, respectively. Compared with the existing ones, the proposed methodology is easier to be implemented in software in the sense that it avoids the symbolic manipulations. Finally, an example is given to illustrate the theoretical results.

Keywords: Discrete event systems, state estimation, detectability, semi-tensor product of matrices.

1. Introduction

State estimation is an important and interesting topic in systems and control theory, and has many applications. For instance, in medical systems, we need to know the disease stage of a patient. Yet in remote and distributed systems, one hopes generally that a central station to be able to determine the state of a remote system with limited communications. The state estimation problem has drawn considerable attentions in the context of discrete event systems (DESs)[1-9].

The problem of state estimation of DESs has been investigated widely in terms of detectability of DESs that are based on finite automata models, which says that whether or not one can know exactly the current and subsequent states of original system after a finite delay. The notion of detectability[4] was initially proposed for the deterministic DESs, where four types of detectability, namely, strong (periodic) detectability and weak (periodic) detectability, were defined. Later on, the concepts of other types of detectability came up and have been further extended to other classes of systems by many others[5-8]. For instance, the problem of verifying detectability has been studied in the frameworks of nondeterministic DESs[5,6] and stochastic DESs[7], respectively. Recently, trajectory detectability[8] of DESs has been investigated. Complexity of determining detectability[9] for DESs has been studied. When the original system is not detectable, some approaches[10,11] have been developed to enforce provably detectability by synthesized optimal supervisor (if exists) that restricts the system's behavior in a manner.

In our previous work[6], we developed a matrix-based approach to discuss strong state detectability of nondeterministic DESs from a new angle. A matrix-based verification criterion with polynomial complexity in the size of system for verifying strong detectability is derived. Strong detectability says that whether or not we can know precisely the current and subsequent states of original system after a finite number of delays. However, the goal may be too rigid in practice. In this regard, extending strong detectability to strong periodic detectability and/or weak (periodic) detectability could be necessary. Additionally, developing a novel matrix-based methodology to tackle simultaneously different types of detectability is still interesting. In this paper, we will solve aforementioned these problems.

Notice that, although the study of detectability[4,5] has been considered, our verification criterions are totally different from them. First, the proposed approach is matrix-algebra-based form by using a new tool, called the semi-tensor product (STP) of matrices[12]. While the existing ones are based on design of algorithms. Second, the proposed approach is easier to be implemented in softwares in the sense that all obtained results are numerically tractable instead of graph-based symbolic manipulations.

The remainder of this paper is arranged as follows. Section 2 introduces some basic notations and concepts needed in this paper. Section 3 provides some matrix-based criterions to verify different types of detectability by means of the developed methodology. In Section 4, an example is presented to illustrate the application of the obtained results. Finally, we conclude the paper in Section 5.

2. Preliminaries

2.1. Notations

$|X|$ denotes the cardinality of set X. $\mathbb{M}_{m\times n}$ denotes the set of $m\times n$ real matrices. $M_{(i,j)}$ denotes the (i,j) element of matrix M. $Col_j(M)$ denotes the j-th column of matrix M. $Col(M)$ denotes the set of all columns of matrix M. A matrix $B\in\mathbb{M}_{m\times n}$ is a Boolean matrix if $B_{(i,j)}\in\mathcal{D}=\{0,1\}$, $\forall\ 1\le i\le m$, $1\le j\le n$. We use $B_{m\times n}$ to denote the set of $m\times n$ Boolean matrices. $\delta_n^0:=[0,0,\cdots,0]^T$. $\delta_n^k:=Col_k(I_n)$, where I_n is the identity matrix of dimension n, $1\le k\le n$. $\Delta_n:=\{\delta_n^1,\delta_n^2,\cdots,\delta_n^n\}$; $\tilde{\Delta}_n:=\{\delta_n^1,\delta_n^2,\cdots,\delta_n^n\}$. $L\in\mathbb{M}_{m\times n}$ is a generalized logical matrix if $Col(L)\subseteq\tilde{\Delta}_m$. We denote the set of $m\times n$ generalized logical matrix by $\mathcal{L}_{m\times n}$. For brevity, $L\in\mathcal{L}_{m\times n}$ is denoted as $L=\delta_m[i_1,\ i_2,\cdots,i_n]$, $i_k\in\{0,1,\cdots,m\}, 1\le k\le n$.

2.2. Semi-tensor product (STP) of matrices

Definition 2.1[12] *Let $A\in\mathbb{M}_{m\times n}$ and $B\in\mathbb{M}_{p\times q}$. The STP of A and B is defined as*

$$A\ltimes B=\left(A\otimes I_{t/n}\right)\left(B\otimes I_{t/p}\right), \tag{1}$$

where t denotes the least common multiple of n and p, i.e., t = lcm(n, p); $\otimes$ is the Kronecker product.

Remark 2.1 *When $n=p$, $A\ltimes B=AB$. Hence, the STP is a generalization of the standard matrix product. We mostly omit the symbol "$\ltimes$" hereinafter.*

Lemma 2.1[12] *Let $X\in\mathbb{R}^m$ and $Y\in\mathbb{R}^n$ be two column vectors. Then*

$$W_{[m,n]}XY=YX,\quad W_{[m,n]}YX=XY. \tag{2}$$

where $W_{[m,n]}=[\delta_n^1\delta_m^1,\cdots,\delta_n^n\delta_m^1,\cdots,\delta_n^1\delta_m^m,\cdots,\delta_n^n\delta_m^m]$.

Lemma 2.2 [12] *Let $\delta_{n_1}^{i_1}\delta_{n_2}^{i_2}=\delta_{n_1n_2}^{i_{12}}$, then we have $i_{12}=(i_1-1)n_2+i_2$.*

2.3. Boolean algebra

Here we introduce some notations from the binary algebra of binary matrices (or called Boolean matrices) that will be used later on.

Definition 2.2[13] *Assume that $A=(a_{ij})_{m\times n}\in\mathcal{B}_{m\times n}$, $B=(b_{ij})_{m\times n}\in\mathcal{B}_{m\times n}$. The Boolean addition of A and B is defined as*

$$A\times_{\mathcal{B}}B:=(a_{ij}\vee b_{ij})_{m\times n}\in\mathcal{B}_{m\times n}, \tag{3}$$

where the symbol "$\vee$" is the logical operators OR.

Definition 2.3[13] *Assume that $A=(a_{ij})_{m\times n}\in\mathcal{B}_{m\times n}$, $B=(b_{ij})_{n\times s}\in\mathcal{B}_{n\times s}$. The Boolean product of A and B is defined as*

$$A\times_{\mathcal{B}}B:=C=(c_{ij})_{m\times s}\in\mathcal{B}_{m\times s}, \tag{4}$$

where $c_{ij}=\vee_{k=1}^{n}\left(a_{ik}\wedge b_{kj}\right)$; the symbols "$\vee$" and "$\wedge$" denote the logical operators OR and AND, respectively.

Definition 2.4 *Assume that $A=(a_{ij})_{m\times n}\in\mathcal{B}_{m\times n}$, $B=(b_{ij})_{p\times q}\in\mathcal{B}_{p\times q}$. The Boolean semi-tensor product of A and B is defined as*

$$A\ltimes_{\mathcal{B}}B:=\left(A\otimes I_{t/n}\right)_{\mathcal{B}}\left(B\otimes I_{t/p}\right), \tag{5}$$

where t = lcm(n, p).

Remark 2.2 *From now on, the following all matrix product (resp., matrix addition) is assumed to be the*

Boolean semi-tensor product (resp., Boolean addition) and the symbol "$\ltimes_{\mathcal{B}}$" (resp., $+_{\mathcal{B}}$) will also be omitted hereinafter when there is no danger of confusion.

2.4. System model

The discrete event system (DES) of interest is modeled as a nondeterministic finite automaton (NFA) that is a five-tuple $G = (X, \Sigma, \delta, X_0, X_m)$, where X is a finite set of states, Σ is a finite set of events, $X_0 \subseteq X$ is the set of initial states, $X_m \subseteq X$ is the set of marked states (or accepted states), $\delta : X \times \Sigma \to 2^X$ is the partial transition function, which describes the system dynamics: given states $x, y \in X$ and an event $\sigma \in \Sigma$, $y \in \delta(x,\sigma)$ means that there is a transition labeled by event σ from state x to state y. Note that $\delta(x,\sigma)$ is undefined when the event σ cannot be executed at state x. We use $\delta(x,\sigma)!$ to denote that $\delta(x,\sigma)$ is well-defined. Obviously, the transition function can be extended to $\delta : X \times \Sigma^* \to 2^X$ in the usual manner, where Σ^* denotes the set of finite strings on the alphabet Σ, including the empty string ϵ.

For brevity, we assume that $G = (X, \Sigma, \delta, X_0, X_m)$ is deadlock free (also called alive), i.e., for each state $x \in X$, there is at least a corresponding event $\sigma \in \Sigma$ such that $\delta(x,\sigma)!$. It should be pointed out that this assumption is without essential loss of generality, since it can be relaxed by adding observable self-loops at terminal states.

When system G is partially observed, its event set is partitioned into two disjoint parts: the observable part Σ_o and the unobservable part Σ_{uo}, i.e., $\Sigma_o \cup \Sigma_{uo} = \Sigma$ and $\Sigma_o \cap \Sigma_{uo} = \emptyset$. The natural projection $P\colon \Sigma^* \to \Sigma_o^*$ is defined by

$$P(\epsilon) = \epsilon \text{ and } P(s\sigma) = \begin{cases} P(s)\sigma, & if \sigma \in \Sigma_o; \\ P(s), & if \sigma \in \Sigma_{uo}. \end{cases} \tag{6}$$

3. Detectability of partially-observed DESs

3.1. Problem statement

In the paper, so-called state estimation is based on observations of some events and states. More explicitly, the event observation is described by the projection $P\colon \Sigma^* \to \Sigma_o^*$, while the state observation is described by the output map $h : X \to Y$ where Y denotes a finite output set. In this regard, a partially-observed nondeterministic DES with the event and state observations can be described as follows.

$$G_o = (G, P, h, \Sigma_o, Y), \tag{7}$$

where $G = (X, \Sigma, \delta, X_0, X_m)$.

Consequently, the problem of state estimation of partially-observed DESs, under the framework of detectability, can be formalized as follows. Given a partially-observed nondeterministic DES (7), we do not know the set of initial states of system G_o, while we have partial event observations (i.e., $\Sigma_o \subset \Sigma$) and some state observations (i.e., $Y \neq \emptyset$). Whether we can know exactly the current and subsequent states of system (7) after a finite number of observations. Formally, we give the concepts of three types of detectability of system (7) below.

Definition 3.1 *A partially-observed nondeterministic DES (7) is said to be weakly detectable, if its current and subsequent states can be precisely determined by some admissible input-output strings after a finite number of observations.*

Definition 3.2 *A partially-observed nondeterministic DES (7) is called strongly (resp., weakly) periodically detectable, if its current state can be periodically determined by all (resp., some) admissible input-output strings after a finite number of observations.*

3.2. Algebraic expression of detector

To investigate aforesaid three types of detectability, our previous work[6] defines the unobservable reach of a state $x \in X$ for partially-observed nondeterministic DES (7).

Definition 3.3 *Given a partially-observed nondeterministic DES (7), the unobservable reach of state $x \in X$, denoted by $UR(x)$, is defined as*

$$UR(x) = \{x\} \cup \begin{array}{c}\{\tilde{x} \in X \mid \exists e \in \Sigma_{uo}^* \; s.t. \; \tilde{x} \in \delta(x, e) \\ and \; h(x) = h(\tilde{x})\}\end{array} \tag{8}$$

Intuitively, $UR(x)$ represents the set of all states that are reachable from x through unobservable strings and they have the same output as state x.

A detector[6] that is a deterministic finite automaton, denoted by G_{det}, is constructed for partially-observed nondeterministic DES (7). Formally,

$$G_{det} = (\tilde{X}, \Sigma_o \cup \{\emptyset\} \times Y, \delta_{det}, \tilde{x}_0), \tag{9}$$

where the state set is $\tilde{X} \subseteq 2^X$, $\Sigma_o \cup \{\emptyset\} \times Y$ denotes the input-output set, δ_{det} is partial transition function, $\tilde{x}_0 = X$ is the initial state.

To obtain the dynamics of detector (9), let us consider the partially-observed nondeterministic DES (7), where $X = \{x_1, x_2, \cdots, x_n\}$; the set of events, without loss of generality, is $\Sigma = \Sigma_o \cup \Sigma_{uo}$ with $\Sigma_o = \{e_1, e_2, \cdots, e_{s-1}\}$ and $\Sigma_{uo} = \{e_s, e_{s+1}, \cdots, e_m\}$; the output set is $Y =$

$\{y_1, y_2, \cdots, y_q\}$. Identify x_i (resp., e_{j_1}, e_{j_2}, y_k) with δ_n^i (resp., $\delta_s^{j_1}$, δ_s^s, δ_q^k) for simplicity, $(1 \le i \le n)$ (resp., $1 \le j_1 \le s-1$, $s \le j_2 \le m$, $1 \le k \le q$). we call δ_n^i, $\delta_s^j(j = j_1, s)$ and δ_q^k are the vector forms of x_i, e_j $(j = j_1, j_2)$ and y_k, respectively. Thus $X \sim \Delta_n$, $\Sigma \sim \Delta_s$ and $Y \sim \Delta_q$. To this end, an admissible input-output pair (δ_s^j, δ_q^k) can be identified with δ_{sq}^p by means of the formula $\delta_{sq}^p = \delta_s^j \ltimes \delta_q^k$ given in Lemma 2.2.

Next, we construct a matrix $F_p \in \mathcal{B}_{n\times n}$, called input-output transition structure matrix associated with the input-output pair (δ_s^j, δ_q^k), as follows.

$$F_{p(i,t)} = \begin{cases} 1, \delta_n^a \in \delta(\delta_n^t, \delta_s^j), \delta_n^i \in UR(\delta_n^a), h(\delta_n^a) = \delta_q^k; \\ 0, otherwise. \end{cases} \quad (10)$$

Further, the input-output transition structure matrix (abbreviated as IOTSM) associated with all admissible input-output pairs is defined as

$$F = [F_1, F_2, \cdots, F_{sq}] \in \mathcal{B}_{n\times nsq}. \quad (11)$$

Proposition 3.1 *Given a partially-observed nondeterministic DES (7), then the dynamics of detector (9) can be equivalently described by the following algebraic equation*

$$\tilde{x}(t+1) = Fu(t)y(t)\,\tilde{x}(t), \quad (12)$$

where the matrix F defined in (11) is called the IOTSM of detector (9), $\tilde{x}(1) = [1,1,\cdots,1]^{\mathrm{T}}$ is the initial state of detector (9), $\tilde{x}(t) = (x_1(\mathrm{t}), x_2(\mathrm{t}), \cdots, x_n(\mathrm{t}))^{\mathrm{T}}$ is the vector form of state of detector (9) at step t, whose i_τ-th $(\tau = 1,2,\cdots,k)$ entry equals to 1 means that the state of detector (9) at step t is $\{x_{i_1}, x_{i_2}, \cdots, x_{i_k}\}$, $u(t)y(t) \in \Delta_{sq}$ is the vector form of input-output pair at step t.

Using Lemma 2.1, we have

$$\tilde{x}(t+1) = FW_{[n,sq]}\tilde{x}(t)u(t)y(t), \quad (13)$$

Define matrix $\tilde{F}^{[t]} =: (FW_{[n,sq]})^{[t]}\tilde{x}(1)$, then Eq.(13) becomes

$$\tilde{x}(t+1) = \tilde{F}^{[t]} \ltimes_{j=1}^{t} u(j)y(j). \quad (14)$$

3.3. Verification of detectability

Now, we further develop a matrix-based methodology in terms of Eq. (13) to verify three types of detectability for partially-observed nondeterministic DES (7). To this end, we need the following some preliminaries.

Using Theorem 2 existed in other paper[14] and Eq. (13) and/or (14), we can obtain easily the state set $\tilde{X}$ of detector (9), and denote by $\tilde{X} = \{z_1, z_2, \cdots, z_n\}$. Identifying $z_i \sim \eta_i$, where η_i is the vector form of state z_i, $1 \le i \le N$. Further, we construct that the vector $A_p \in \mathcal{B}_{N\times 1}(1 \le p \le N)$ is as follows.

$$A_{p(i,j)} = \begin{cases} 1, \eta_i \in Col(FW_{[n,sq]})\eta_j, 1 \le i,j \le N; \\ 0, otherwise. \end{cases}$$

(15)

Intuitively, the matrix $A = [A_1, A_2, \cdots, A_N] \in \mathcal{B}_{N\times N}$ is the adjacency matrix of the state transition diagram of the detector (9).

On the other hand, we call a state $z \in \tilde{X}$ a single state if $|z| = 1$ and denote by $X_{single} = \{z \in \tilde{X} \mid |z| = 1\}$. Let $X_{single} = \{z_{i_1}, z_{i_2}, \cdots, z_{i_w}\}$. $\nu = \{i_1, i_2, \cdots, i_w\}$ is said to be the subscript set of X_{single}.

The following results provide the matrix-based criteria of verifying aforesaid four types of detectability for partially-observed nondeterministic DES (7).

Theorem 3.1 *A partially-observed nondeterministic DES (7) is weakly detectable if and only if*

$$\{\sigma \mid (\textstyle\sum_{k=1}^{\omega}(P^TAP)^K)_{(\sigma,\sigma)} = 1\} \neq \emptyset \quad (16)$$

where $P = \delta_N[i_1, i_2, \cdots, i_w]$.

Proof. Since $\nu = \{i_1, i_2, \cdots, i_w\}$ is the subscript set of the set of single states X_{single}, then (16) holds if and only if there exists at least a loop in G_{det} in which all states belongs to X_{single}. This means that there exists at least an admissible infinite input-output string by which the current and subsequent states of system (7) can be know exactly after a finite delay. Therefore, by Definition 3.1, system (7) is weakly detectable if and only if (16) holds.

Theorem 3.2 *A partially-observed nondeterministic DES (7) is strongly periodically detectable if and only if*

$$\{\sigma \mid (\textstyle\sum_{k=1}^{N-\omega}(Q^TAQ)^K)_{(\sigma,\sigma)} = 1\} = \emptyset \quad (17)$$

where $Q = \delta_N[1, \cdots, i_1 - 1, i_1 + 1, \cdots i_\omega - 1, i_\omega + 1, \cdots, N]$.

Proof. We know that (17) holds if and only if there is no loop of all states belong to $\tilde{X}/X_{single}$ in G_{det}. This implies that we determine periodically the current state of system (7) for all admissible input-output strings after a finite number of delays. Consequently, by Definition 3.2, (17) holds if and only if system (7) is strongly periodically detectable.

Theorem 3.3 *A partially-observed nondeterministic DES (7) is weakly periodically detectable if and only if*

$$\{\sigma \mid (\textstyle\sum_{k=1}^{N}A^K)_{(\sigma,\sigma)} = 1\} \cap \nu \neq \emptyset \quad (18)$$

Proof. The proof of Theorem 3.3 follow directly from Definition 3.2 and Theorem 3.1, here we omit it.

Remark 3.1 Consider a partially-observed nondeterministic DES (8) with n state nodes, $s-1$ observable event nodes and q output nodes. In other paper[9], author proved that there exists no polynomial algorithm for verifying weak (periodic) detectability unless $P = PSPACE$. In our paper, since the size of matrix F in (12) is $n \times nsq$, the complexity of constructing the matrix-based detector (12) is $O(n^2 sq)$, which is polynomial with respect to the size of system (7). On the other hand, in the worst case, the cardinality of the state set $\tilde{X}$ of detector G_{det} is 2^n. Therefore, the complexity of constructing matrix A is $O(2^n \times 2^n)$, which is exponential with respect to the size $\tilde{X}$. Overall, the total complexity of implementing Theorems 3.1-3.3 to verify aforementioned three types of detectability is $O(n^2 sq + 2^{2n})$.

Remark 3.2 A Matlab toolbox on the numerical computation of the STP of matrices has been created at *http://lsc.amss.ac.cn/~dcheng/stp/STP .zip*. In this paper, the implementation of the following an example is based on this Matlab toolbox.

4. Illustrative example

4.1. Example 4.1

Consider a partially-observed DES shown in Fig.1, where $e_3 \in \Sigma_{uo}$.

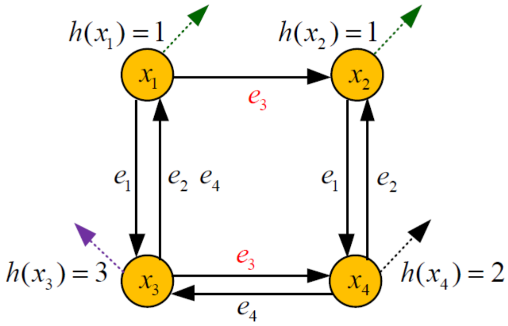


Fig.1 A partially-observed DES with output observations

Identifying $x_i \sim \delta_4^i \ (1 \le i \le 4)$, $e_j \sim \delta_4^j \ (1 \le j \le 4)$, $k \sim \delta_3^k \ (1 \le k \le 3)$. Using Eq. (8), we obtain that $UR(x_i) = \{x_i\}$, $i = 1,2,3,4$. By Proposition 3.1, we obtain that the dynamics of detector G_{det}

$$\tilde{x}(t+1) = Fu(t)y(t)\,\tilde{x}(t). \qquad (19)$$

Using Lemma 2.1, Eq.(19) becomes

$$\tilde{x}(t+1) = FW_{[4,12]}\tilde{x}(t)u(t)y(t), \qquad (20)$$

where $F = [F_1, F_2, \cdots, F_{11}, F_{12}]$ with $F_i = 0_{4\times 4} (i = 1,5,6,7,9,11)$, $F_2 = \delta_4[0,4,0,0]$, $F_3 = \delta_4[3,0,0,0]$, $F_8 = \delta_4[0,0,4,0]$, $F_{12} = \delta_4[0,0,0,3]$,

$F_4 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, $F_{10} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. We can obtain further that $\tilde{X} = \{z_1, z_2, z_3, z_4, z_5\}$, where $z_1 = \{x_1, x_2, x_3, x_4\}$, $z_2 = x_2$, $z_3 = x_3$, $z_4 = x_4$, $z_5 = \{x_1, x_2\}$, $X_{single} = \{z_2, z_3, z_4\}$, $\nu = \{2,3,4\}$,

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \end{bmatrix},$$

$\{\sigma | (\sum_{k=1}^{5} A^K)_{(\sigma,\sigma)} = 1\} = \{2,3,4,5\} \cap \nu \neq \emptyset$,

$\{\sigma | (\sum_{k=1}^{3} (P^T AP)^K)_{(\sigma,\sigma)} = 1\} = \{1,2,3\} \neq \emptyset, where\ P = \delta_5[2,3,4]$,

$\{\sigma | (\sum_{k=1}^{2} (Q^T AQ)^K)_{(\sigma,\sigma)} = 1\} = \emptyset, where\ Q = \delta_5[1,5]$.

By Theorems 3.1-3.3, the system shown in Fig.1 is strongly periodic detectable and weakly (periodic) detectable.

5. Conclusion

In this paper, we developed a matrix-based methodology to verify various types of detectability for partially-observed nondeterministic DESs. By resorting to our previous work [6], several matrix-based criterions for verifying three types of detectability were derived. These verification criterions all of are based on closed-forms

The proposed methodology in this paper is only viewed as a start of the related issues for partially-observed DESs, it may be provide a new theoretical framework for verifying and synthesizing of other types of system-theoretic property. For instance, extending the developed methodology to trajectory detectability[8] for partially-observed DESs modeled by finite automata (resp., Petri nets[15]) is an interesting direction.

Acknowledgements

This work is supported in part by the National Natural Science Foundation of China under Grant 61903274 and the Tianjin Natural Science Foundation of China under Grant 18JCQNJC74000.

References

1. Ramadge Peter J., Observability of discrete event systems, *Proceedings of 25th IEEE Conference on Decision and Control*, Athens, Greece, Dec. 1986, pp. 1108–1112.
2. Ozveren, C. and A. Willsky, Observability of discrete event dynamic systems, *IEEE Transactions on Automatic Control*, 1990, 35(7): pp. 797-806.
3. Feng Lin, Diagnosability of discrete event systems and its applications, *Discrete Event Dynamic Systems*, 1994, 4(2): pp. 197-212.

4. Shaolong Shu, Feng Lin, Hao Ying, Detectability of discrete event systems, *IEEE Transactions on Automatic Control*, 2007, 52(12): pp. 2356-2359.
5. Shaolong Shu, Feng Lin, Generalized detectability for discrete event systems, *Systems Control Letters*, 2011, 60(5): pp.310-317.
6. Xiaoguang Han, Zengqiang Chen, Jiemei Zhao, Matrix approach to detectability of discrete-event systems under partial observation, *Proceedings of 13th IEEE Conference on Automation Science and Engineering*, Xi'an, China, Aug. 2017, pp.187-192.
7. Christoforos Keroglou, Christoforos N. Hadjicostis, Detectability in stochastic discrete event systems, *Systems Control Letters*, 2015, 84(2): pp.21-26.
8. Xiang Yin, Zhaojian Li, Weilin Wang, Trajectory detectability of discrete-event systems, *Systems Control Letters*, 2018, 119: pp.101-107.
9. Kuize Zhang, The problem of determining the weak (periodic) detectability of discrete event systems is PSPACE-complete, *Automatica*, 2017, 81: pp. 217-220.
10. Shaolong Shu, Feng Lin, Enforcing detectability in controlled discrete event systems, *IEEE Transactions on Automatic Control*, 2013, 58(8): pp. 2125-2130.
11. Xiang Yin, Stéphane Lafortune, A uniform approach for synthesizing property-enforcing supervisors for partially observed discrete-event systems, *IEEE Transactions on Automatic Control*, 2016, 61(8): pp. 2140-2154.
12. Daizhan Cheng, *An Introduction to Semi-tensor Product of Matrices and Its Applications*, Singapore: World Scientific, 2012.
13. K Rosen, *Discrete Mathematics and Its Applications*, McGraw-Hill, 2003.
14. Xiangru Xu, Yiguang Hong, Matrix expression and reachability analysis of finite automata, *Journal of Control Theory and Applications*, 2012, 10(2): pp.210-215.
15. Xiang Yin, Stéphane Lafortune, On the decidability and complexity of diagnosability for labeled Petri nets, *IEEE Transactions on Automatic Control*, 2017, 62(11): pp. 5931-5938.

Authors Introduction

Mr. Jinliang Wang

He received the B.S. degree in electrical engineering and intelligent control from Shandong University of Science and Technology, China, in 2019. He is currently pursuing the M.S. degree in instrument science and technology in Tianjin University of Science and Technology, China. His research interests include discrete event systems, supervisory control, and state estimation.

Ms. Jiawei Wei

She received the B.S. degree in control science and engineering from Tianjin University of Science and Technology, China, in 2000. She is currently pursuing the M.S. degree in control science and engineering in Tianjin University of Science and Technology, China. Her research interests include discrete-event systems, cyber-security and privacy, and formal methods.

Dr. Xiaoguang Han

He received the B.S. and M.S. degrees in mathematics from Harbin Normal University, China, and Harbin Engineering University, China, in 2007 and 2009, respectively, and the Ph.D. degree in control science and engineering, from Nankai University, Tianjin, China, in 2017. In 2017, he joined Tianjin University of Science and Technology, China, where he is an Associate Professor in the College of Electronic Information and Automation. His research interests include discrete-event systems, cyber-security and privacy, formal methods, and Boolean control networks.

Hardware Circuit Design Of Tracking Car Based On K60

Peng Jia*, Yande Xiang
Tianjin University of Science and Technology, China
*E-mail:*1563337147@qq.com*
www.tust.edu.cn

Abstract

The development of the robot industry is booming. For many repetitive and difficult jobs, using industrial robots to complete them can greatly save money, manpower and time costs. Industrial tracking vehicle has been applied in many fields and has shown good results. Taking the industrial automated guided vehicle as the model, this paper designs a hardware system of intelligent tracking vehicle based on k60. The main work includes hardware circuit scheme design, component selection, schematic diagram and PCB design, PCB board welding.

Keywords: AGV, k60, path search, hardware circuit design

1. Introduction

With the development of information technology and economic globalization, the traditional logistics transportation is far from meeting the fierce competitive logistics market. The development and innovation of modern logistics technology has gradually become a key factor affecting the development of national economy. Therefore, optimizing the enterprise transportation mode and designing the intelligent car and its control scheme for logistics demand will be the focus in the field of enterprise intelligent transportation at home and abroad.

Since the discovery of the world's first New Coronavirus case in December 2019, the epidemic has spread to many industries, and manufacturing and logistics industry is the first to bear the brunt. However, challenges and opportunities coexist. Intelligent manufacturing and UAV distribution show strong growth potential during the epidemic period. Jingdong warehouse is a good example. Among them, automated guided vehicle is widely used to carry and sort goods, which greatly saves the material and human resources required by factories and warehouses for daily production and storage activities.

2. Hardware Circuit Design

2.1 Overall design idea of the system

According to the application requirements of intelligent tracking car in industry, the main functions of hardware are divided into tracking, driving, power supply and control circuit, or human-computer interaction. As shown in Fig.1, it is an idea for the hardware design framework of a simple tracking car:

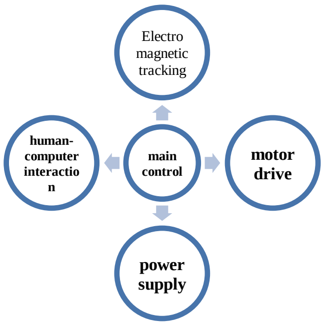


Fig.1. Overall design idea of the system

2.2 Main control chip

MK60DN512ZVLQ10 is used to be the external MCU, which is the Cortex-M4 ARM core and has a 144-pin. The chip has 512KB flash memory and 128KB RAM. And it has more than 100 general-purpose IO ports, multiple timers and serial ports, which fully meet the design of the required pin and memory requirements. The design of the main control chip is shown in Fig.2.

Fig.2. MK60DN512ZVLQ10 chip

2.3 Electromagnetic tracking

The functions of intelligent tracking car in industrial applications are mostly handling and patrol inspection. The electromagnetic tracking method is adopted here. The advantage of this method is that the camera can go out independently to do other identification and monitoring work, saving space, budget and accurate tracking.

The experimental environment is set indoors. The enamelled wire laid on the site is connected with an alternating current of 20kHz and 100mA. The induction module is composed of 10MH I-shaped inductor and 6.8nf capacitor. The I-shaped inductor is a cylinder, occupies a small space and can sensitively detect the change of the surrounding magnetic field.

According to Biot Savart law, the induced magnetic field distribution excited by a straight wire with alternating current is a group of concentric circles with the straight wire as the axis. The magnetic field intensity at all points on the circle with the same radius is the same, and will decrease inversely with the increase of radius. Therefore, five inductors are set in the induction part of the tracking car and installed at the front end of the tracking car according to the way[1] shown in Fig.3.

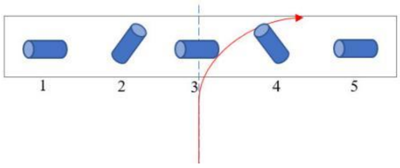


Fig.3. Inductance installation diagram

Among them, inductors 1, 3 and 5 are a group to detect the deviation angle of the forward direction. Inductors 2 and 4 are used to detect the curve.

The AC voltage signal collected by the electromagnetic detection module is amplified and rectified and input to MCU. The amplifier circuit adopts LM386 chip, which is an audio set success rate amplifier chip with low power consumption and adjustable gain. The LM386 can be powered by a single 5V. The gain can be adjusted by the external capacitance and resistance between pins 1 and 8, up to 200, and the total harmonic distortion of the output signal is small. The rectifier part adopts board bridge rectifier, parallel capacitor and resistance to transmit the voltage signal on the resistance to MCU. Fig.4 is a part of the schematic diagram of the electromagnetic amplification circuit. The electromagnetic detection part is composed of the same 5 parts in total.

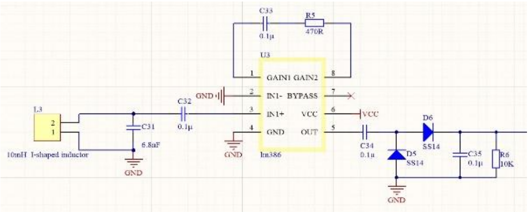


Fig.4. Part of schematic diagram of electromagnetic amplification circuit

2.4 Human-computer interaction

This paper uses three methods to output information: active buzzer, a group of color LEDs and an OLED screen. At the same time, the dial switch and key are used to input information.

There is a simple oscillation circuit inside the active buzzer, which can make sound only by connecting the DC voltage signal. Because the way of displaying information is very intuitive and obvious, it is very suitable for alarm. Since the default solder resist of PCB is green, the LED colors are white, red and blue. This led is very suitable for displaying and adjusting some parameters, especially information related to the "binary" value of threshold. The reserved OLED display interface is used for external OLED to display some key information, such as voltage signal data returned by electromagnetic detection circuit, binary image collected by camera, motor speed collected by encoder, PID parameters, etc., which is convenient for debugging tracking car. Fig.5 is a partial circuit for displaying output information. Fig.6 shows the peripheral interface for MCU to obtain information.

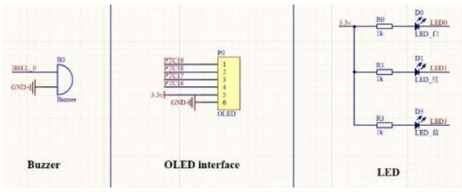


Fig.5. Design of display and output information part

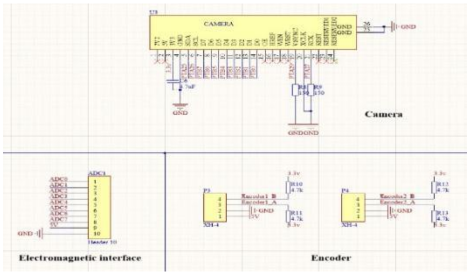


Fig.6. Peripheral Interface

2.5 Motor drive

(1) H-bridge drive principle

Because the DC motor rotates in different directions according to the polarity of the applied DC voltage. Therefore, in order to realize differential turning, H-bridge circuit needs to be used to drive the motor[2]. Fig.7 is a schematic diagram of H-bridge.

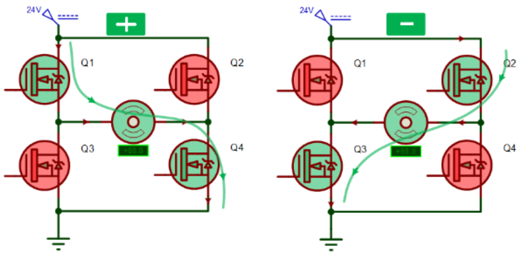


Fig.7. H-bridge

As shown in Fig.7, the circuit is a driving circuit composed of four N-CHANNEL POWER MOS tubes. It is assumed that the left side of the motor is (+) terminal and the right side is (-) terminal. If Q1 = Q4 = on and Q2 = Q3 = off, the (+) terminal of the motor is connected to the positive pole of the power supply and the (-) terminal is connected to GND, so the motor runs in the positive direction at this time. Conversely, if Q2 = Q3 = on and Q1 = Q4 = off, the (+) terminal of the motor is connected to GND and the (-) terminal is connected to the positive pole of the power supply. At this time, the motor runs in the opposite direction. Because the MOS tube and motor that make up the circuit are placed like the letter "H", it is called H-bridge drive circuit.

(2) H-bridge drive design

The scheme design of H bridge is shown in Fig.8.

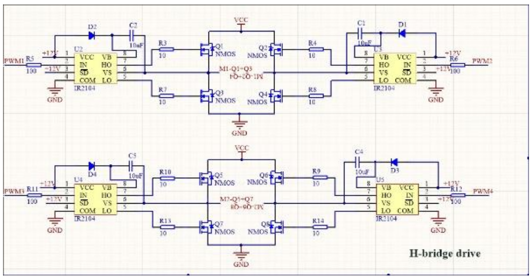


Fig.8. H-bridge scheme design

In the application of H-bridge, the driving performance of N-channel power MOSFET is better than that of P-channel power MOSFET. This design adopts IRLR7843, and its internal integrated diode can be used to protect MOSFET. When the DS two-stage voltage is too high, the bulk diode will be broken down to protect the MOSFET. A reverse path can also be provided for overcurrent caused by inductive load freewheeling or parasitic parameters.

In addition to the MOS transistor, the driver chip is also required to drive the bridge. Here, the chip IR2104 is selected. Fig.9 is an input / output timing diagram of IR2104.

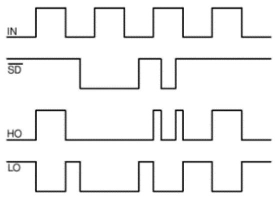


Fig.9. IR2104 I / O sequence diagram

When pin 3 of the chip inputs a high level, the enable is effective and the chip starts to work. When the enable is effective, the high side output HO is consistent with the PWM input by pin 2, and the low side output LO is opposite to HO. Due to the different positions of the motor relative to the MOSFET of the upper bridge arm and the lower bridge arm, the conduction condition of the MOSFET is VGS > Vth. This leads to a large voltage required for the gate to ground if the upper bridge arm MOSFET is to be turned on[3]. At this time, the bootstrap loop is introduced, which makes use of the unidirectional conductivity of the diode and the characteristics that the voltage at both ends of the capacitor can not change suddenly. Combined with the design structure of the driver chip, the upper bridge arm MOSFET can be turned on. Two opposite PWM signals are used to control the two drive chips to make the two half bridge circuits in opposite directions. Then they are combined to form a complete H-bridge circuit.

2.6 Power supply

The tracking car plans to use the most commonly used power supply on the market, which is mainly composed of two 18650 batteries. The rated voltage is 7.4V and the capacity is about 2000mAh. According to the power supply required on the board, the following schemes are designed:

- The battery is used to directly supply power to the motor.
- 7805 chip is used to output 5V voltage to supply power to encoder and operational amplifier.
- AMS1117-3.3 chip is used to output 3.3V voltage to supply power to MCU, camera, OLED and other components.
- SX1308 chip is used to output 12V voltage to supply power to H-bridge drive chip.
- AMS1086-ADJ chip is used to output 6V to supply power to the front wheel steering gear.

3. PCB design

The EDA software used in this design is Altium Designer. After designing the PCB, package the manufacturing documents and give them to the manufacturer. Then purchase components for welding. Fig.10 shows the designed PCB and Fig.11 shows the welded PCB.

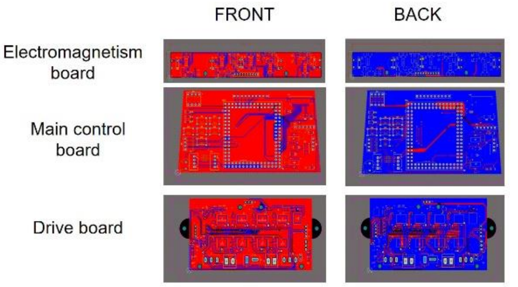


Fig.10. Designed PCB

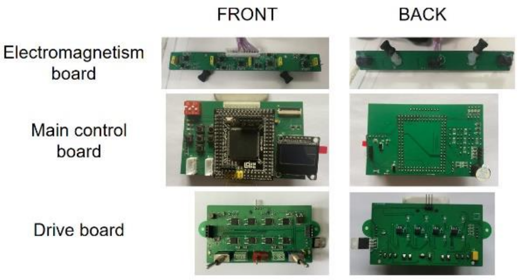


Fig.11. Welded PCB

4. Conclusion

For the design of hardware circuits related to vehicles, their composition is roughly the same. Main control part, driving part, information acquisition part and human-computer interaction part. Among them, the driving part will probably use H-bridge.

Choose different design schemes according to different needs and take them as needed. When selecting devices, it is necessary to carefully study the datasheet and comprehensively consider the packaging, price and other elements.

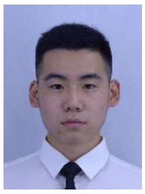
During PCB testing, first ensure that the power supply part can work normally, and then weld and test other components.

References

1. Yuxin Yin. Design of intelligent tracking vehicle based on electromagnetic signal. Inner Mongolia University, 2019.
2. Dogan Ibrahim. ARM-Based microcontroller projects using MBED. 2019, pp. 261-279.
3. Tao Wu. Design of a high voltage IGBT half bridge driver chip. University of Electronic Science and Technology ,2015.

Authors Introduction

Mr. Peng Jia



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.

Mr. Yande Xiang

He is currently pursuing his undergraduate degree at the School of Electronic Information and Automation, Tianjin University of Science and Technology.

Detachable IoT Garbage Sorting Device Based on Machine Vision

Tao Zhu *, Yang Su, Zhiqing Xiao, Fengzhi Dai

Tianjin University of Science and Technology, Tianjin, China

*E-mail: * atust@qq.com*

www.tust.edu.cn

Abstract

China is promoting a garbage sorting system, but people have to spend more time changing their habits, and the garbage bins on the road do not have the function of automatic sorting. Therefore, a detachable IoT waste sorting device based on machine vision is proposed. Through machine vision and a unique mechanical structure design, it can be directly installed on existing trash cans, and is suitable for two and four classifications that meet the Chinese classification standards. Inside the device, it is fixed with an existing trash can with a mechanical electronic lock, and the switch lock operation can be performed through the mechanical key and the Internet of Things applet, which is convenient for the replacement of the device.

Keywords: Garbage classification, detachable,Internet of Things, machine vision

1. Introduction

In recent years, with the accelerated urbanisation of life in China, coupled with the rising living standards of residents, the amount of urban domestic waste generated has maintained a steady growth trend, and in recent years China has become one of the countries with the greatest pressure on urban domestic waste disposal in the world.

According to statistics, China produces hundreds of millions of tonnes of domestic waste every year, and more than 600 cities across the country are piling up 8 billion tonnes of waste, taking up more than 500 million square metres of land. Waste separation and disposal has become a hot social issue, and in order to solve this problem, starting from Shanghai, to 46 cities across the country on a pilot basis, waste separation has been introduced and regulations related to domestic waste separation have been mandatorily implemented.

At present, waste separation in China is mainly carried out manually, i.e. by placing several bins with different categories of labels, and pedestrians are required to make their own judgement before putting them into the bins.

Although manual sorting has been effective in the early stages, it is not a long-term solution. It is not only costly in terms of manpower and resources, but it is also a waste of resources. Smart bins have been proposed and designed for this purpose, but they are still not widely used in public places. The main reasons for this are the high cost, the large size and the low level of intelligence, which does not allow for automatic recognition of the sorting function. In addition, the existing smart bins on the market are capable of sorting recyclable and non-recyclable waste, but are not widely available due to their complex structure and high cost.

2. Overview of Waste Separation Units

The paper discusses a detachable IoT waste sorting device based on machine vision. A convolutional neural network is built through machine learning, and a large amount of waste training data is input into the convolutional neural network for training and classification [1].

In terms of mechanical structure, the paper designs a detachable sorting device that adapts to the majority of bins in the market for both Class II and Class IV bins. A load platform for waste recognition is added above the drop-off opening, which is connected to the acrylic plate using a tiller, and a fixed steel frame is attached underneath the plate to increase the weight of the load. The support structure is made of aluminium alloy to prevent damage to the device.

When sorting waste, the platform can be turned left and right after the recognition device has determined the type of waste to be sorted. The four types of rubbish bin adds a screw bar to the two types of rubbish bin, and the classification of the four types of rubbish is achieved by moving the load-bearing platform on the screw bar through a stepper motor. The whole identification process is simple in structure, short in time to move the device and highly accurate in identification.

After the device has been placed in the bin, the system sends a signal and the electronic lock is activated to attach the device to the bin and is equipped with a mechanical lock on the outside to prevent locking. In addition the unit uses the IoT module SIM7020C for remote upgrades and electronic locking control of the bins, allowing for real-time knowledge of the bin status and updating of the bin training data in the Raspberry Pi during post maintenance [2].

3. Mechanical Structural Design

3.1. Two types of waste separation units

As shown in Fig.1 and Fig.2, the integrated design of the device has two parts. The top plate of the bin can be directly stuck into the middle when retrofitting. The top of the bin is equipped with a hardware box, the sorting structure is designed in the middle drop-off area of the bin, and the upper surface of the lower box is equipped with a camera.

The middle is equipped with an electronically controlled platform for left and right tipping, the tipping board is the same slotted structure that can take the waste. An electronic lock is installed in the upper part of the device, the hardware box can be opened via a remote mini-programmed terminal, and the side is equipped with a mechanical lock at the same time [3].

The model diagram of the two type waste separation unit is shown in Fig.1, and the combination diagram is shown in Fig.2.

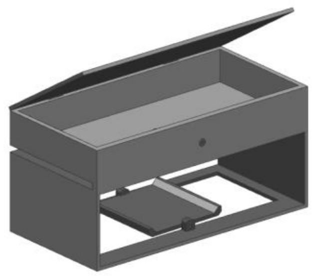


Fig.1 The model diagram of the secondary waste separation unit

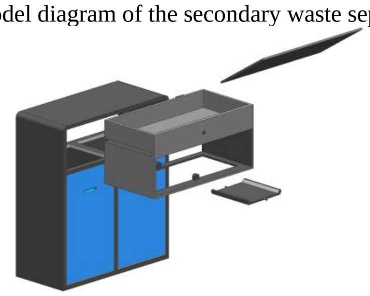


Fig.2 Combination diagram

3.2. Four types of waste separation devices

The sorting structure of the device is installed in the middle input area of the four types of waste bins. The lower surface of the top plate is fitted with a camera, side box is provided with several sorting compartments for collecting different types of waste, side sorting compartments follow the parallel arrangement of common waste bins and therefore have a stronger applicability.

The sorting device can also be dismantled and directly retrofitted to the existing four-type waste classification bin. The sorting compartment is equipped with an electronically controlled platform in the middle of the upper part of the compartment, which can be moved from side to side and tipped downwards, and the tilting and dumping plate is a slotted structure that can receive waste.

Exploded view of the structure of the four types of waste separation units is shown in Fig.3.

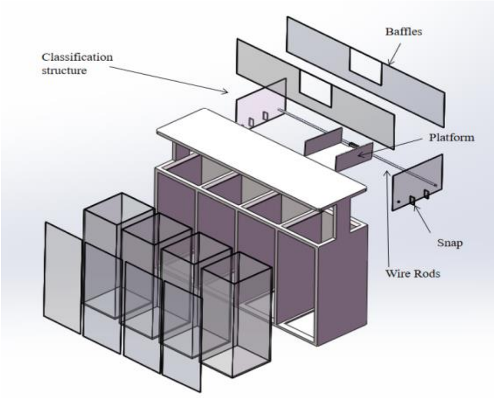


Fig.3 Exploded view of the structure of the four types of waste separation units

The sorting compartment is equipped with an electronically controlled platform in the middle of the upper part of the compartment, which can be moved left and right and turned downwards for dumping. The platform can be turned down 90 degrees, which is shown in Fig.4.

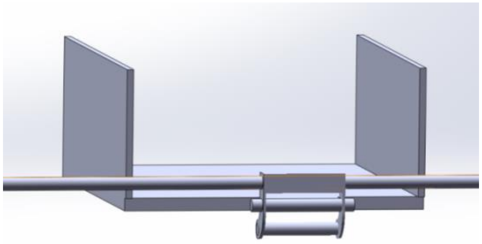


Fig.4 Platform can be turned down 90 degrees

4. Hardware Design

4.1. Master control STM32F103C8T6

The STM32F103C8T6 is chosen as the core processor of the hardware control system. The main functions of the STM32F103C8T6 are: data connection of the Internet of Things, control of the rotating platform servo, system power detection and switching of the mechanical electronic lock.

Its peripheral configuration is powerful, supporting SP interface, I2C interface, USART interface and other communication interfaces, as well as ADC, timer and other peripherals, and reserved for redundant I/O ports, to facilitate the later expansion of the device and maintenance. The master control STM32F103C8T6 is shown in Fig.5.

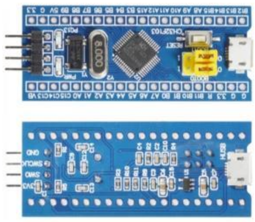


Fig.5 Master control STM32F103C8T6

4.2. Raspberry Pi 4B

The camera sensor pixel is 1080p with 130 degree wide angle, which makes the shooting range perfectly cover the drop-off opening; in addition, it also has a flash function, which is enough to work properly in both day and night.

The camera is mounted directly above the bin drop-off point and takes interval photos to capture images of the rubbish on the platform and transmits the data to the Raspberry Pi.

The Raspberry Pi processes the original images, converting them into grey-scale, binary and RGB images, and matches them with a training set of rubbish that has been pre-input into a convolutional neural network to find common features based on contour recognition, colour recognition and other recognition methods [1].

The data is then transferred to the STM32 to identify the category of waste to be placed. The data is then transferred to the STM32 control board to perform the corresponding operations. The Raspberry Pi 4B is shown in Fig.6.

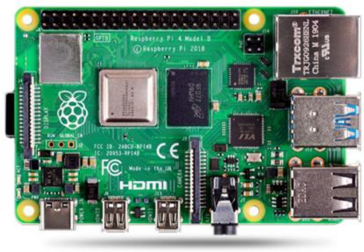


Fig.6 Raspberry Pi 4B

4.3. Electronic lock

The electronic lock is installed in the upper middle of the device, the locking method is magnetic induction. The starting current is 0.9A, the stable current is 0.15A, and the working voltage is DC 12V. The power on is locked while the power off is unlocked, and the power is connected to the microcontroller.

According to the instructions of the microcontroller for on and off, it plays the function of automatic switch lock. The electronic lock is shown in Fig.7.

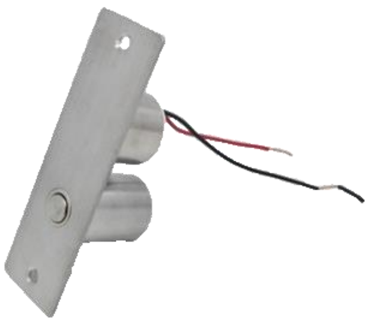


Fig.7 Electronic lock

4.4. IoT module SIM7020C

The module uses the License band and can be deployed in three ways (in-band, protected-band or independent carrier) to coexist with existing networks. It can be deployed directly on GSM, UMTS or LTE networks to reduce deployment costs and enable smooth upgrades. 20db increase in signal gain compared to GPRS Capable of supporting tens of thousands of connections in a single sector, supporting low latency sensitivity, low equipment cost, low equipment power consumption and optimised network architecture [2]. the IoT module SIM7020C is shown in Fig.8.

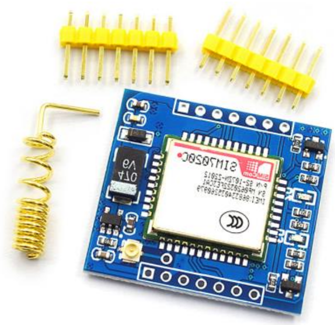


Fig.8 IoT module SIM7020C

5. Conclusion

This paper addresses the problems of difficult waste classification and high replacement cost of similar waste bins in the market, and designs a detachable IoT waste classification device based on machine vision. Through the combination of detachable design and machine vision recognition, it has a wide range of application and high practicability compared with existing intelligent waste sorting bins.

The propossed structure improves the recycling rate of recyclable waste, reduces the loss of material resources and has a certain value of use.

Overall, the innovation of this system lies in the design of Class II and Class IV detachable devices, the combination of machine vision and the Internet of Things, which reduces the cost of waste bin intelligence and improves the maximum benefit of the waste separation.

Acknowledgements

We would like to thank to Mr. Gao Longyu, who gave us the technical support and showed us the way forward when the project encountered difficulties.

References

1. Jianghai Liu, Yadong Jiang. "Design of intelligent trash can be based on machine vision." *2020 International Conference on Image, Video Processing and Artificial Intelligence. Vol. 11584.* International Society for Optics and Photonics, 2020.
2. Yuan, S. C., Yang, S. R., Yang, I. F.,et al., DTMFTalk: A DTMF-Based Realization of IoT Remote Control for Smart Elderly Care. *In International Conference on Smart Grid and Internet of Things* (pp. 16-27). Springer, Cham, 2019.
3. Sun, B., Ma, Q., Zhu, G.,et al., Design and fabrication of an automatic classifying smart trash bin based on Internet of Things. *In Cognitive Computing for Human-Robot Interaction* (pp. 355-375). Academic Press,2021.

Authors Introduction

Mr. Tao Zhu

He is a third year undergraduate student at Tianjin University of Science and Technology. His research interests are hardware design and automation.

Miss. Yang Su

She is a third year undergraduate student at Tianjin University of Science and Technology. She is mainly responsible for the visual recognition.

Miss. Zhiqing Xiao

She majored in automation in college of Electronic Information and Automation, Tianjin University of Science and Technology. She has learned MCU programming, machine vision and other related knowledge, and has made certain achievements in related professional competitions.

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.

A Visual Measurement Algorithm of Approaching Vehicle Speed Based on Deep Learning

Zhu Yurong
School of Electrical and Electronic Engineering,
Shanghai Institute of Technology, Shanghai,201418, China

Zhao Huailin✉
School of Electrical and Electronic Engineering,
Shanghai Institute of Technology, Shanghai, 201418, China

Liu Junjie
School of Electrical and Electronic Engineering,
Shanghai Institute of Technology, Shanghai, 201418, China

Zhang Jinping
School of Electrical and Electronic Engineering,
Shanghai Institute of Technology, Shanghai, 201418, China

Ji Xiaojun
Shanghai Gazecloud Software Technology Co.Ltd.
Shanghai Institute of Technology, Shanghai, 201418, China
E-mail: zhao_huailin@yahoo.com
www.sit.edu.cn

Abstract

With the urbanizational process expediting and the national economy developing rapidly and healthily, the amount of private cars is on the rise, and traffic accidents occur frequently due to speeding and other reasons, and the difficulty of traffic supervision has also increased. This topic will use semantic segmentation and feature extraction and matching. Based on the video data of the traffic surveillance camera, an algorithm is designed to quickly calculate the matching of feature points in adjacent frames with low computing power to achieve the calculation. The same vehicle moves within the two frames of the target, so as to calculate the speed of the vehicle. Firstly, performing semantic segmentation based on deep learning, we choose a fully convolutional network to achieve semantic segmentation of depth maps, and distinguish the picture’s principal part. After that, we can realize features extraction and mapping. The HOG algorithm is used on the matching step, the target’s relative movement is calculated based on these matched point pairs to measure the moving speed of the vehicle. The experiment and the test prove that the system can realize the efficient speed measurement of moving vehicles.

Keywords: Vehicle speed measurement; semantic segmentation; HOG feature extraction; SVM classification

1. Introduction

There are more and more cars on the road, and the traffic has become heavier and heavier, which has led to the frequent traffic accidents and the difficult traffic supervision. Some data show that on the road sections equipped with "electronic eyes", the number of people and vehicles observing traffic regulations has increased significantly, indicating that the use of some technological methods will enhance people's cautiousness and have a certain binding force on people's irregular behaviors[1]. Strengthen people's attention to traffic rules, thereby reducing the incidence of accidents and reducing conflicts between personnel. The "electronic eye" for the speed monitoring system of mobile vehicles is one of them, which reduces the incidence of traffic accidents caused by speeding, indicating that the installation of these monitoring facilities can avoid the occurrence of some major traffic accidents and effectively reduce the difficulty of traffic supervision[2,3]. In order to improve the level of public security and reduce the incidence of traffic incidents, it is necessary to research and improve vehicle detection and monitoring technologies.

The world is strengthening the development of emerging technologies, science and technology are changing with each passing day, society is becoming more and more informatized, network information technology, communication transmission technology and multimedia technology and other high-tech are constantly advancing with the times, penetrating all aspects of life, and intelligent transportation systems are gradually gaining popularity and become favored by people. As the number of vehicles on the road increases, the monitoring of their speed becomes more and more important[4].

The core of this subject is to realize the speed measurement of moving vehicles. For a few frames of image speed measurement, the image needs to be processed first. If the entire image is processed directly, the amount of calculation is large and complicated, so semantic segmentation is needed to divide the main body of the image. As a key technology, semantic segmentation is used in computer vision and other fields. The realization is that the pixels of the image are processed and classified to obtain several different semantic categories, so that the image is divided into many sub-regions, and the semantics of these sub-regions have different specific meanings. Since the advent of Convolutional Neural Network (CNN), this technology has been continuously improving. Researchers have tried to use it to solve various problems[5,6]. Through unremitting efforts, it has proved its potential in the processing of semantic segmentation problems, whether in automatic driving, medical image segmentation, pedestrian and vehicle detection all have good application prospects.

This paper is based on the traffic monitoring video and uses low computing power to calculate the speed of moving vehicles. It helps to enhance the supervision of vehicle speed, better determine the responsibility of the accident, and can also serve as a warning to the driver. It can reduce the incidence of accidents. It has good application prospects in traffic monitoring, autonomous driving and other fields.

2. System scheme

2.1. *Overall design*

For the video data obtained by the traffic camera, deframe processing is first performed to obtain a frame-by-frame image sequence. Then perform semantic segmentation based on deep learning on the image sequence to obtain the main body of the image, that is, the vehicle. Then for the image sequence containing the vehicle, the vehicle detection based on HOG and SVM is performed[7]. Finally, some algorithms are used to measure the speed of moving vehicles. The overall design is shown in Figure 1.

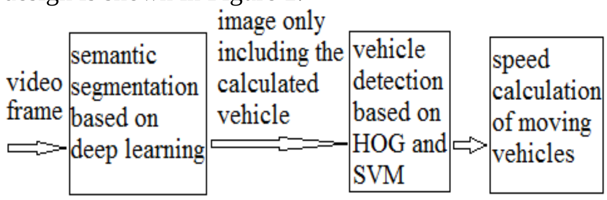


Fig. 1. The overall design

2.2. *Semantic segmentation based on deep learning*

Semantic segmentation, as a key research field of computer vision, has good applications for flat data, three-dimensional or high-dimensional data. In daily life, more and more scenes need to be understood, such as unmanned autonomous driving, reality augmentation, medical image processing, etc. We need to consider relevant semantics from each scene to effectively solve problems, and semantic segmentation is very important. Conducive to the understanding of the scene, so we need to continue to develop semantic segmentation technology.

The network model of semantic segmentation is usually composed of an encoder and a decoder. The encoder, which is a trained network that can be used in classification. The decoder determines the different structures of semantic segmentation. Its work is mainly to process the features. The low-resolution obtained in the encoder is based on semantic projection to obtain high-resolution features, so that it can be realized classification.

Semantic segmentation based on CNN is generally divided into three stages[8,9], namely pixel-level classification, down-sampling, and up-sampling. Figure 2 shows a common semantic segmentation architecture. For the convolutional network, down sampling achieves a feature map that reduces its resolution and increases the number of channels. It is mainly used to extract low-level semantic features in the image, as well as some abstract features. The characteristics of up-sampling are completely different. The size of the feature map will increase and the number of channels will decrease. After up-sampling is followed by down-sampling, the semantic content can be recovered slowly. In the classification stage, multi-classification can be achieved, by classifying all pixels, mainly based on semantic categories.

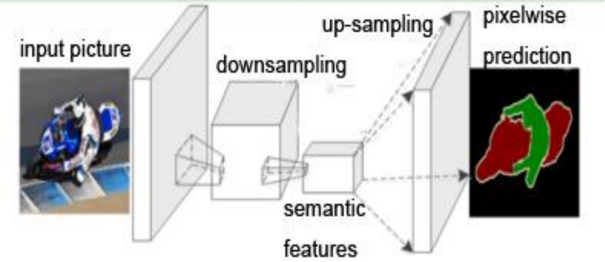


Fig. 2. Semantic segmentation based on CNN

2.3. *HOG feature extraction*

HOG is called Histogram of oriented gradient. It can use gradient information to describe part of the limited features of the image, and then use a series of algorithms to obtain the gradient feature information of the entire image, which is mainly used in computer vision field.

When performing HOG feature extraction on an image, the local information of the image can be depicted more completely through gradient information or edge information. To achieve HOG feature extraction need four steps: image preprocessing, calculation of the gradient vector of the image, extracting the histogram of the gradient direction, block normalization. The details are as follows.

2.3.1. *Image preprocessing*

The first step is to process the input image to achieve grayscale. Because the effect of color information in the HOG algorithm is not obvious, the impact on detection accuracy is also small, and the amount of information contained in the grayscale image will be less, and the calculation pressure will also be reduced. So usually the color image is grayed out first.

The second step is to perform Gamma normalization processing on the grayscale image. After the implementation of the first step, there are still lighting, shadows, and noise that affect HOG feature extraction. In order to reduce the impact of these factors, Gamma normalization (normalization) of the image obtained in the first step is used to solve the problem. The Gamma compression formula is shown in the following formula.

$$I(x,y) = I(x,y)^{gamma} \tag{1}$$

Where, $I(x, y)$ represents the value of (x, y). And gamma index refers to gamma correction.

2.3.2. The gradient vector of the image

The difference in gradient value can reflect the change of discontinuous features, such as gray value. Based on the pixel level, if the grayscale difference is larger for the image, then the gradient difference will be larger accordingly. For a certain object to be tested, the gradient change of the edge feature is generally the largest, and we usually use this feature to determine the edge position. Generally speaking, the first-order derivative of the function is used to express the image gradient, and the second-order derivative is used to express its change.

The usual method is to first use convolution in the original image to calculate the gradient components of the plan image, and then divide it horizontally and vertically. For the pixel (x, y) in a certain image, the pixel can be calculated by the following formula.

$$G_x(x,y) = H(x+1,y) - H(x-1,y) \tag{2}$$

$$G_y(x,y) = H(x,y+1) - H(x,y-1) \tag{3}$$

$$G(x,y) = \sqrt{G_x(x,y)^2 + G_y(x,y)^2} \tag{4}$$

$$\alpha(x,y) = \tan^{-1}\left(\frac{G_y(x,y)}{G_x(x,y)}\right) \tag{5}$$

Among them, for a certain pixel in the input image, $G_x(x,y)$ is its horizontal component, $G_y(x,y)$ is vertical component, $H(x,y)$ is pixel value, $G(x,y)$ is gradient magnitude, and $\alpha(x,y)$ is gradient direction.

2.3.3. Extracting the histogram of the gradient direction

In HOG feature extraction, each picture is divided into a finite number of cells of the same size and independent of each other. Different shapes can be divided into rectangles and stars. The calculation methods used are different, and the feature dimensions are also different. Generally, the picture is first divided into rectangular unit blocks of the same size, and then the unit blocks are equally divided into 9 directions, and the direction of the unit block can be obtained by weighting according to the gradient magnitude of each direction.

2.3.4. Block (interval) normalization within the image block

In the third step, the image has been divided into many Cell blocks, and then the characteristics of each Cell block are analyzed. Then, according to the obtained feature attributes of the Cell blocks, the Cell blocks need to be combined, the Cell blocks are concatenated to obtain a Block block, and the features of each Cell block are concatenated to obtain the entire HOG feature vector of the corresponding Block. Block blocks are allowed to overlap, but Cell blocks must be independent of each other. From this it can also be known that in the entire feature descriptor, there may be situations where the same feature gets different description results. In order to obtain the HOG feature vector of the complete image, each Block block needs to be connected in series, and this descriptor can be used for the processing of classification problems.

Through the above process, it can be known that Block blocks will overlap. When this phenomenon occurs, the pixels within it will be calculated multiple times, and the obtained gradient values will be more different, so that the accuracy of classification will be effectively improved.

2.4. *SVM classification*

SVM is called Support Vector Machine (Support Vector Machine). SVM has a good application when dealing with classification and extraction problems. It is a good algorithm that can be used for machine learning. It was first studied in 1995. Support vector machine is mainly used to deal with the problem of feature classification, a supervised learning algorithm based on boundary classification. The purpose of this algorithm is to find a boundary in a plane or a boundary in a high-dimensional space to realize the classification of features. When dealing with simple binary classification problems, this dividing line is easy to find and meet the requirements, but when dealing with multi-feature classification or complex binary classification problems, it is relatively more complicated. For dividing lines or interfaces that need to be expressed by mathematical expressions, it is important to master the relevant theories of support vector machines.

SVM needs the support of mathematical tools to effectively solve super-high-dimensional classification problems. Practice has proved that SVM works best when the number of samples is lower than the spatial dimension. In addition, SVM only needs to store support vectors, so it can effectively save memory space. Usually, in the construction of the classifier, the method of using HOG and SVM together is used. Because at this time, the gradient histogram can be used as a very robust descriptor, which can accurately reflect the characteristics of a certain class, which will make the classification effect better.

For linearly separable data, the classification process of SVM is shown in Figure 3. In addition to the processing of linearly separable problems, it can also handle data linearly inseparable. The usual method is to use some non-linear transformations to map this difficult problem in low-dimensional space to high-dimensional space, so that It becomes a linearly separable data to be solved, as shown in Figure 4. In order to reduce the complexity, the kernel function will be quoted in high-dimensional operations.

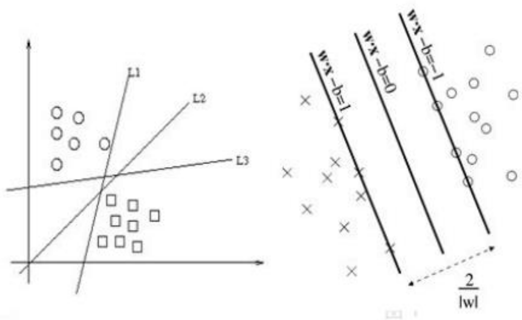


Fig. 3. Linear case

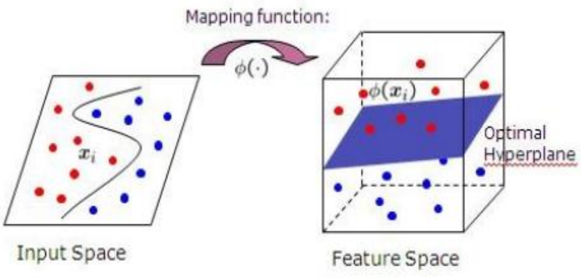


Fig. 4. Non-linear case

3. Programming

According to the analysis and implementation steps of the above-mentioned related technologies. First of all, for a traffic video data, we use an algorithm to extract the images of adjacent frames in the video data, and then perform semantic segmentation based on the full convolutional neural network to divide the main body and the background of the image. Next, input the vehicle (the main body of image) into the SVM classifier for training to further ensure the accuracy of the segmentation. After that, extract and match the HOG features of the main body in the image. When the gap in the number of video frames is within 10 frames, target tracking is performed to obtain a certain distance between the different frames. When moving distance of a vehicle is calculated, we could get its moving speed. Finally, the video where the speed of the moving vehicle in the original video has been measured is stored, and the export the video.

From the above methods, the program flow chart of this article can be made, based on the algorithm realization process, as shown in Figure 5.

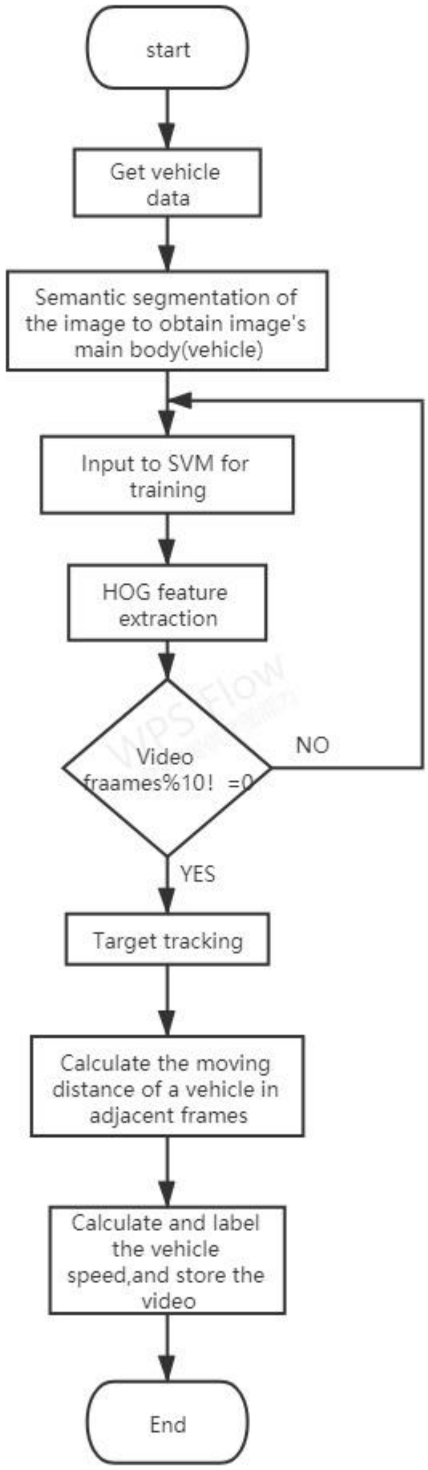


Fig. 5. Program flow chart

4. Experiment

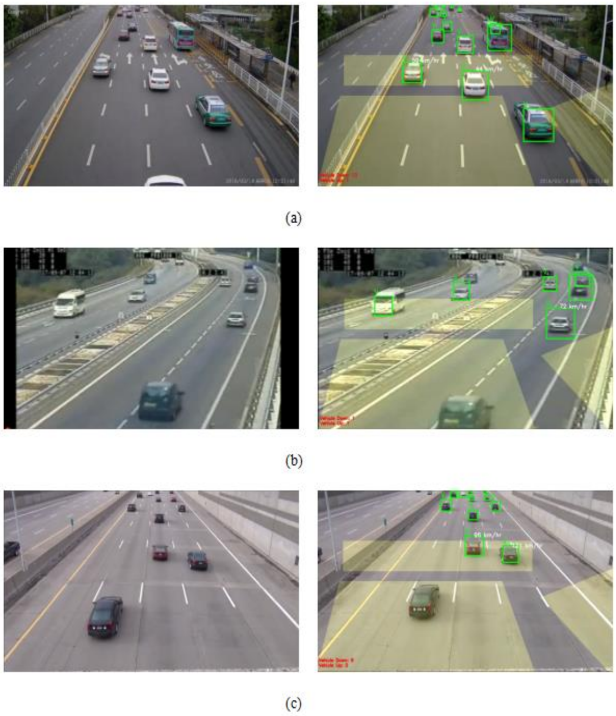
In the previous sections, we mainly introduced and analyzed the related technologies and algorithms used in the speed measurement of moving vehicles, including CNN, HOG, and SVM classifiers and so on. In this section, it is mainly based on the following implementation process to obtain and analyze the corresponding results.

4.1. *Experimental Results*

First of all, for a traffic video data, we first use an algorithm to extract the images of adjacent frames in the video data, and then perform semantic segmentation based on the full convolutional neural network to achieve the segmentation of the vehicle body and the background. Then for the image body, also That is the vehicle. Then input into the SVM classifier for training, and then extract and match the HOG features of the main image. When the difference in the number of video frames is within 10 frames, perform target tracking to obtain a certain vehicle moving distance between adjacent frames of images, and then calculate Find out its moving speed, and finally store the video where the speed of the moving vehicle in the original video has been measured, and export the video.

After the original video data is run and debugged by the program, the video data that has been marked with the moving speed of the vehicle will be output. The screenshot of the video part is shown in Figure 6. Through the video data obtained from the data set, debugging and running in the program, the result shown in Figure 6 has been obtained, which can meet the requirements for the speed measurement of moving vehicles. Next, add some programs and add a yellow area to the output to better see how the program implements the vehicle speed measurement. Figure 7 shows three original video frames compared with the corresponding experiment results.

Fig. 6. Output video screenshot



(a)

(b)

(c)

Fig. 7. Comparison of three groups of experiments

4.2. *Experiment Analysis*

Given the corresponding number of samples and the number of pictures to be tested, the number of samples is about 700, and the corresponding number of pictures to be tested is about 1800. The running results are shown in Table 1.

Table 1. Statistics of running results

No. of samples	No. Of pictures to be deteced	No. Of recognition	Accuracy rate
100	300	231	77.0%
100	400	324	81.0%
200	400	324	82.0%
300	700	657	93.9%

It can be seen from the Table1 that as the number of samples increases, the recognition rate is significantly improved, and the accuracy rate is also improved. Therefore, for the relevant data, the target of the vehicle can basically be identified. And the number to be tested is not as large as possible. It is necessary to select an appropriate value so that it can be better applied to engineering practice.

According to the results of the above-mentioned different traffic monitoring videos, it can be known that the program can detect more than 90% of the vehicles in the image, with a high detection rate, and the results will also show the moving speed of the vehicles. Next, the detection accuracy under light changes is further verified. In the three cases of medium, strong, and weak light intensity, 600 test samples are intercepted for accuracy measurement. The results are shown in Table 2.

Table 2. Detection degree under different illumination

Light intensity	Detection accuracy rate(%)	False detection rate(%)	Missing detection rate(%)
Normal light	97.22	2.32	0.46
Strong light	94.31	3.46	2.23
Weak light	91.87	1.77	6.36

It can be seen from Table 2 that the detection accuracy of HOG features can be maintained above 90% under the influence of different illumination, and its anti-interference ability is strong.

The speed measurement of the highway video with known speed is used to judge the accuracy of the moving vehicle speed calculation algorithm used in this article. Figure 8 shows the speed measurement result of a blue truck on a certain highway.

It is known that the actual speed of the blue truck in the picture is 80km/h. In Fig. 8(c), in the first 6 speed calculations, the error is kept within ±3km/h. However, it can be seen that in the 7th and 8th calculations, the data fluctuates greatly. Then it gradually calmed down. On the whole, from the line graph data, we can see that in multiple vehicle speed calculation experiments, the accuracy of the vehicle speed measured by the experiment has reached more than 90%, so the speed measurement method adopted in this paper can better meet the test requirements.

(a) Blue truck speed measurement start frame

(b) Blue truck speed test end frame

(c) Blue truck speed test line chart

Fig. 8. Blue truck speed measurement

5. Conclusion

By using the data in different data sets for testing, and then using the above method for result analysis and evaluation, it can be known that the program can basically detect all vehicles with a high detection rate, and it can also successfully display the moving speed of most vehicles. However, the correctness of the speed needs to be further verified, and the program will be further optimized and improved afterwards, so as to achieve a more accurate speed measurement of moving vehicles.

In sum, the proposed algorithm can effectively achieve the speed calculation of the moving vehicles. The image segmentation, HOG feature extraction, and the speed measurement are of high accuracy.

References

1. Tong, Zheng, Philippe Xu, and Thierry Denœux. "Evidential fully convolutional network for semantic segmentation." Applied Intelligence (2021): 1-24.
2. D. Y. Chan, D. C. Morris, T. J. Polascik, M. L. Palmeri and K. R. Nightingale, "Deep Convolutional Neural Networks for Displacement Estimation in ARFI Imaging," in IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 68, no. 7, pp. 2472-2481.
3. Lahiani, Houssem, and Mahmoud Neji. "Hand gesture recognition method based on HOG-LBP features for mobile devices." Procedia Computer Science 126 (2018): 254-263.
4. Gan, Jiangzhang, Jiaye Li, and Yangcai Xie. "Robust SVM for Cost-Sensitive Learning." Neural Processing Letters (2021): 1-22.
5. Ye, Lu, Ting Duan, and Jiayi Zhu. "Neural network-based semantic segmentation model for robot perception of driverless vision." IET Cyber-systems and Robotics 2, no. 4 (2020): 190-196.
6. Wang, Qi, Xiang Gao, Fan Wang, Zhihang Ji, and Xiaopeng Hu. "Feature Point Matching Method Based on Consistent Edge Structures for Infrared and Visible Images." Applied Sciences 10, no. 7 (2020): 2302.
7. Pozzer, Sandra, Ehsan Rezazadeh Azar, Francisco Dalla Rosa, and Zacarias Martin Chamberlain Pravia. "Semantic segmentation of defects in infrared thermographic images of highly damaged concrete structures." Journal of Performance of Constructed Facilities 35, no. 1 (2021): 04020131.
8. Krishnakumar, K., and S. Indira Gandhi. "Video stitching based on multi-view spatiotemporal feature points and grid-based matching." The Visual Computer 36, no. 9 (2020): 1837-1846.
9. Wang, Qi, Xiang Gao, Fan Wang, Zhihang Ji, and Xiaopeng Hu. "Feature Point Matching Method Based on Consistent Edge Structures for Infrared and Visible Images." Applied Sciences 10, no. 7 (2020): 2302.

Authors Introduction

Ms. Yurong Zhu

She received the B.S. degree from Shanghai Institute of Technology, Shanghai, China. And now she is a master course student. Her main research interests are deep learning and Control Science and Engineering.

Dr. Huailin Zhao

He received his PhD from Oita University, Japan in 2008. He is a professor in the School of Electrical & Electronic Engineering, Shanghai Institute of Technology, China. His main research interests are robotics, multi-agent system and artificial intelligence. He is the members of both IEEE and Sigma Xi.

Mr. Junjie Liu

He received the B.S. degree from Shanghai Institute of Technology, Shanghai, China. His main research interests are deep learning and Control Science and Engineering.

Mr. Jinping Zhang

He received the B.S. degree from Shanghai Institute of Technology, Shanghai, China. And now he is a master course student. His main research interests are deep learning and Control Science and Engineering.

Mr. Xiaojun Ji

He received the master degree from Southeast University, Nanjing, China. He is a senior software engineer, system architect and product manager. And he is the founder of Shanghai Gazecloud Software Technology Co. Ltd..

Target Search Based on Scene Priors

Shengyang Lu, Lanjun Liang, Huailin Zhao✉, Fangbo Zhou, Feng yao
School of Electrical and Electronic Engineering,
Shanghai Institute of Technology, Shanghai, China
E-mail: 405838753@qq.com, zhao_huailin@yahoo.com
www.sit.edu.cn

Abstract

Aiming at the problems of reinforcement learning algorithm in target search tasks, such as low accuracy and low fault tolerance, this article mainly introduces a method of reinforcement learning target search based on scene prior in simulation environment. This method mainly uses graph convolutional neural network to extract the current object relationship as the input of prior knowledge. Secondly, it uses the actor-critic algorithm to take the agent's vision, position and prior knowledge as input to decide the agent's next navigation. Finally, use path planning to navigate to the target point to find the target. Through experiments conducted in Habitat and compared with the previous algorithm, the experiment shows that this method is better than the previous algorithm in target search accuracy and navigation efficiency.

Keywords: Target search, Reinforcement Learning, Scene priors, visual navigation

1. Introduction

In recent years, the field of robot research has been committed to expanding the ability of robots to explore the environment, understand the environment, interact with the environment and communicate with people. As one of the most important tasks of agent visual navigation, visual semantic navigation requires the robot to navigate to the target location by using the observed visual information according to the given target in the unknown environment. The agent solves two problems at the same time. The first is where to go, that is, where the target object is, and the second is how to get there, that is, planning an effective path to navigate to the target object. This has far-reaching significance for both the research in the field of artificial intelligence and the application in real life, such as disaster and battlefield rescue, smart home, unmanned driving and so on. Moreover, it is also of far-reaching significance for the research in other fields, such as embodied question answering [1], visual semantic navigation [2], and Visual Dialogue navigation [3].

The traditional navigation method [4] usually uses the environment map for navigation, and divides the navigation task into three steps: drawing map, positioning and path planning. This method usually needs to build a 3D map in advance, as well as reliable map positioning and path tracking. However, the map can not be used in the invisible environment. Recently, the success of data-driven machine learning strategies for various control and perception problems has opened up a new way to overcome the limitations of the previous method [4]. The key point of these methods is to directly learn the mapping between the original observation and operation of the end-to-end way of the task. These methods take advantage of the ability of previous navigation experience in a new similar environment, whether there is a map or not. Reinforcement learning (RL) is often used in visual navigation. However, reinforcement learning still has some problems, such as low generalization ability, low navigation efficiency and low accuracy.

Recently, chaplot [5] proposed a target driven navigation method based on learning model and won the first place in CVPR 2020 habitat objectnav challenge, called "goal oriented semantic exploration" "Visual navigation using semantic mapping and reinforcement learning can cover the whole scene for target search to the greatest extent. However, this method still has the problems of low learning efficiency and low navigation accuracy.

However, human navigation still exhibits excellent generalization performance, which cannot be explained by spatial or topological memory. For example, people who come home for the first time will naturally go to the kitchen (instead of outdoor or toilet) to pick up plates; From the kitchen to the bedroom, they know that the living room may be a midway point. Although visually different, this kind of semantic knowledge, the "close" relationship of semantic entities, is naturally shared across environments and can be learned from past experience to guide future navigation.

Prior knowledge can not only help to navigate to known objects, but also help agents find targets based on the current visual exploration of unknown environment. Based on this, we propose a target driven navigation model based on semantic prior. The model uses semantic prior knowledge to assist navigation, and improves the ability of agent target search by learning the relationship between objects power.

In order to evaluate our model, we trained and tested our model in the habitat [6] simulation environment. Habitat simulation environment allows us to train AI agents in realistic and efficient 3D simulators. The simulation environment includes Gibson [7] and matterport3d [8] data sets, including "kitchen", "toilet", "living room" and other scenes, allowing agents to explore the whole room to find targets, which greatly increases the difficulty of the task. The agent is initialized to the random position in the scene, and explores through visual information to find the target. Experiments show that our target search algorithm based on scene a priori is superior to other methods in both navigation success rate and average step.

In this paper, our innovations can be summarized as follows:

- We propose a target search algorithm based on semantic a priori. The algorithm uses semantic a priori knowledge to assist the agent to explore the unknown environment, and helps the agent to search the target through the relationship between objects.
- We train and test our method in the habitat simulation environment. Experiments show that our method is superior to other algorithms in both the success rate of target search and the efficiency of navigation.

2. Related Work

Navigation is one of the most fundamental issues for mobile robots. Traditional methods such as slam construct metric maps by perceiving signals, which are subsequently used in planning. Recently, due to advances in deep learning, end-to-end methods have been applied to navigation in several fields, such as mazes, indoor scenes [24,25], autonomous driving [26,27]. There is also an excellent summary of recent progress [2,37]. We focus on indoor navigation scenarios, using the house3d environment [28], which contains relationships between semantic entities that are consistent with the real world and provides ground truth labels of objects and scenarios.

2.1. Semantic navigation

Semantic navigation and target search tasks have far-reaching significance for both the development of artificial intelligence and the application of daily life. Recently, many methods use reinforcement learning as a navigation decision module to control agents to navigate in unknown environments. Such as A3C [9], PPO [10], DQN [11], etc. In addition, Liang [12] et al. Proposed an algorithm to explicitly model the scene a priori using the confidence perception semantic scene completion module to complete the scene and guide the navigation planning of the agent. Moreover, many opportunistic memory mechanisms[13],attention mechanisms [14] and algorithms of transfer learning [15] and imitation learning [16] are also applied to visual navigation. Although a large number of algorithms have improved the performance of searching objects in agent exploration environment, they still face the characteristics of small target search range, low accuracy and low robustness. And these algorithms do not take into account the relationship characteristics between objects.

2.2. Scene priors

The target search algorithm based on scene a priori takes more account of the relationship between objects, and uses this relationship to assist agents in environment exploration and target search. In recent years, many people use prior knowledge to assist various tasks, and have made good achievements in such tasks. For example, scene atlas is used for traffic light search [17], road detection [18] and dialogue tasks [19]. Moreover, in terms of visual navigation, Yang [20] et al. Proposed a reinforcement learning model framework that uses graph convolution neural network coding to assist exploration of scene prior knowledge. Liu [21] and others applied the scene prior knowledge to the multi-agent system, which effectively improved the efficiency of multi-agent system navigation and target search. Li [22] et al. Combined prior knowledge with meta learning and made great progress in map free visual navigation.

3. Problem description

In unknown scene s , The visual information of the agent is recorded as O , $O=\{O_t^1, O_t^2, ..., O_t^k\}$. At the beginning, the agent is initialized to the random position and random posture in the room, which is recorded as $P_0=\{x_0, z_0, \theta_0\}$. (x_0, z_0) Represents the random position of the agent initialized in the room at the initial time, θ_0 is the initial rotation angle. Initially, the agent receives the tag T of the target as input . At each time point, the agent receives visual information from the visual sensor O_t , and its odometer receives attitude information P_t . Visual information O_t consists of first person RGB and depth information.

At each time point t , agent A receives the first person visual information O_t from the visual sensor, obtains the pose information P_t from the odometer. Visual information and pose information are used to predict the corresponding semantic map m_t . Each agent learns its corresponding navigation strategy π to determine the corresponding expected navigation point L , $L_t=\pi(m_{t-1}, T, L_{t-1})$. In order to reach the expected target point, the agent obtains the next action A_t through path planning. When action A_t is performed in each agent, its visual input is updated to O_{t+1} and the odometer input P_{t+1} . Go back and forth until you find all the targets.

4. Problem description

This paper proposes an actor critic target search model based on scene a priori, which is mainly composed of four parts. As shown in Figure 1, they are semantic mapping module, prior knowledge extraction module, feature fusion module and action decision module. The semantic decision module maps the first person vision into top-down semantic map features. The object relationship feature uses the scene prior knowledge extraction module to extract the object relationship in the first person RGB information. The feature fusion module is to fuse the current object feature relationship and semantic mapping vector with the previous state. The action decision module is to decide the next expected goal and generate the next action that the agent needs to perform.

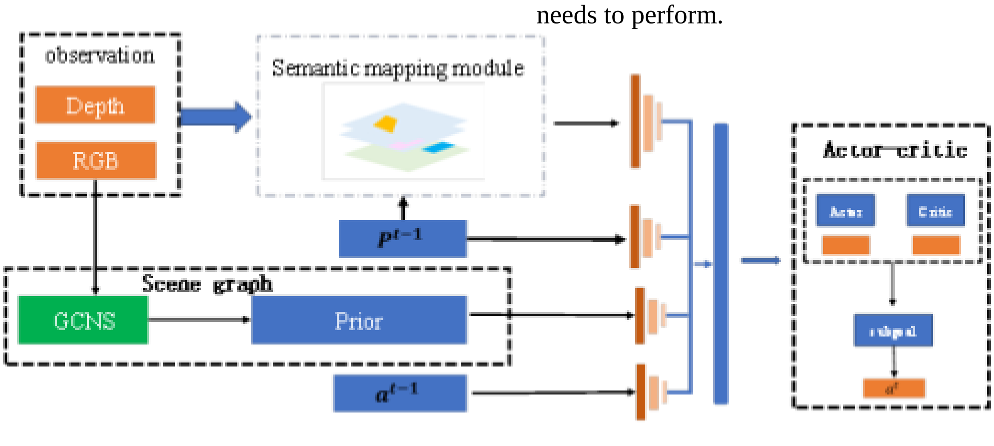


Fig. 1: overall framework

4.2.Semantic mapping module

In this module, the method similar to literature [5] is mainly used to generate idiom semantic mapping map. Our agent uses RGB and depth to generate point cloud information, voxels it, and finally turns it into a semantic map. Then, reinforcement learning is used to generate the expected goals of each agent. In this process, semantic maps are constantly being improved.

We mainly use spatial semantic maps and can access the location of agents at any time. The structure of spatial semantic map is a matrix of $K \times M \times N$, in which the size of $M \times N$ is the size of the actual map. Firstly, the size of $K \times 240 \times 240$ is generated, then the matrix is mapped to the size of the actual map, and finally the spatial semantic map m_t is generated. The semantic map is composed of voxels, and the size of each voxel block is $16m^2$. The number of channels of spatial semantic map k = 2 + C, C represents the semantic information of each object in the actual environment, and the first two layers represent the map size and explored places respectively. Each element on the map represents the place that has been explored and corresponds to the position of the semantic channel object in the map. During initialization, the map is initialized with an all 0 matrix, such as $[0]^{k \times M \times N}$, and the agent is initialized in a random position

4.3. Prior knowledge extraction

The object relationship feature extraction module adopts a method similar to that in literature [31], brings semantic knowledge into the reinforcement learning framework, integrates the relationship features between current environmental objects using graph convolution networks (GCNs) [32], and dynamically updates and saves them when the agent receives the environmental information.

With the constant navigation of the agent, the agent's understanding of the whole environment is increasingly clear, and the relationship between the object and the object is gradually improved. These relational vector assistant agents are used to find the target objects. Based on these relational agents, they can find objects more efficiently and with a higher success rate.

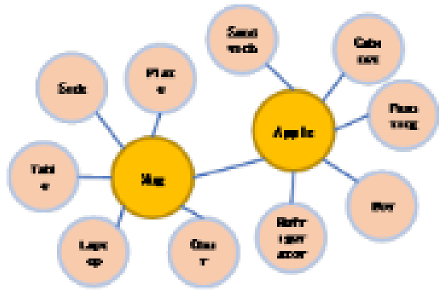


Fig 2. Scene priors

In this paper, the prior knowledge of the scene is represented in the form of undirected graph G={V,E}.The node in V represents different types of objects, and the edge E represents the special positional relationship between the two types of objects.

As an extension of graph neural network to graph structure, GCNs aims to learn the functional representation of a given graph G={V,E}. We generalize all nodes into characteristic matrices. The graph structure is represented by binary adjacency matrix $F=[F_1, F_2 ... F_{|V|}]$. we standardize the matrix A to obtain the matrix $\hat{A}$. Each node of GCNs output is represented as $Z=[z_1, z_2, ..., z_{|V|}]$. So we can get:

$$H^{(l+1)} = f(\hat{A}H^{(l)}W^{(l)})$$

Where, $H^{(0)} = X$, $H^{(L)} = Z$, $W^{(l)}$ is the parameter of layer l and L is the total number of layers of GCNs.

In this section, three layers of GCNs are used, the input is RGB image, the output of the first two layers is 1024 dimensional potential features, and the output of the last layer is the value output by each node to obtain the feature vector $|V|$.

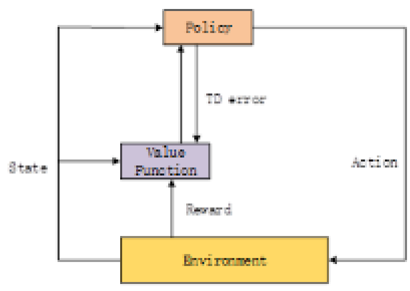


Fig 3. Navigation model

The feature vector is the semantic coding information of the current scene and environment context. Finally, this eigenvector is mapped to a 512 dimensional eigenvector $f_{K,t}{}^{k}$. The 512 dimensional feature vector is used as the input feature to assist the agent in visual navigation.

4.3. Navigation decisions

At this stage, we mainly use reinforcement learning to solve the problem of our multi-agent visual navigation. The reinforcement learning model is shown in the figure below. Our long-term goal is to let the agent team find the goal of concern, so the agent must navigate to the area where the goal is located. The multi-agent spatial semantic map described in Section 4.2 is used as the input, and the short-term goals that can be reached by each agent are generated through the network above. The output short-term goals are in the form of {x, z}. The short-term goals will lead the agent team to conduct visual guidance in an interactive environment. When the agent sees the target concerned by the problem, the channel corresponding to the target in the spatial semantic map will be non-0. When the channel corresponding to the target in the spatial semantic map is 0, the short-term goal will guide the agent to explore more places until the information of the spatial semantic map is enough to answer the problem. As shown in Fig 3.

5. Experimental setup

We use habitat simulation environment to test our model. Haibitat contains Gibson and matterport3d data sets, but we only choose Gibson data set as simulation environment to train and test the effect of our model.

The agent team is initialized at any position in the environment, so the observation space is composed of the first person vision (RGB and depth) of each agent, which is a 4 * 480 * 640 matrix. The coordinates of the agent mainly include the position and attitude information of the agent, which is composed of $\{x, z, \theta\}$, which is a 1×3 matrix. The action space of agent is continuous. When the agent finds the target, that is, there is a semantic mapping of the target label in the semantic map, it means that the navigation is successful, and then the navigation will end, or the length of the action sequence of the agent exceeds 3000, then we think that the agent is very familiar with the scene this time, can not find the target, and there is no other significance to continue searching, Therefore, it will also end this navigation and determine that the navigation is wrong.

In order to take more account of the relationship between objects, the semantic map we designed includes 15 categories. During the retraining process of our hierarchical reinforcement learning network, the reward is saved every 200 actions, the batch size is 36 in each reinforcement learning update weight process, and it is updated every 4 epochs. We use the Adam optimizer with a learning rate of 0.000025, and the weight size is set to γ=0.95.

6. Experimental setup

This article uses two evaluation functions: success rate (SR) and average path length (SPL) to test the effect of the algorithm. The success rate is defined as:

$$SR=\frac{1}{N_{task}}\sum_{i=1}^{N_{task}} R_i \qquad (2)$$

Among them, when the i round of experiment is successful, $R_i = 1$ otherwise $R_i = 0$, N_{task} is the total number of rounds of the experiment. The higher the success rate, the better the search effect of agents.

The average path length is defined as:

$$SPL=\frac{1}{N}\sum_{i=1}^{N} S_i \frac{l_i}{\max(p_i, l_i)} \qquad (3)$$

Among them p_i is the length of the path that the agent moves in the first round, l_i is the shortest path length from the initial position of the agent to the target, and N represents the total number of rounds. S_i defines whether this round is successful. The lower the average path length, the higher the efficiency of multi-agent searching for targets.

Distance to succes is defined as:

$$DST=\max(\|x_T - G\|_2 - d_s, 0) \qquad (4)$$

Where $\|x_T - G\|_2$ is the L2 distance of the agent from the goal location at the end of the episode, d_s is the success threshold.

6.1. Navigation decisions

We use two end-to-end Reinforcement Learning (RL) methods as baselines:

RGBD + RL: A common recursive RL strategy initialized with the resnet18 trunk,. Proxy poses and target object categories are passed through the embedded layer and attached to the loop layer input.

RGBD + Semantics + RL: This baseline passes semantic segmentation and object detection prediction and rgbd input to recursive RL strategy. We use the pre trained mask RCNN, which is the same as the RCNN used in the proposed model, for semantic segmentation and object detection in this baseline. Rgbd observations are encoded using resnet18 backbone visual encoder, and proxy pose and target object are encoded using the above embedded layer.

Both RL-based baselines are trained using proximity strategy optimization [10], using a dense reward that reduces the distance to the nearest target. We designed two additional baselines based on a combination of goal-independent exploration methods and heuristic-based local goal-driven strategies.

SemExp[5] : This baseline maps the visual inputs onto a three-dimensional semantic map, predicting the semantic map by using a method similar to the active nerual slam. Meanwhile, semantic map information is coded as vectors and trained using PPO's algorithm, and a navigation strategy is learned.

Our method mainly improves on this algorithm, adds scene prior knowledge on the basis of semantic map, and codes as knowledge vector to fuse with the semantic map, puts it into the network for learning, and finally learns a single-agent navigation strategy.

6.2. Quantitative experiment

TABLE1 Comparison of quantitative experimental results

	SPL	SR	DTS
Random	0.004	0.004	3.893
RGBD + RL	0.027	0.082	3.310
RGBD + Semantics + RL	0.049	0.159	3.203
SemExp	0.199	0.544	1.723
ours	**0.304**	**0.621**	**0.555**

We trained our algorithm over 10,000 times on Gibson and evaluated its performance in our test set. As shown in the table above, our algorithm has the highest performance compared to several algorithms, whether SPL, SR or DTS. By analyzing the table above, we can draw a few conclusions：

For random walk, the scene is larger in Gibson simulation environment, and the agent needs to navigate to find the target. For these three criteria, random walk is the worst. For RGBD+RL, RL models the process of exploration as a partially visible Markov model, but the model can not fit a navigation model perfectly due to less information, and the training speed is slow.

RGBD + semantic + RL and semexp take more into account the semantic information and spatial information between scenes, so they have a certain improvement in accuracy. However, they do not take into account the relationship characteristics between objects, so our model is superior to other methods in the above three indicators. Therefore, we draw a conclusion, Our model is more suitable for semantic exploration in unknown environment.

6.2. Qualitative experiment

In this section, we visualize first person vision for intelligent navigation as well as semantically predicted maps and navigation paths, with the specific structures shown in figure 4.

In this experiment, given an intelligent one semantically tagged as a " toilet " intelligent achieves this goal by predicting a semantic map and navigating, eventually finding a target, when it is found, turning the color of the target blue, providing us with the maximum priority that we need to navigate, and then navigating directly to the location of that target.

As shown in figure4, at the beginning the intelligent is randomly initialized at random locations throughout the environment, and it is not clear to the intelligent what the environment looks like, so the whole semantic map is blank. As the intelligence receives first person visual information, the intelligence begins to predict the semantic map, predict the situation across the scene, and compute an intended target based on the current incomplete semantic map, as well as scene prior information, which is where the intelligence wants to go. As shown in figure 4, at the beginning the intelligent is randomly initialized at random locations throughout the environment, and it is not clear to the intelligent what the environment looks like, so the whole semantic map is blank.

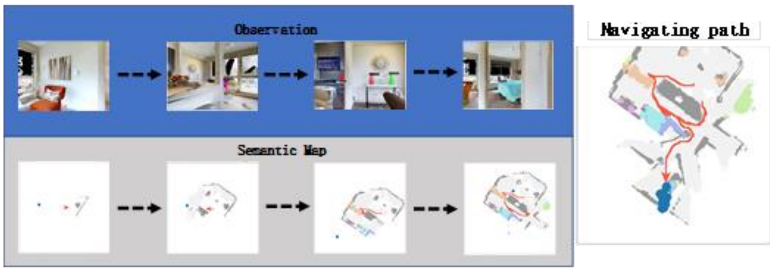


Fig 4.visual semantic navigation path

As the intelligence receives first person visual information, the intelligence begins to predict the semantic map, predict the situation across the scene, and compute an intended target based on the current incomplete semantic map, as well as scene prior information, which is where the intelligence wants to go.

In the algorithm design, this design takes LK optical flow tracking as an extension of the better key point. Therefore, the LK optical flow tracking node can be used to subscribe the key points that have been obtained in the better key point node. The algorithm can be used to pass the current key point. Two grayscale pictures predict the next set of key points and perform backward prediction.

5.Conclusion

This paper mainly introduces a visual navigation method based on scene a priori. Firstly, this method uses GCNs to calculate the relationship between objects and encode them. Secondly, the method calculates the prediction semantic map of the scene through the first person visual information and its own pose information. Encode the above information, fuse features, and conduct visual navigation through actor critical algorithm to finally find the target. Through qualitative and quantitative experiments, it is concluded that the algorithm proposed in this paper is effective in both success rate and average path length.

6.References

1 Das, Abhishek, et al. "Embodied question answering." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018.

2 Anderson P, Wu Q, Teney D, et al. Vision-and-language navigation: Interpreting visually-grounded navigation instructions in real environments[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 3674-3683.

3 Thomason J, Murray M, Cakmak M, et al. Vision-and-dialog navigation[C]//Conference on Robot Learning. PMLR, 2020: 394-406.

4 Bailey T, Durrant-Whyte H. Simultaneous localization and mapping (SLAM): Part II[J]. IEEE robotics & automation magazine, 2006, 13(3): 108-117.

5 Chaplot D S, Gandhi D P, Gupta A, et al. Object goal navigation using goal-oriented semantic exploration[J]. Advances in Neural Information Processing Systems, 2020, 33.

6 Savva M, Kadian A, Maksymets O, et al. Habitat: A platform for embodied ai research[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019: 9339-9347.

7 Xia F, Zamir A R, He Z, et al. Gibson env: Real-world perception for embodied agents[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 9068-9079.

8 Chang A, Dai A, Funkhouser T, et al. Matterport3d: Learning from rgb-d data in indoor environments[J]. arXiv preprint arXiv:1709.06158, 2017.

9 Babaeizadeh M, Frosio I, Tyree S, et al. GA3C: GPU-based A3C for deep reinforcement learning[J]. CoRR abs/1611.06256, 2016.

10 Schulman J, Wolski F, Dhariwal P, et al. Proximal policy optimization algorithms[J]. arXiv preprint arXiv:1707.06347, 2017.

11 Hernandez-Garcia J F, Sutton R S. Understanding multi-step deep reinforcement learning: a systematic study of the DQN target[J]. arXiv preprint arXiv:1901.07510, 2019.

12 Liang Y, Chen B, Song S. Sscnav: Confidence-aware semantic scene completion for visual semantic navigation[C]//2021 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2021: 13194-13200.

13 de Andrade Santos I B, Romero R A F. Deep Reinforcement Learning for Visual Semantic Navigation with Memory[C]//2020 Latin American Robotics Symposium (LARS), 2020 Brazilian Symposium on Robotics (SBR) and 2020 Workshop on Robotics in Education (WRE). IEEE, 2020: 1-6.

14 Seymour Z, Thopalli K, Mithun N, et al. MaAST: Map Attention with Semantic Transformersfor Efficient Visual Navigation[J]. arXiv preprint arXiv:2103.11374, 2021.

15 Wang L, Zhao L, Huo G, et al. Visual semantic navigation based on deep learning for indoor mobile robots[J]. Complexity, 2018, 2018.

16 Wang X, Huang Q, Celikyilmaz A, et al. Reinforced cross-modal matching and self-supervised imitation learning for vision-language navigation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 6629-6638.

17 Barnes D, Maddern W, Posner I. Exploiting 3D semantic scene priors for online traffic light interpretation[C]//2015 IEEE Intelligent Vehicles Symposium (IV). IEEE, 2015: 573-578.

18 Alvarez J M, Gevers T, Lopez A M. 3D scene priors for road detection[C]//2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2010: 57-64.

18 Mishra A, Alahari K, Jawahar C V. Scene text recognition using higher order language priors[C]//BMVC-British Machine Vision Conference. BMVA, 2012.

19 Yang W, Wang X, Farhadi A, et al. Visual semantic navigation using scene priors[J]. arXiv preprint arXiv:1810.06543, 2018.

20 Liu X, Guo D, Liu H, et al. Multi-Agent Embodied Visual Semantic Navigation with Scene Prior Knowledge[J]. arXiv preprint arXiv:2109.09531, 2021.
21 Li F, Guo C, Luo B, et al. Multi goals and multi scenes visual mapless navigation in indoor using meta-learning and scene priors[J]. Neurocomputing, 2021, 449: 368-377.
22 He K, Gkioxari G, Dollár P, et al. Mask r-cnn[C]//Proceedings of the IEEE international conference on computer vision. 2017: 2961-2969.
23 Wang L, Zhao L, Huo G, et al. Visual semantic navigation based on deep learning for indoor mobile robots[J]. Complexity, 2018, 2018.
24 Liang Y, Chen B, Song S. Sscnav: Confidence-aware semantic scene completion for visual semantic navigation[C]//2021 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2021: 13194-13200.
25 Wang J, Liu J, Kato N. Networking and communications in autonomous driving: A survey[J]. IEEE Communications Surveys & Tutorials, 2018, 21(2): 1243-1274.
26 Yurtsever E, Lambert J, Carballo A, et al. A survey of autonomous driving: Common practices and emerging technologies[J]. IEEE access, 2020, 8: 58443-58469.
27 Panayi N. In-house 3D designing of lingual and labial customized brackets using composite-made customized bases[J].

Authors Introduction

Mr. Shengyang Lu

He received the B.S. degree from Shanghai Institute of technology, Shanghai, China, in 2018. He is currently pursuing the M.S. degree in Control Science and Engineering at Shanghai Institute of Technology. His main research interests are deep learning and ROS robots.

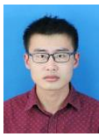
Ms.Lanjun Liang

She received B.S. degree in automation from Shanghai Polytechnic University, Shanghai, China, in 2020. She is studying for the master degree of control science and engineering at Shanghai Institute of Technology, Shanghai, China. her main research interests are deep learning and ROS robots.

Dr. Huailin Zhao

He received his PhD from Oita University, Japan in 2008. He is a professor in the School of Electrical & Electronic Engineering, Shanghai Institute of Technology, China. His main research interests are robotics, multi-agent system and artificial intelligence. He is the members of both IEEE and Sigma Xi.

Mr. Fangbo Zhou



He received the B.S. degree in automation from Chuzhou University, Chuzhou, China, in 2019. He is currently pursuing the M.S. degree in safety engineering at Shanghai Institute of Technology, Shanghai, China. His research interests include crowd counting and ROS robots.

Mr.Feng Yao

He received the B.S. degree from Shanghai Institute of technology， Shanghai, China, in 2020. He is currently pursuing the M.S. degree in Control Science and Engineering at Shanghai Institute of Technology, Shanghai, China. His main research interests are reinforcement learning and human-robot collaboration.

A Generalized Hamiltonian Conservative System with Multi-scroll Chaotic Flows

Jingwen Liu*, Zhonggao Chen
College of Electronic Information and Automation, Tianjin University of Science and Technology,
Tianjin 300222, China
*E-mail: * 1229778495@qq.com*
www.tust.edu.cn

Abstract

By analyzing mechanics and energy of a three-dimensional volume conservative chaotic system proposed by Vaidyanathan and Volos, a new generalized conservative chaotic system with multi-scroll chaotic flows is found based on the corresponding Hamiltonian energy. The new system satisfies both volume conservation and energy conservation. By analyzing the equilibrium characteristics of the system, equilibrium points of the new system are found to be a line. In addition, the number of scrolls of conservative chaotic flows of the new system depend on the corresponding Hamiltonian energy. The paper provides a new conservative chaotic model for chaos application.

Keywords: conservative, multi-scroll, Hamiltonian energy, equilibrium

1. Introduction

Chaos is a special physical phenomenon in nonlinear systems[1], which widely exists in the fields of physics and life science. Therefore, people are committed to constructing chaotic systems with better performance[2]. Compared with single scroll chaotic systems or double scroll chaotic systems, multi-scroll chaotic systems have more control parameters and more corresponding key parameters. Moreover, it can present complex multi-direction grid scroll in phase space. The number and shape of scroll can also be controlled and adjusted by the parameters of the system[3]. In practical application, the increase of scrolls number in chaotic systems are realized by increasing the number of equilibrium points. Multi-scroll chaotic attractors can show more complex chaotic dynamic behaviors[4-6], and has high application value in chaotic information processing[10], chaotic neural network[7] and chaotic secure communication[12], which makes the multi scroll chaotic system have a very broad application prospect in practical engineering[10-13]. Therefore, the research of multi scroll chaotic system is becoming a research hotspot in the field of chaos[14-18].

However, there are few studies on the multi-scroll flows of conservative chaotic systems. In this paper, single direction multi-scroll flows and double direction multi-scroll flows are constructed and their direction and number are both controllable. At the same time, the feasibility and effectiveness of the method of constructing multi-scroll flows is verified by Lyapunov exponent spectrum and phase diagram, which provides a new method for the construction of multi-scroll flows and a new conservative chaotic model for chaos application.

2. Construction of Four-dimension Conservative Chaotic System

Firstly, a new four-dimension conservative system is found by analyzing mechanics and energy of a three-dimensional volume conservative chaotic system proposed by Vaidyanathan and Volos[19]. It can be expressed as

$$\dot{\mathbf{x}} = J(\mathbf{x})\nabla H(\mathbf{x}). \tag{1}$$

Where, $J(\mathbf{x}) = \begin{bmatrix} 0 & a & x & 0 \\ -a & 0 & y & 0 \\ x & -y & 0 & 1+2w \\ 0 & 0 & -1-2w & 0 \end{bmatrix}$,
$\nabla H(\mathbf{x}) = [x\ y\ z\ 1]^T$.

Secondly, Equation (1) can be described as

$$\begin{cases} \dot{x} = ay + xz \\ \dot{y} = -ax + yz \\ \dot{z} = 1 - x^2 - y^2 - 2w \\ \dot{w} = -z - 2zw \end{cases}. \tag{2}$$

Where x, y, z and w are state variable, and a is a system parameter, the divergence of the system (2) is

$$\nabla f = \frac{\partial \dot{x}}{\partial x} + \frac{\partial \dot{y}}{\partial y} + \frac{\partial \dot{z}}{\partial z} + \frac{\partial \dot{w}}{\partial w} = 0. \tag{3}$$

Moreover, the derivative of Hamiltonian energy is

$$\dot{H} = \nabla H(\mathbf{x})^T J(\mathbf{x}) \nabla \mathrm{H}(\mathbf{x}) = 0. \tag{4}$$

According to the above analysis, the system satisfies both Hamiltonian energy conservation and volume conservation. Meanwhile, the equilibrium equation of system (2) is

$$\begin{cases} ay + xz = 0 \\ -ax + yz = 0 \\ 1 - x^2 - y^2 - 2w = 0 \\ -z - 2zw = 0 \end{cases}. \tag{5}$$

It can be found that the equilibrium point of the system (2) is obtained as $(0, 0, z, -0.5)$, which are found to be a line.

3. Construction of Conservative Chaotic Systems with Multiple multi-scroll flows

In this part, multi-scroll flows are obtained by changing Hamiltonian energy. According to Equation (1), the Hamiltonian energy of system (2) is obtained as

$$H(\mathbf{x}) = \frac{1}{2}(x^2 + y^2 + z^2) + w. \tag{6}$$

Set system parameter $a = 1$, the initial value $(x, y, z, w) = (x(0), 1, 1, 1)$, Lyapunov exponent spectrum of the system (2) is shown in Fig. 1.

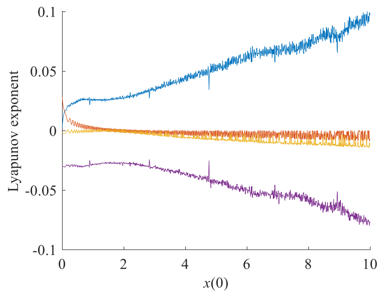


Fig. 1. Lyapunov exponent spectrum of system (2)

Set the initial value $(x, y, z, w) = (\frac{\pi}{2}, 1, 1, 1)$, Single scroll flow is shown in Fig. 2.

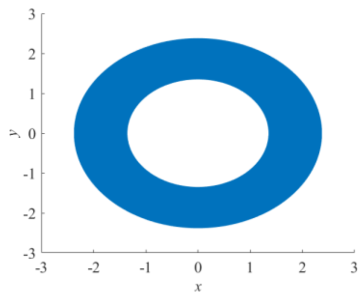


Fig. 2. Single scroll flow

3.1. Characteristic analysis of the single direction multi-scroll conservative system

In order to obtain the multi-scroll conservative system, the Hamiltonian energy is firstly changed by introducing piecewise functions. Set $\nabla H(\mathbf{x}) = [f(x)\ y\ z\ 1]^T$, and Equation (2) can be described as

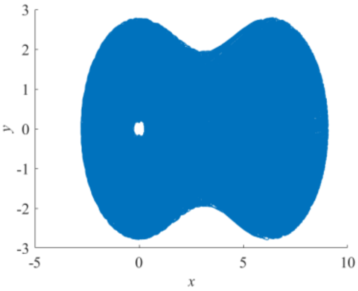
$$\begin{cases} \dot{x} = ay + xz \\ \dot{y} = -af(x) + yz \\ \dot{z} = 1 - xf(x) - y^2 - 2w \\ \dot{w} = -z - 2zw \end{cases}, \tag{7}$$

$$f(x) = \begin{cases} x + Q, & x < -Q \\ sin(x), & -N \le x \le Q \\ x - N & x > N \end{cases}. \tag{8}$$

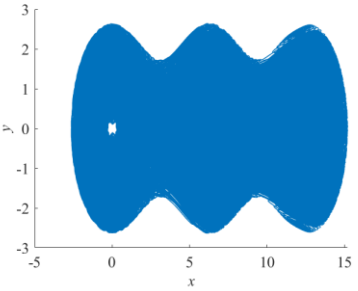
Where $f(x)$ satisfy mapping, and $Q = n_1\pi, N = n_2\pi, n_1, n_2 \in Z^*$.The Hamiltonian energy of system (2) is obtained as

$$H(\mathbf{x}) = \frac{1}{2}(y^2 + z^2) + w + \int f(x)\, dx. \tag{9}$$

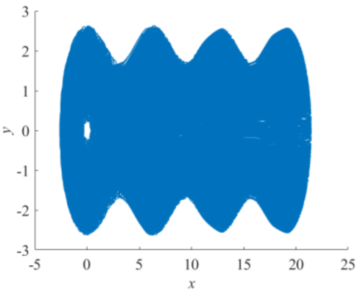
To further explain the process of generating multi-scroll in system (7), set $Q = 0$ and the initial value $(x, y, z, w) = (\frac{\pi}{2}, 1, 1, 1)$. When $N = 2\pi, 4\pi$ and 6π, a 2-scroll flow, a 3-scroll flow and a 4-scroll flow extended in the positive direction of x-axis are respectively shown in Fig. 3.



(a) 2-scroll flow



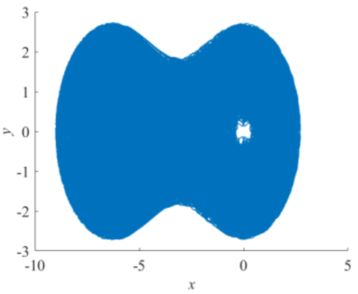
(b) 3-scroll flow



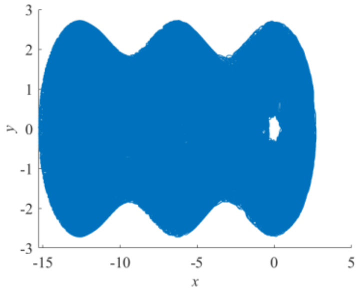
(c) 4-scroll flow

Fig. 3. Multi-scroll flows of the system (7)

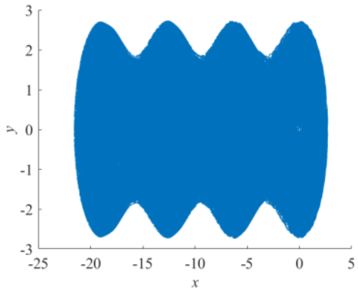
Set $N = 0$ and the initial value $(x, y, z, w) = (\frac{\pi}{2}, 1, 1, 1)$.When $Q = 2\pi, 4\pi$ and 6π, a 2-scroll flow, a 3-scroll flow and a 4-scroll flow extended in the negative direction of x-axis are respectively shown in Fig. 4.



(a) 2-scroll flow



(b) 3-scroll flow



(c) 4-scroll flow

Fig. 4. Multi-scroll flows of the system (7)

In summary, it can be concluded that flows with different numbers of scrolls can be obtained by changing the values of N and Q to change the Hamiltonian energy of the system (7), and both the direction and the number of scrolls is controllable.

3.2. Characteristic analysis of the double direction multi-scroll conservative system

In the above, multi-scroll flows along the x-axis direction are obtained in system (7) by changing the Hamiltonian energy of x variable. Similarly, grid type multi-scroll flows along x and y axis directions can be simultaneously obtained by changing the Hamiltonian energy of x variable and y variable. Set $\nabla H(\mathbf{x}) = [f(x)\, f(y)\, z\, 1]^T$, and Equation (2) can be described as

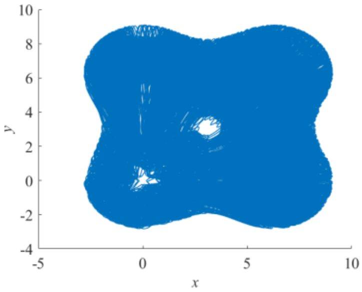
$$\begin{cases} \dot{x} = af(y) + xz \\ \dot{y} = -af(x) + yz \\ \dot{z} = 1 - xf(x) - yf(y) - 2w' \\ \dot{w} = -z - 2zw \end{cases} \tag{10}$$

$$f(x) = \begin{cases} x + Q_1, & x < -Q_1 \\ sin(x), & -N_1 \le x \le Q_1, \\ x - N_1, & x > N_1 \end{cases} \tag{11}$$

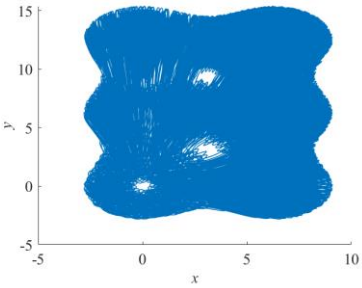
$$f(y) = \begin{cases} y + Q_2, & y < -Q_2 \\ sin(y), & -N_2 \le y \le Q_2, \\ y - N_2, & y > N_2 \end{cases} \tag{12}$$

Where $f(x)$ and $f(y)$ respectively satisfy mapping (11) and mapping (12) , and $Q_1 = m_1\pi, Q_2 = m_2\pi, N_1 = m_3\pi, N_2 = m_4\pi, m_1, m_2, m_3, m_4 \in Z^*$. The Hamiltonian energy of system (10) is obtained as

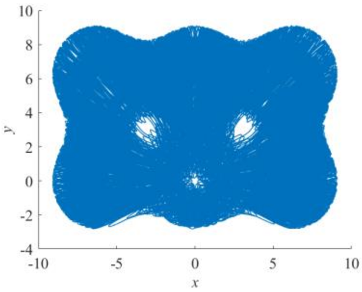
$$H(\mathbf{x}) = \frac{1}{2}z^2 + w + \int f(y)\,dy + \int f(x)\,dx. \quad (13)$$



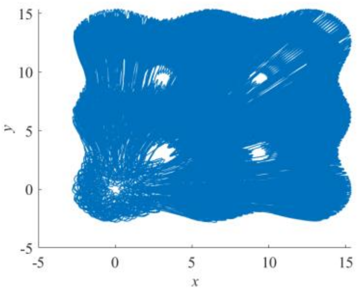
(a) 2×2 grid type multi-scroll flow



(b) 2×3 grid type multi-scroll flow



(c) 3×2 grid type multi-scroll flow



(d) 3×3 grid type multi-scroll flow

Fig. 5. Grid type multi-scroll flows of the system (10)

To further explain the process of generating multi-scroll in system (10), set the initial value $(x, y, z, w) = (\frac{\pi}{2}, \frac{\pi}{2}, 1, 1)$. When $Q_1 = Q_2 = 0, N_1 = N_2 = 2\pi$, a 2×2 grid type multi-scroll flow is shown in Fig. 5 (a). When $Q_1 = Q_2 = 0, N_1 = 2\pi, N_2 = 4\pi$, a 2×3 grid type multi-scroll flow is shown in Fig. 5 (b). When $Q_2 = Q_1 = 0, N_1 = 4\pi, N_2 = 2\pi$, a 3×2 grid type multi-scroll flow is shown in Fig. 5 (c). When $Q_1 = Q_2 = 0, N_1 = N_2 = 4\pi$, a 3×3 grid type multi-scroll flow is shown in Fig. 5 (d).

In summary, it can be concluded that different grid type multi-scroll flows can be obtained by changing the values of Q_1, Q_2, N_1 and N_2 to change the Hamiltonian energy of the system (10),, and both the direction and the number of scrolls is controllable.

4. Conclusion

In this paper, a new generalized conservative chaotic system with multi-scroll chaotic flows is found based on the corresponding Hamiltonian energy through the mechanical and energy analysis of the three-dimensional volumetric conservative chaotic system proposed by Vaidyanathan and Volos. The new system satisfies the conservation of volume and energy. At the same time, the feasibility and effectiveness of the method of constructing multi-scroll flows is verified by Lyapunov exponent spectrum and phase diagram, which provides a new method for the construction of multi-scroll flows and a new conservative chaotic model for chaos application.

References

1. Norbert Euler,Maria Clara Nucci & Da Jun Zhang(2021).*Nonlinear Systems and Their Remarkable Mathematical Structures, Volumes 1, 2, and 3*.CRC Press
2. Xiu Zhao, et al."Complex generalized synchronization of complex-variable chaotic systems." *The European Physical Journal Special Topics* .prepublish(2021): doi:10.1140/EPJS/S11734-021-00129-6.
3. Complexity Research; New Complexity Research Study Findings Reported from Civil Aviation University of China (A New Memristor-based 5d Chaotic System and Circuit Implementation. *Technology News Focus* .(2019): doi:
4. Chen,D.Y;Sun,Z.T.;Ma,X.Y.;Chen,L.Circuit implement ation and model of a new multi-scroll chaotic system.Int.J.CircuitTheoryAppl.2014,42,407–424.
5. Deng, W.H.; Lü, J.H. Generating multi-directional multi-scroll chaotic attractors via a fractional differential hysteresis system. Phys. Lett. A 2007, 369, 438–443.

6. Zhang, Y .; Yu, S.M.; Liu, M.H. Generating multi-scroll hyperchaotic attractors based on FPGA technology. Circuits Syst. 2007, 12, 39–43.
7. Samaneh Gholami,etal."An efficient image-based verification scheme by fusion of double random phase encoding and dynamic chaotic map." Multimedia Tools and Applications 78.17(2019),10.1007/s11042-019-7714-1.
8. Yongfeng Cui,,Zhongyuan Zhao,,Yuankun Ma & Shi Dong.Resource allocation algorithm design of high quality of service based on chaotic neural network in wireless communication technology. Cluster Computing(5),2019,10.1007/s10586-017-1285-6.
9. Lili Zhou & Fei Tan.A chaotic secure communication scheme based on synchronization of double-layered and multiple complex networks. Nonlinear Dynamics(2), 2019,10.1007/s11071-019-04828-7.
10. Yu, S.M.; Lü, J.H.; Leung, H.; Chen, G.R. Design and implementation of n-scroll chaotic attractors from a general jerk circuit. IEEE T rans. Circuits Syst. I Regul. Pap. 2005, 52, 1459–1476.
11. He, S.B.; Sun, K.H.; Wang, H.H.; Ai, X.X.; Xu, Y .X. Design of n-dimensional multi-scroll Jerk chaotic system and its performance. J. Appl. Anal. Comput. 2016, 6, 1180–1194.
12. Liu, X.Z.; Shen, X.M.; Zhang, H.T. Multi-scroll chaotic and hyperchaotic attractors generated from Chen system. Int. J. Bifurc. Chaos 2012, 22, 1250033.
13. Chen, Z.; Wen, G.L.; Zhou, H.A.; Chen, J.Y . A new M x N-grid double-scroll chaotic attractors from Rucklidge chaotic system. Optik 2017, 136, 27–35.
14. Tlelo-Cuautle, E.; Rangel-Magdaleno, J.J.; Pano-Azucena, A.D.; Obeso-Rodelo, P .J.; Nunez-Perez, J.C. FPGA realization of multi-scroll chaotic oscillators. Commun. Nonlinear Sci. Numer. Simul. 2015, 27, 66–80.
15. Luo, X.H. Circuitry implementation of a novel nonautonomous hyperchaotic Liu system based on sine input. Chin. Phys. B 2009, 18, 3304–3308.
16. Hu, X.; Liu, C.; Liu, L.; Ni, J.K.; Li, S.L. Multi-scroll hidden attractors in improved Sprott A system. Nonlinear Dyn. 2016, 86, 1725–1734.
17. Tang, W.K.S.; Zhong, G.Q.; Chen, G.; Man, K.F. Generation of n-scroll attractors via sine function. IEEE T rans. Circuits Syst. I Fundam. Theory Appl. 2001, 48, 1369–1372.
18. Chen Z G, Song M H. Dynamic Characteristics Analysis of a Multi-scroll Conservative Chaotic System with Sinusoidal Nonlinearity. The 2021 International Conference on Artificial Life and Robotics (ICAROB2021), Jan. 21-24.
19. Jia H Y, Chen Z G, Shi W X, et al. A generalized Hamiltonian conservative chaotic system with multi-stable flows. Journal of Shandong University (Engineering Science), 2021, 51 (6).

Authors Introduction

Ms. Jingwen Liu

She is studying for a master's degree in electronic information at Tianjin University of Science and Technology.

Mr. Zhonggao Chen

He received the B.S. degree from Tianjin University of Science and Technology, Tianjin, China. And now she is a master course student. Her main research interests are research and circuit implementation of multi-scroll Hamiltonian energy conservative system.

Multi-stability and FPGA Implementation of a Conservative Chaotic System

Minghan Song

College of Electronic Information and Automation, Tianjin University of Science and Technology, 300222，China;

*E-mail: *892095883@qq.com*

www.tust.edu.cn

Abstract

The paper first studies the reason why a three-dimensional volume conservative chaotic system proposed by Vaidyanathan and Volos can generate chaos by analyzing mechanics and energy. Then, based on numerical methods including balance characteristics, Lyapunov exponents, bifurcation diagrams, phase trajectories and so on, multi-stability of the three-dimensional volume conservative chaotic system are discovered. In addition, the three-dimensional volume conservative chaotic system is realized by using FPGA, and all the results from FPGA implementation are consistent with those from numerical analysis.

Keywords: Chaos,Conservative,Multi-stability, FPGA implementation

1. Introduction

In 1964, Hénon and Heiles proposed a conservative chaotic system for the first time in the study of three-body motion, which was called the Hénon-Heiles system[1,2]. In 1994, Sprott proposed some simple three-dimensional chaotic systems,it is often called Sprott-A system[3]. This system is actually a special case of the Nosé-Hoover system, also known as the Nose-Hoover conservative oscillator[4-6]. Subsequently, in 1997 and 1999, Sprott and Thomas respectively gave the jerk conservative chaotic system and the conservative chaotic system with sine function[7-8]. Subsequent research on conservative chaotic characteristics is gradually attracting researchers' attention[9-15].

In addition, in recent years, chaotic systems with multi-stability characteristics are gradually becoming a hot issue in chaotic theory and application research[16-21]. Multi-stability phenomenon usually refers to the phenomenon that different manifolds coexist in the system when different initial values are taken when the system parameters are unchanged. In particular, when the initial value of the system is changed, the number of manifolds tends to be infinite, this phenomenon is called super multi-stability[22-24]. Multi-stability is ubiquitous in many natural systems and usually has an important impact on system performance. Although there are no attractors in conservative systems, they are extremely sensitive to small fluctuations under initial conditions. Due to the existence of multiple stability, the system is extremely sensitive to small disturbances under initial conditions, which often leads to the coexistence of multiple streams. This uncertainty and complexity also attract more and more attention[25,26].

The rest of the paper is organized as follows: in Section 2, we conduct a mechanical analysis of a three-dimensional conservative chaotic system and reveal the cause of chaos. Then We conduct research on whether the system produces multi-stability in Section 3. After that, FPGA implementation of this system is showed in Section 4, while Section 5 concludes the paper.

2. A three-dimensional conservative chaotic system and its analysis

In 2013, Vaidyanathan and Volos[9] proposed a three-dimensional volumetric conservative chaotic system, which was described as:

$$\begin{cases} \dot{x} = xz + ay \\ \dot{y} = yz - bx \\ \dot{z} = 1 - x^2 - y^2 \end{cases} \tag{1}$$

Among them, x, y, z are state variables, a and b are state parameters. When setting a=0.05, b=1, the calculated Lyapunov exponent is:

$$L_1 = 0.0395, L_2 = 0, L_3 = -0.0395 \tag{2}$$

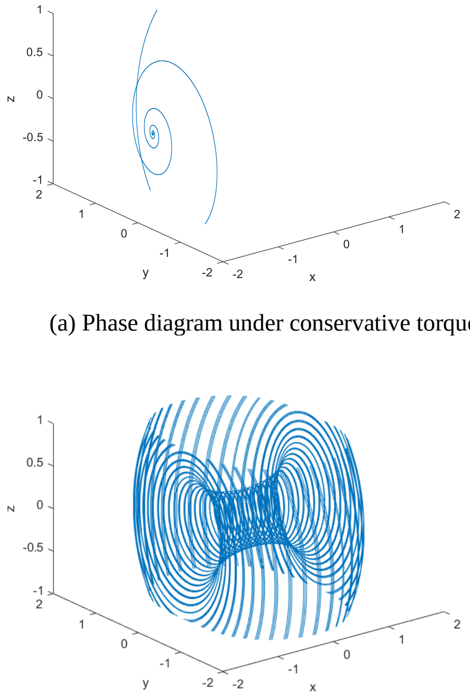
It can be seen that the sum of Lyapunov exponents of system (1) is 0, so the system is a conservative chaotic system.

When choosing different initial values and parameters of the system, the system can also present rich periodic or chaotic dynamics. In order to further study the reasons for the chaotic dynamics of the system (1) from the perspective of energy characteristics, first convert it into the Kolmogorov form as:

$$\dot{x} = J(x)\nabla H(x) + u = \begin{bmatrix} 0 & a & x \\ -b & 0 & y \\ -x & -y & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{3}$$

where, $J(x) = \begin{bmatrix} 0 & a & x \\ -b & 0 & y \\ -x & -y & 0 \end{bmatrix}$, $u = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, $H(x) = \frac{1}{2}(x^2 + y^2 + z^2)$. Through the analysis, it can be found that the system (3) consists of two parts: the conservative moment $J(x)\nabla H(x)$ and the external moment *u*.

When we set a=b=10, take the initial value $(x, y, z) = (1,1,1)$, the phase diagram under the action of conservative moment and the phase diagram under the action of conservative moment and external moment are obtained, as shown in Fig. 1. From Fig.1, it can be seen that the external torque is the main reason for the chaotic phenomenon of the system.



(a) Phase diagram under conservative torque

(b) Phase diagram under the combined action of conservative torque and external torque

Fig.1 Phase diagram

3. Multi-stability analysis of the three-dimensional conservative chaotic system

Generally, systems with hidden attractors are extremely sensitive to initial values, and there will often be situations where multiple attractors coexist, that is called multi-stability. Multi-stability means that when the system parameters are fixed, corresponding to different initial values, the system presents different dynamic characteristics, and the corresponding phase diagram will also show different attractors.

3.1. a≠b

When we select the initial value of the system (1) $(x, y, z) = (1, y_0, -3)$, the parameters a=1, b=1.6, where y_0 varies from -5 to 5, and the Lyapunov exponent diagram and bifurcation diagram that vary with y_0 are shown in Fig.2 and Fig.3. It can be seen that when $y_0 \in (-5,-4.3) \cup (-3.3, 0.2) \cup (1.5, 5)$, the maximum Lyapunov exponent of the system is greater than 0, and the system is in a chaotic state; when $y_0 \in (-4.3, -3.3) \cup (0.2, 1.5)$, the motion state of the system changes between cycles and counter cycles. When we select the initial value $(x, y, z) = (1, -2.75, -3)$, the phase diagram and Poincaré cross-sectional diagram drawn are shown in Fig.4 and Fig.5. It can be seen that the system (1) is in a state of chaos.

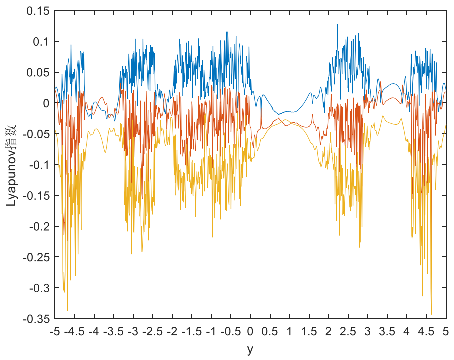


Fig.2 Lyapunov exponent diagram with y_0

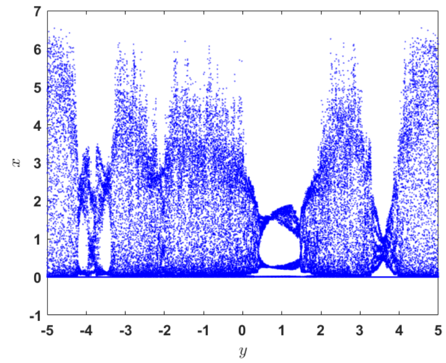


Fig.3 Bifurcation diagram

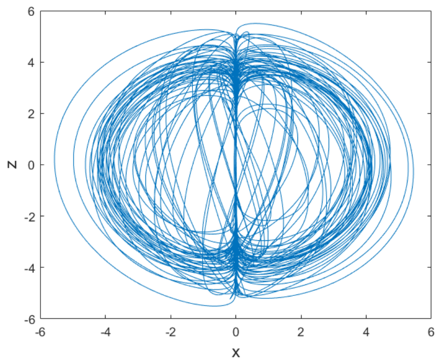


Fig.4 Phase diagram

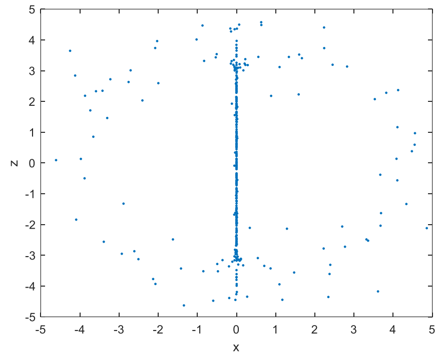


Fig.5 Poincaré cross-sectional diagram

We further analyze system (1), when selecting different initial values, we find hidden attractors in the system. For example, when we select the initial value $(x, y, z) = (1, -4, -3)$, $(1, 1, -3)$, $(1, 3.75, -3)$, the system will have three quasi-periodic attractors in different states as shown in Fig.6, which are represented by red, blue and green lines respectively. When we select $(x, y, z) = (1, -4.5, -3)$, $(1, -1, -3)$, the system will show two quasi-periodic attractors in different states as shown in Fig.7, which are represented by blue and red lines respectively. When we select $(x, y, z) = (1, 0.76, -3)$, $(1, 3.64, -3)$ is selected, the system will show two different states of periodic attractors as shown in the figure, which are represented by blue and red lines respectively. The above numerical analysis fully shows that when a≠b, system (1) has multi-stability.

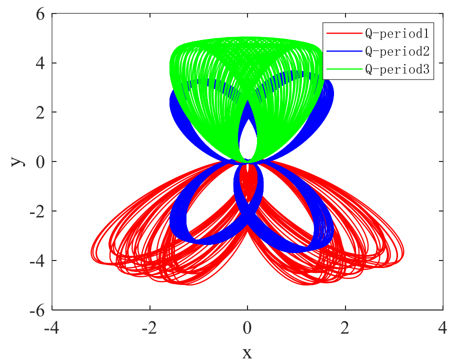


Fig.6 Three quasi-periodic attractors

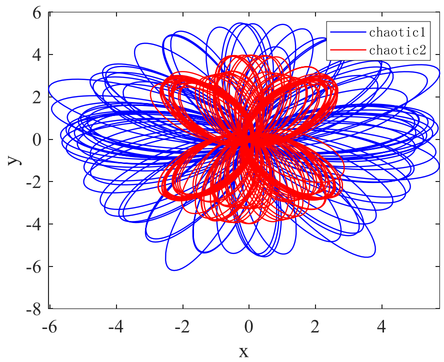


Fig.7 Two chaotic attractors

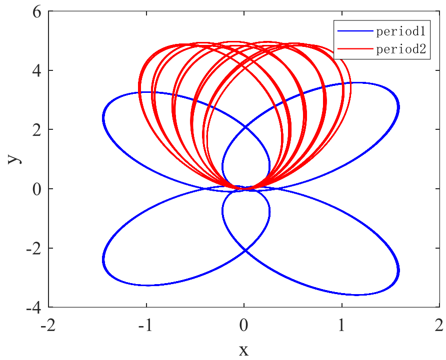


Fig.8 Two period attractors

3.2 a=b

When we select the initial value of the system (1) $(x,y,z)=(x_0,1,1)$, the parameters a=b=1, where x_0 varies from -5 to 5, and the Lyapunov exponent diagram and bifurcation diagram that vary with x_0 are shown in Fig.9 and Fig.10. It can be seen that when $x_0 \in (-5,-0.5) \bigcup (0.5,5)$, the maximum Lyapunov exponent of the system is greater than 0, and the system is in a chaotic state; when $x_0 \in (-0.5,0.5)$, the motion state of the system changes between cycles and counter cycles.

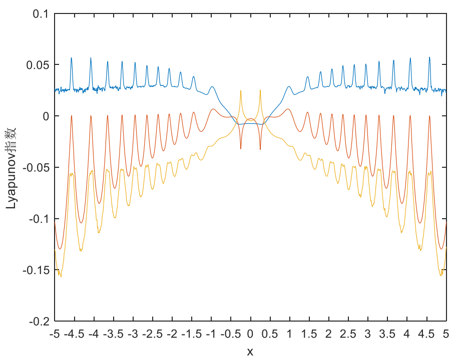


Fig.9 Lyapunov exponent diagram with x_0

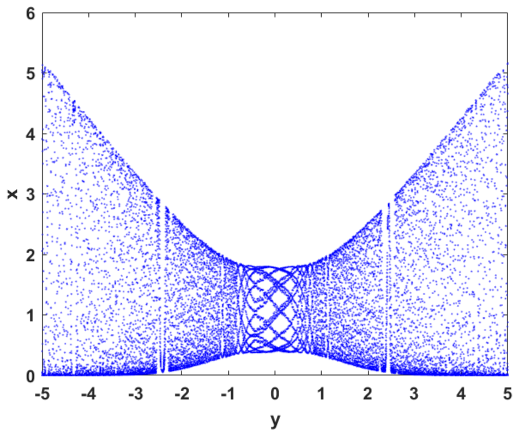
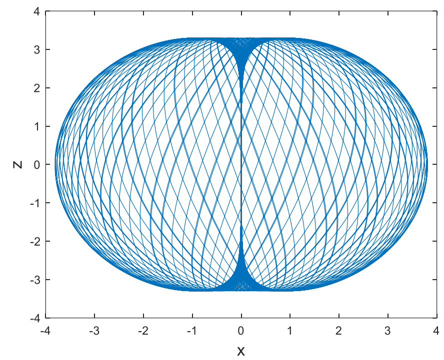
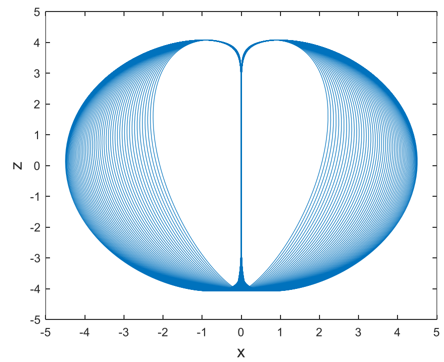


Fig.10 Bifurcation diagram

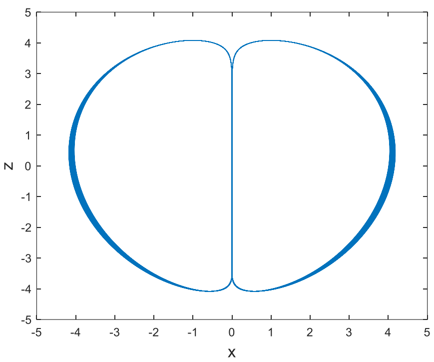
Then we further study the multi-stability of system (1) in this state. Select the initial value $(x,y,z)=$ $(2,-1.45,-2.75)$, $(2,-1.85,-3.5)$, $(2,-3.8,1.45)$, respectively, and system (1) will appear chaotic, quasi-periodic, and periodic attractors, as shown in Fig.11. Through the above numerical analysis, when a=b, system (1) exhibits multiple stability phenomena.



(a) Phase diagram with $(x,y,z)=(2,-1.45,-2.75)$



(b) Phase diagram with $(x,y,z)=(2,-1.85,-3.5)$



(c) Phase diagram with $(x, y, z) = (2, -3.8, 1.45)$

Fig.11 System phase diagram under different initial values

4. FPGA implementation of the three-dimensional conservative chaotic system

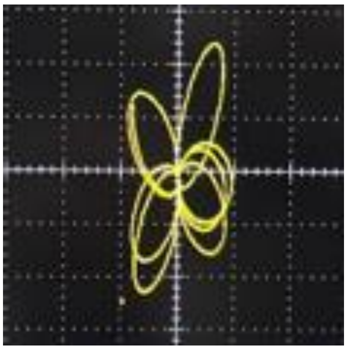
The FPGA hardware device used in this article is the DE2i-150 development board. First, we model and design system (1) and build its circuit structure model.Then the circuit model is converted into the corresponding FPGA hardware description language VHDL through Signal Compiler, for synthesis, compilation, adaptation and simulation. Finally, download the generated file to the FPGA hardware development board, and observe the image by debugging the oscilloscope.

When designing the circuit model, since system (1) is a continuous-time chaotic system, it needs to be discretized and converted into a digital circuit, which is built through the MATLAB/Simulink library. The equation discretized by Euler algorithm is:

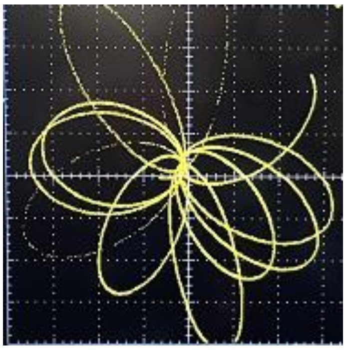
$$\begin{cases} x(n+1) = x(n) + \Delta T[x(n)z(n) + ay(n)] \\ y(n+1) = y(n) + \Delta T[y(n)z(n) - bx(n)] \\ z(n+1) = z(n) + \Delta T[1 - x(n)x(n) - y(n)y(n)] \end{cases} \quad (4)$$

among them, ΔT is the sampling time; $x(n)$, $y(n)$, $z(n)$ are the iterative sequence in the current state; $x(n+1)$, $y(n+1)$, $z(n+1)$ are the iterative sequence in the next cycle state.

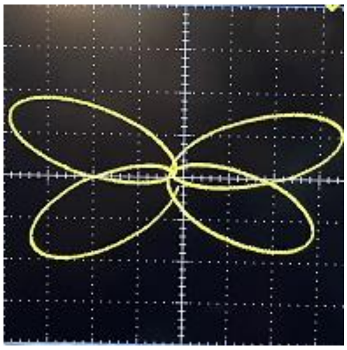
Take the parameters a=1, b=1.6, the initial value $(x, y, z) = (1,1,-3)$, $(1,-1,-3)$, $(1,0.76,-3)$, and the phase diagram observed by the oscilloscope is shown in the Fig.12(a)-(c). It can be found from Fig.12 that the FPGA hardware implementation results are consistent with the numerical analysis results, which further verifies the physical feasibility and multi-stability of system (1).



(a) Phase diagram with $(x, y, z) = (1,1,-3)$



(b) Phase diagram with $(x, y, z) = (1,-1,-3)$



(c) Phase diagram with $(x, y, z) = (1,0.76,-3)$

Fig.12 When taking different initial values and parameters, the system phase diagram observed by the oscilloscope

5. Conclusion

By analyzing the mechanical properties of the three-dimensional conservative chaotic system proposed by Vaidyanathan and Volos, the reasons for the chaos of the system are revealed. Then by changing the system parameters and initial values, it is found that the system has rich dynamic behaviors, including periodic characteristics, quasi-periodic characteristics, and chaotic characteristics. In addition, the rich dynamic characteristics of the three-dimensional system are verified from the perspectives of numerical analysis and FPGA implementation, and the FPGA hardware experimental results are consistent with the numerical analysis results, which further demonstrates the physical feasibility of the system in a physical sense.

References

1. Hénon M, Helles C. The applicability of the third integral of motion: Some numerical experiments. *Astrophys J*, 1964, 69: 73-79
2. Lakshmanan M, Rajasekar S. Nonlinear Dynamics - Integrability, Chaos, and Patterns. *Springer-Verlag Berlin Heidelberg*, 2003.
3. Sprott J C. Some simple chaotic flows. *Phys Rev E*, 1994, 50(2): 647-650.
4. Hoover W G. Remark on "some simple chaotic flows". *Phys Rev E*, 1995, 51(1):759–760.
5. Posch H A, Hoover W G, Vesely F J. Canonical dynamics of the Nosé-oscillator: stability, order, and chaos. *Phys Rev A*, 1986, 33: 4253-4265.
6. Hoover W G. Canonical dynamics: equilibrium phase-space distributions. *Phys Rev A*, 1985, 31(3):1695-1697.
7. Sprott J C. Some simple chaotic jerk functions. *Am J Phys*, 1997, 65: 537-543.
8. Thomas R. Deterministic chaos seen in terms of feedback circuits: Analysis, synthesis,`labyrnth chaos'. *Int J Bifurcat Chaos*, 1999, 9: 1889-1905.
9. Vaidyanathan S, Volos C. Analysis and adaptive control of a novel 3-D conservative no-equilibrium chaotic system. *Arch Control Sci*, 2015, 25(3): 333-353.
10. Mahmoud G M, Ahmed M E. Analysis of Chaotic and hyperchaotic conservative complex nonlinear systems. *Miskolc Math. Notes*, 2017, 18: 315-326.
11. Singh J P, Roy B K. Five new 4-D autonomous conservative chaotic systems with various type of non-hyperbolic and lines of equilibria. *Chaos Soliton Fract*, 2018, 114: 81-91.
12. Vaidyanathan S, Pakiriswamy S. A 3-D Novel Conservative Chaotic System and its Generalized Projective Synchronization via Adaptive Control. *J Eng Sci Technol Rev*, 2015, 8(2):52-60.
13. Cang shijian, Wu A G, Wang Z H, et al. On a 3-D generalized Hamiltonian model with conservative and dissipative chaotic flows. *Chaos Soliton Fract*, 2017, 99: 45-51.
14. Gugapriya G, Rajagopal K, Karthikeyan A, et al. A family of conservative chaotic systems with cyclic symmetry. *Pramana*, 2019, 92(4):48.
15. Cang shijian, Li Y, Xue Wei, et al. Conservative chaos and invariant tori in the modified Sprott A system. *Nonlinear Dyn*, 2020, 99(2): 1699-1708.
16. Xian Y J, Xia C, Guo T T, et al. Dynamical analysis and FPGA implementation of a large range chaotic system with coexisting attractors. *Results Phys*, 2018, 11: 368-376.
17. Li C B, Sprott J C. Multistability in the Lorenz System: A Broken Butterfly. *Int J Bifurcat Chaos*, 2014, 24: 1450131, 7pages.
18. Bao B C, Hu A H, Bao H, et al. Three-Dimensional Memristive Hindmarsh-Rose Neuron Model with Hidden Coexisting Asymmetric Behaviors. *Complexity*, 2018, Article ID 3872573, 11 pages.
19. Jia Hongyan, Shi W X, Qi G Y. Coexisting attractors, Energy Analysis and Boundary of Lü System, *Int J Bifurcat Chaos*, 2020, 30(3): 2050048, 12 pages.
20. Wang F P and Wang F Q. Multistability and coexisting transient chaos in a simple memcapacitive system. *Chin. Phys. B*, 2020; 29(5): 058502.
21. Zhang S, Zeng Y H, Li Z J, et al. Generating one to four-wing hidden attractors in a novel 4D no-equilibrium chaotic system with extreme multistability. *Chaos*, 2018, 28: 013113, 11 pages.
22. Bao B, Bao H, Wang N, Chen M, Xu Q. Hidden extreme multistability in memristive hyperchaotic system. *Chaos Soliton Fract* 2017;94:102–111.
23. Chen M, Sun M, Bao B, Wu H, Xu Q, Wang J. Controlling extreme multistability of memristor emulator-based dynamical circuit in flux–charge domain. *Nonlinear Dyn* 2018;91(2):1395–1412.
24. Bao H, Jiang T, Chu K, Chen M, Xu Q, Bao B. Memristor-based canonical Chua's circuit: extreme multistability in voltage-current domain and its controllability in flux-charge domain. *Complexity* 2018; 2018(5935637):1-13.
25. Jafari, S., Sprott, J.C., Dehghan, S.: Categories of conservative flows. *Int. J. Bifurc. Chaos* 29(02), 1950021 (2019).
26. Singh, J.P., Roy, B.K.: Five new 4-D autonomous conservative chaotic systems with various type of non-hyperbolic and lines of equilibria. *Chaos Solitons Fractals* 114, 81–91(2018).

Authors Introduction

Mr. Minghan Song

He received his bachelor's degree in building Electrical and Intelligent engineering from Shandong Jianzhu University, China in 2020. He is studying for a master's degree at Tianjin University of Science and Technology, China.

A New Hyperchaotic Financial System

Lei Gong

Tianjin University of Science and Technology, 300222, China

*E-mail: * 1442064639@qq.com*

www.tust.edu.cn

Abstract

In this paper, a new hyperchaotic financial system is obtained based on a financial system. It is first transformed into Kolmogorov model, which is composed of conservative torque, dissipative torque and external torque, to study the reason why the new system can generate chaos. Then, by studying energy exchange and combining different torques, dynamics of the new system is analyzed, the external torque is found to be the main reason the new system generate chaos. The paper provides a new method of analyzing chaotic dynamics in financial system, and further promotes new strategies are found to control chaos in financial market.

Keywords: hyperchaos, Kolmogorov model, Hamiltonian energy, financial market

1. Introduction

Nowadays, the economic society is becoming more and more complex and full of uncertainty. In 1980, Stutzer, an American economist, for the first time revealed the chaotic phenomenon of macroeconomic system in the economic growth equation in reference [1]. This made people realize the limitations of the economic model based on the traditional economic theory, and applied the chaotic model to economics. In 2007, the U.S. subprime mortgage crisis triggered the global economic crisis, which once again showed the existence of butterfly effect and chaos in the financial system. This has attracted extensive attention of researchers [2-8].

As we all know, most physical models are open systems [9-11], which can dissipate energy, generate energy, store energy and exchange energy with the external environment. Financial system is a complex nonlinear system, which is composed of many elements. It is open and far from equilibrium. In this nonlinear system, there is energy exchange. Therefore, chaos is a very common phenomenon. Since the discovery of chaos in economics, it has had a great impact on the famous economics. So far, there are few research results on the mechanics and energy analysis of hyperchaotic financial system.

In this paper, a new dynamic model of hyperchaotic financial system is constructed and studied. The Hamiltonian energy of the system under different torque combinations is studied by numerical method and MATLAB. It is found that the external torque is an important reason for the chaotic behavior of the system. The result analysis is helpful to predict and solve some crisis problems in the economic market, and has important theoretical and practical significance in the financial field.

2. A New Hyperchaotic Financial System Model

References [12] reported that a hyperchaotic financial system model consists of four state variables: the interest rate x, the investment demand y, the price exponent z, and the average profit margin u. The hyperchaotic finance system is given as follows:

$$\begin{cases}\dot{x}=z+(y-a)x+u\\ \dot{y}=1-by-x^2\\ \dot{z}=-x-cz\\ \dot{u}=-dxy-ku\end{cases} \tag{1}$$

Based on system (1), the change of interest rate is not only affected by the average profit margin, but also easily affected by the interaction between investment cost and profit margin. A new hyperchaotic financial system is found, which more accurately reflects the actual financial market. The model is as follows:

$$\begin{cases}\dot{x}=z+(y-a)x+duy\\ \dot{y}=1-by-x^2\\ \dot{z}=-x-cz\\ \dot{u}=-dxy-ku\end{cases} \tag{2}$$

Where a is the saving, b is the per investment cost, c is the elasticity of demands of commercials, and they are positive constants.

3. Transformation of Hyperchaotic Financial System

Every parameter and state variable of financial system has practical significance. Studying and analyzing their impact on chaotic behavior is of great value to the prevention of financial risk. Convert system (2) into Kolmogorov form [13-15]:

$$\dot{X}=\{X,H\}-\Lambda X+f$$
$$=\begin{pmatrix}0 & x & 1 & dy\\ -x & 0 & 0 & 0\\ -1 & 0 & 0 & 0\\ -dy & 0 & 0 & 0\end{pmatrix}\begin{pmatrix}x\\ y\\ z\\ u\end{pmatrix}-\begin{pmatrix}ax\\ by\\ cz\\ ku\end{pmatrix}+\begin{pmatrix}0\\ 1\\ 0\\ 0\end{pmatrix} \tag{3}$$

Generally, $\{X,H\}$ refers to the conserved parts of (3), it includes the inertial torque and the internal torque. ΛX and f represents the dissipative part and the external torque of (3), respectively. Furthermore, there are three torques in (2): the conservative torque, the dissipative torque and the external torque.

4. Torque and Energy Analysis

4.1. The conservative torque

When there is only conservative moment, the system equation is:

$$\begin{cases}\dot{x}=xy+z+duy\\ \dot{y}=-x^2\\ \dot{z}=-x\\ \dot{u}=-dxy\end{cases} \tag{4}$$

Eq. (4) is a conservative system only affected by the inertial torque and the internal torque, and its corresponding Hamiltonian energy is a constant and $\dot{H}=0$. Which means the Hamiltonian energy is not exchanged with the dissipative energy and external supplied energy. When the initial value is $(1,2,0.5,0.5)$, the $H=2.75$. The system state variables and Hamiltonian energy are shown in Fig.1. All these results from numerical analysis further verify system under the conservative torque is conservative. Therefore, it can be obtained that the system can't generate chaos only under the conservative torque, and the corresponding Hamiltonian energy is invariable.

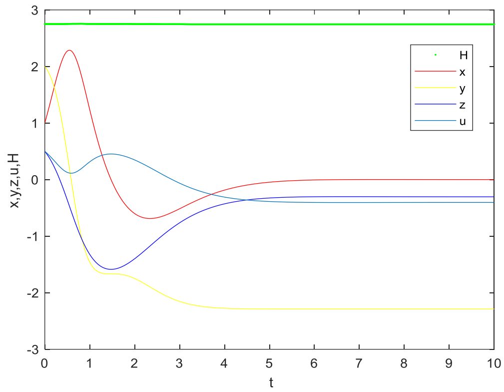


Fig.1 State variables and Hamiltonian energy of system (4)

In this case, chaos is not found to exist in the system. The interest rate, price index and average profit margin tend to zero, indicating that the economic market has lost its vitality.

4.2. The conservative torque and the dissipative torque

Under both the conservative torque and the dissipative torque, the system equation is:

$$\begin{cases}\dot{x}=z+(y-a)x+duy\\ \dot{y}=-by-x^2\\ \dot{z}=-x-cz\\ \dot{u}=-dxy-ku\end{cases} \tag{5}$$

At this time, the Hamiltonian energy function of the system (5) is selected as the Lyapunov function. Hamiltonian derivative is

$\dot{H} = -(ax^2 + by^2 + cz^2 + ku^2) \leq 0$. According to Lyapunov stability theory, the system (5) is asymptotically stable. The state variable converges to zero. Because there is no external input, the system energy dissipates, and the phase space volume shrinks to zero over time. The state variables and Hamiltonian energy of the system (5) are shown in Fig.2. It can be seen that the interest rate, investment cost, price index, average profit margin and Hamiltonian energy all tend to zero. The economic market is paralyzed and completely lost its vitality.

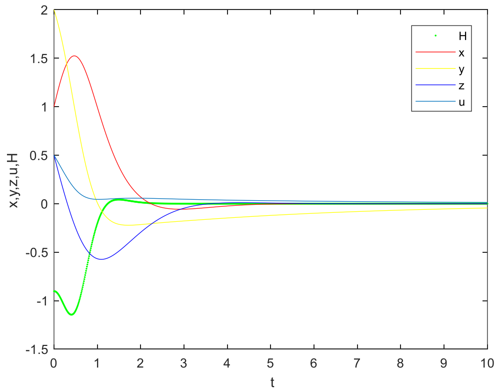


Fig.2 State variables and Hamiltonian energy of system (5)

4.3. The conservative torque and the external torque

Under both the conservative torque and the external torque, the system equation is:

$$\begin{cases} \dot{x} = xy + z + duy \\ \dot{y} = 1 - x^2 \\ \dot{z} = -x \\ \dot{u} = -dxy \end{cases} \tag{6}$$

At this time, the Hamiltonian energy derivative is $\dot{H} = y$, Hamiltonian energy is $H = \frac{1}{2}y^2$, Hamiltonian energy and its derivatives are shown in Fig.3. Investment demand is also affected by interest rate, price index and average profit margin. Hamiltonian energy will change and alternate with the irregular movement of investment cost. Therefore, it is chaotic in a limited area, as shown in Fig.4.

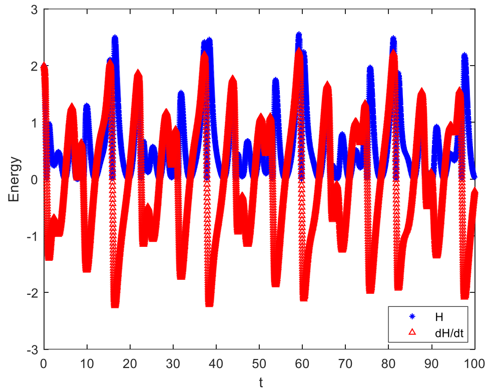


Fig.3 Hamiltonian energy and its derivative of system (6)

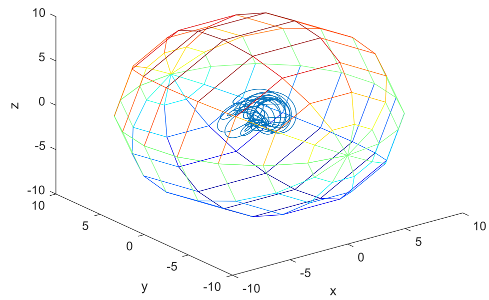


Fig.4 Phase diagram in finite region of system (6)

The external torque f is equivalent to an external excitation, which adds the investment rate 1 to the system, increases the dynamic behavior of the system. This makes the state variable no longer tend to zero, and the system appears chaos, which means that it adds vitality to the economic market. It can be seen that external incentive is essential to market economy, and it is also the main reason for the chaos of financial market.

4.4.The full torque

Under the full torque, the equation is system (3). At this time, the Hamiltonian energy derivative $\dot{H} = -(ax^2 + by^2 + cz^2 + ku^2) + y$ is shown in Fig.5. When the system parameter is $a = 0.9$, $b = 0.2$, $c = 1.5$, $d = 0.33$, $k = 0.17$, and the initial value point is selected $(1, 2, 0.5, 0.5)$, the Hamiltonian energy distribution of the system (3) is shown in Fig.6. Different colors represent different Hamiltonian energies, and the color from yellow to dark blue represents the energy from high to low, indicating the energy change of the attractor. Dissipative torque dissipates energy, and external torque promotes energy absorption. The two interact, and their energy changes are disordered, and the Hamiltonian energy derivative is high and low.

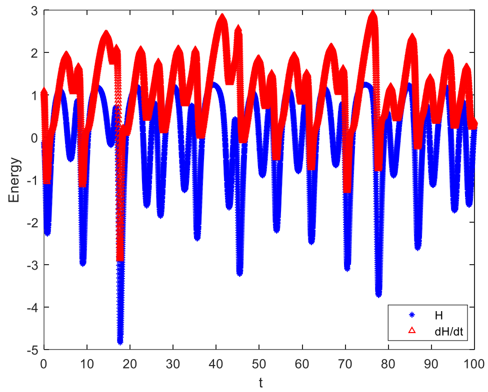


Fig.5 Hamiltonian energy and its derivative of system (3)

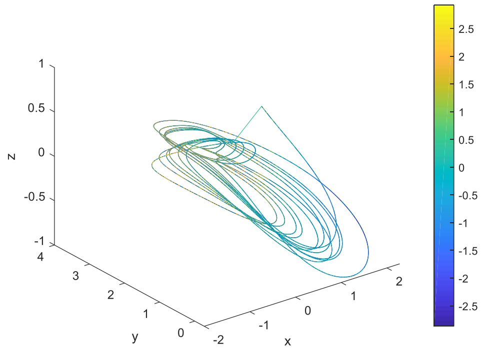


Fig.6 Chaos attractor with Hamiltonian energy

By further analyzing, we find that the interaction between the external torque and other torques is main reason why hyperchaotic financial system can generate chaos and its dynamics characteristics is related to the variation of Hamiltonian energy.

5. Conclusion

Based on the original financial system and combined with the actual situation, a new hyperchaotic financial system is established. In addition, by transforming the system into Kolmogorov type system, the dynamic characteristics of the system are analyzed. It is found that the interaction between external torque and other torques is the main reason for the chaos of the system. At the same time, the influence of energy change on dynamic characteristics is studied by energy analysis. Finally, combined with the practical significance of various parameters and variables, this paper makes a reasonable explanation for various phenomena, which is of great value to the regulation of financial market.

References

1. Frey, S Bruno, Stutzer, et al., Happiness, Economy and Institutions. 2000, 7: pp.4-6.
2. Mei, X.H., Yu, J.N., Zhang, J.G.: The Linear and Nonlinear Feedback Synchroniza-tion of a Class of Chaotic Finance System. *Journal of Tianjin Normal University(Natural Science Edition)* 28(3), 49–51 (2008)
3. Huang, D.S., Li, H.Q.: Theory and Method of the Nonlinear Economics. Publishing House of Sichuan University, Chengdu (1993)
4. Fanti, L., Manfredi, P .: Chaotic business cycles and fiscal policy: An IS-LM model with distributed tax collectionlags. Chaos Solitons Fractals 32(2), 736–744 (2007)
5. Yu, P., Liao, X.X.: New Estimates for Globally Attractive and Positive Invariant Set of the Family of the Lorenz. *Int. J. Bifurcation and Chaos* 16(11), 3383–3390(2006)
6. Chian, A.C.L, Rempel, E.L., Rogers: Complex economic dynamics: Chaotic saddle, crisis and intermittency.Chaos Solitons Fractals 29(5), 1194–1218 (2006)
7. Guégan, D.: Chaos in economics and finance. *Annu. Rev. Control* 33, 89–93 (2009).
8. Igor, E., Michael, T.: Dynamic interaction models of economic equilibrium. *Dyn. Control* 33, 166–182
9. G. Chen, T. Ueta, Yet another chaotic attractor. *Int. J. Bifurcat, Chaos* 9(7), 1465 (1999),
10. J. Lu, G. Chen, A new chaotic attractor coined. *Int. J. Bifurcat, Chaos* 12(3), 659 (2002)
11. G. Qi, G. Chen, M.A.V . Wyk, et al, A four-wing chaotic attractor generated from a new 3-D quadratic autonomous system. *Chaos Soliton. Fract*. 38(3), 705 (2008)
12. Haojie Yu, Guoliang Cai, Yuxiu Li, Dynamic analysis and control of a new hyperchaotic finance system. *Nonlinear Dyn* (2012) 67:2171–2182.
13. V. Arnold, Kolmogorov's hydrodynamic attractors. *Proc. Soc, R., Lond.* A 434, 19 (1991).
14. G. Qi, X. Liang, Mechanical analysis of Qi four-wing chaotic system. *Nonlinear Dyn.* 86(2), 1095 (2016).
15. Arnold, V. Kolmogorov hydrodynamic attractors, *Proc. Roy. Soc. London* 434, 19–22(1991).

Authors Introduction

Mr. Lei Gong

He received his Bachelor of engineering degree from the school of Ludong University in 2019. He is receiving a master's degree in engineering from Tianjin University of science and technology in China.

New Hybridization Algorithm of Differential Evolution and Particle Swarm Optimization for Efficient Feature Selection

Koon Meng Ang, Mohd Rizon Bin Mohamed Juhari, Wei Hong Lim*, Sew Sun Tiang, Chun Kit Ang, Eryana Eiyda Hussin, Li Pan, Ting Hui Chong

Faculty of Engineering, Technology and Built Environment, UCSI University, 1, Jalan Puncak Menara Gading, UCSI Heights, 56000 Cheras, Kuala Lumpur, Malaysia.

E-mail: 1001436889@ucsiuniversity.edu.my, mohdrizon@ucsiuniversity.edu.my, limwh@ucsiuniversity.edu.my, tiangss@ucsiuniversity.edu.my, angck@ucsiuniversity.edu.my, eryanaeiyda@ucsiuniversity.edu.my, 1002060534@ucsiuniversity.edu.my, 1001747667@ucsiuniversity.edu.my

http://www.ucsiuniversity.edu.my/

Abstract

Feature selection is a popular pre-processing technique applied to enhance the learning performances of machine learning models by removing irrelevant features without compromising their accuracies. The rapid growth of input features in big data era has increased the complexities of feature selection problems tremendously. Given their excellent global search ability, differential evolution (DE) and particle swarm optimization (PSO) are considered as the promising techniques used to solve feature selection problems. In this paper, a new hybrid algorithm is proposed to solve feature selection problems more effectively by leveraging the strengths of both DE and PSO. The proposed feature selection algorithm is reported to achieve an average accuracy of 89.03% when solving 7 datasets obtained from UCI Machine Learning Repository.

Keywords: Feature Selection; Particle Swarm Optimization (PSO); Differential Evolution (DE); Metaheuristic Search Algorithm; Hybridization.

1. Introduction

Feature selection[1] is a popular pre-processing technique used to address the "curse of dimensionality" issue by eliminating redundant features from large-scale datasets. Feature selection is widely used in real-world problems due to its ability to reduce storage space and computational time required for training the predictive models without sacrificing their performances[2,3]. Feature selection is formulated as a non-deterministic polynomial-time (NP) hard combinatorial problem that is not trivial to solve, especially when it involves the large input feature size of $D > 100$ [4].

Nature-inspired algorithms have emerged as effective approach to solve the complex real-world optimization problems due to their promising global search ability. Differential evolution (DE)[5] and particle swarm optimization (PSO)[6] are the two most popular nature-inspired algorithms widely used to solve different types of optimization problems[7,8,9,10,11,12,13,14]. However, the capability of conventional DE and PSO to tackle large-scale feature selection problems remains unexplored. The presence of excessive irrelevant features in original datasets can introduce massive number of local optima in search space and increase the complexity of feature selection problem. For conventional DE and PSO, the random initialization scheme adopted do not fully consider any information around their search environments, therefore the quality of initial population obtained is questionable[15]. The "No Free Lunch Theorem"[16] is another factor that restrict the performances of conventional DE and PSO to tackle various optimization problems. More robust optimization algorithms are required to handle the feature selection problems with different complexity levels effectively.

In this article, a hybrid DE and PSO with chaotic-opposition-based initialization scheme (HDPCIS) is proposed to address the aforementioned challenges in performing feature selection. A chaotic-opposition-based initialization scheme (CIS) is first incorporated into

HDPCIS to generate an initial population with better solution quality. A hybridization scheme that aims to leverage the benefits of DE and PSO is also introduced to achieve better tradeoff in terms of search efficiency and diversity preservation. The performance of HDPCIS to solve different feature selection problems are assessed with 7 datasets of UCI Machine Learning Repository.

2. Related Works

2.1. *Conventional DE*

Given a set of randomly initialized DE solution with the population size of N, each n-th solution of $X_n = [X_{n,1}, \ldots, X_{n,d,\ldots,} X_{n,D}]$ represents a candidate solution of a given problem with total dimension size of D, where $d \in [1, D]$ and $n \in [1, N]$ refer to the dimension and solution indices, respectively. During the evolution process of DE, a mutation process is first performed using "DE/rand/1" strategy to generate a donor vector $U_n = [U_{n,1}, \ldots, U_{n,d,\ldots,} U_{n,D}]$ for each n-th target vector $X_n = [X_{n,1}, \ldots, X_{n,d,\ldots,} X_{n,D}]$ as follow:

$$U_n = X_a + F\left(X_b - X_c\right) \quad (1)$$

where F is a scaling factor in range of 0 to 1; X_a, X_b and X_c are three randomly selected solutions from population with $n \neq a \neq b \neq c$.

An offspring vector $Y_n = [Y_{n,1}, \ldots, Y_{n,d,\ldots,} Y_{n,D}]$ is then produced for each n-th solution by performing crossover on the target vector $X_n = [X_{n,1}, \ldots, X_{n,d,\ldots,} X_{n,D}]$ and donor vector $U_n = [U_{n,1}, \ldots, U_{n,d,\ldots,} U_{n,D}]$. Define $C_r \in [0.5, 1]$ as the crossover probability, the d-th dimension of each n-th offspring vector can be computed as:

$$Y_{n,d} = \begin{cases} U_{n,d}, & \text{if } rand_d \le Cr \\ X_{n,d}, & \text{otherwise} \end{cases} \quad (2)$$

For selection process, the fitness value of Y_n is compared with that of X_n in terms of their fitness values as a selection process. The latter solution is replaced if the former one has better fitness value.

2.2. *Conventional PSO*

Each n-th candidate solution or particle consists of two vectors define its current state, i.e., the velocity $V_n = [V_{n,1}, \ldots, V_{n,d,\ldots,} V_{n,D}]$ and position $X_n = [X_{n,1}, \ldots, X_{n,d,\ldots,} X_{n,D}]$, where $d \in [1, D]$ and $n \in [1, N]$. Each PSO particle can memorize the best solution found by itself and population that are denoted as $P_n^{best} = \left\lfloor P_{n,1}^{best}, \ldots, P_{n,d}^{best}, \ldots, P_{n,D}^{best} \right\rfloor$ and $G^{best} = \left\lfloor G_1^{best}, \ldots, G_d^{best}, \ldots, G_D^{best} \right\rfloor$, respectively. The velocity V_n and position X_n of each particle n are updated as:

$$V_n^{new} = \omega V_n + c_1 r_1 \left(P_n^{best} - X_n\right) + c_2 r_2 \left(G^{best} - X_n\right) \quad (3)$$

$$X_n^{new} = X_n + V_n^{new} \quad (4)$$

where ω js inertia weight; c_1 and c_2 are acceleration coefficients; r_1 and r_2 are two random numbers obtained from uniform distribution in range of 0 to 1. The fitness value of X_n is compared with those of P_n^{best} and G^{best}. The latter two solutions are replaced if the former one has better fitness value.

2.3. *Feature Selection Problem*

Feature selection is considered as a bi-objective optimization problem, aiming to minimize the number of selected features and maximize the classification accuracy, simultaneously. In order to satisfy these objectives, a fitness function is formulated to measure the quality of each candidate solution as follow[17]:

$$f(\bullet) = \chi\varepsilon + \gamma \frac{|F_S|}{|F_T|} \quad (5)$$

where $\chi \in [0,1]$ and $\gamma = (1-\chi)$ refer to the parameters measuring the weightage of classification quality and subset length, respectively; ε represents the classification error; $|F_S|$ and $|F_T|$ indicate the selected subset of features and the total number of features in original dataset, respectively.

3. The Proposed HDPCIS

At the beginning of search process, a chaotic-opposition-based initialization scheme (CIS)[8] is incorporated to replace random initialization scheme. A chaotic swarm Ψ^{CS} and an opposite swarm Ψ^{OS} are produced by CIS based on the modified sine map and opposition-based-learning strategy, respectively. Ψ^{CS} and Ψ^{OS} are then combined to form a merged population Ψ^{M}. After all the solution members of Ψ^{M} are sorted from the worst to the best based on their fitness values, the first best N members of Ψ^{M} are selected to construct the initial population $\Psi^{I} = \left[X_1, \ldots, X_n, \ldots, X_N\right]$ of HDPCIS.

For the proposed hybridization scheme, DE and PSO are employed as the primary and secondary algorithms used to evolve candidate solutions, respectively. During the DE stage, a mutation scheme of Eq. (1) is performed to generate a donor vector ∂_n for each solution n. The corresponding offspring vector $\wp_n$ is computed based on Eq. (2). A greedy selection scheme is applied to compare the fitness of $\wp_n$ with that of X_n. The latter solution is replaced by the former solution if the former solution has better fitness value. The greedy selection scheme is also used to update the best solution in population G^{best}.

If $\wp_n$ computed by DE has better fitness than original X_n, PSO is triggered as secondary optimizer to refine X_n. In PSO stage, the velocity component of solution n is updated as follow:

$$V_n^{new} = \omega V_n + c\left(G^{best} - X_n\right) \tag{6}$$

where c refers to acceleration coefficient. Notably, the velocity update equation of HDPCIS in Eq. (6) only considers social component. For HDPCIS, X_n is essentially equivalent to self-cognitive component P_n^{best} because it only updated when a better solution is found in optimization process. Given V_n^{new}, the new position X_n^{new} is calculated with Eq. (4). Both of X_n and G^{best} are replaced by X_n^{new}, if the latter solution is more superior than the former solutions. Otherwise, X_n^{new} with worse fitness will be discarded.

The overall framework of HDPCIS is summarized in Fig. 1. The optimization process is iterated until the termination criteria $\tau > \tau^{\max}$ is satisfied, where τ and $\tau^{\max}$ represent the fitness evaluation counter and the predefined maximum fitness evaluation number.

Algorithm: HCPCIS

Inputs: *D, N, Ub, Lb Cr, F,* ω, c, τ, $\tau^{\max}$

Initialize $\tau \leftarrow 0$;
Produce $\Psi^I = [X_1, ..., X_n, ..., X_N]$ using CIS;
$\tau \leftarrow \tau + 2N$;
while $\tau \le \tau^{\max}$ **do**
  for each solution n **do** *//Execute DE.*
    Produce ∂_n using Eq. (1);
    Produce $\wp_n$ using Eq. (2);
    Evaluate fitness of $\wp_n$ using Eq. (5);
    $\tau \leftarrow \tau + 1$;
    Update X_n and G^{best} with greedy selection;
    if $f(X_n) < f(\wp_n)$ **then** *// Execute PSO*
      Calculate V_n^{new} using Eq. (6);
      Calculate X_n^{new} using Eq. (4);
      Evaluate fitness of X_n^{new} using Eq. (5);
      $\tau \leftarrow \tau + 1$;
      Update X_n and G^{best} with greedy selection;
    end if
  end for
end while

Outputs: G^{best}

Fig. 1. Pseudocode of HDPCIS.

4. Performance Evaluations of HDPCIS

4.1. *Simulation settings*

The performance of HDPCIS to solve feature selection problem is evaluated using seven datasets obtained from the UCI Machine Learning Repository[18], i.e., (a) glass identification, (b) lymphography, (c) lung cancer, (d) multiple features, (e) statlog (heart), (f) ionosphere and (g) iris. The proposed HDPCIS is compared with four peer algorithms known as: chaotic-opposition-based hybridized DE with PSO (CO-HDEPSO)[8], chaotic-opposition-based differential evolution (CO-DE), conventional DE (DE)[5] and conventional PSO (PSO)[6], in terms of the mean accuracy Acc^{mean} and average numbers of selected features nF^{avg}. The population size and maximum fitness evaluations numbers of all algorithms are set as N = 10 and $\tau^{\max} = 1000$, respectively. All algorithms are simulated for 30 times to solve the selected datasets.

Table 1. Mean accuracy Acc^{mean}

Datasets	HDPCIS	CO-HDEPSO	CO-DE	DE	PSO
(a)	**0.7952**	0.7762	0.7619	0.7143	0.7286
(b)	**0.5725**	0.5724	0.5448	0.4621	0.5448
(c)	**1.0000**	**1.0000**	**1.0000**	0.9600	0.6400
(d)	**0.9825**	0.9730	0.9715	0.9705	0.9695
(e)	**0.9074**	0.8926	0.8667	0.7852	0.8556
(f)	**0.9743**	0.9714	0.9543	0.9457	0.9114
(g)	**1.0000**	**1.0000**	**1.0000**	**1.0000**	0.9667

Table 2. Average number of selected features nF^{avg}.

Datasets	HDPCIS	CO-HDEPSO	CO-DE	DE	PSO
(a)	**3.8**	4.8	4.8	**3.8**	5.4
(b)	**4.2**	7.0	4.8	6.0	8.4
(c)	**12.4**	14.8	12.8	14.2	21.2
(d)	**273.4**	308.2	295.2	285.6	308.6
(e)	**3.2**	5.0	4.4	4.4	4.8
(f)	**9.6**	10.4	10.8	11.0	14.4
(g)	**1.0**	**1.0**	2.0	**1.0**	**1.0**

4.2. *Comparisons between selected algorithms*

The Acc^{mean} and nF^{avg} values obtained by all algorithms in solving 7 selected image datasets are reported in Tables 1 and 2, respectively. The best and second-best results are indicated by boldface and underlined fonts, respectively. Table 1 reported that the HDPCIS has the best feature selection performance by producing the Acc^{mean} values for seven datasets. This is followed by CO-HDEPSO, CO-DE, DE, and PSO that produce the best Acc^{mean} values in 2, 2, 1, and 0 dataset, respectively. On the other hand, Table 2 reported that the proposed HDPCIS has the best performance in selecting optimal number of features by producing best nF^{avg} for all seven datasets. This is followed by DE, CO-HDEPSO, PSO, and CO-DE that obtain the best nF^{avg} values for 2, 1, 1, and 0 datasets, respectively.

5. Conclusions

A new hybridization algorithm of HDPCIS is introduced to solve feature selection problem effectively. A CIS module is first employed to initialize a population with better quality to reduce the possibility of premature convergence. A hybridization scheme is designed with DE as the primary algorithm and PSO as the secondary algorithm, to achieve better balancing of exploration and exploitation search behaviors. The simulation studies reported that the proposed HDPCIS can outperform its peer algorithms by solving feature selection problems with higher accuracy and lesser 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 Proj-FRGS/1/2019/TK04/UCSI/02/1 and Proj-FRGS/1/2020/TK0/UCSI/02/4.

References

1. B. Xue, M. Zhang, W. N. Browne, and X. Yao, "A survey on evolutionary computation approaches to feature selection," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 4, pp. 606-626, 2015.
2. J. Yang and V. Honavar, "Feature subset selection using a genetic algorithm," in *Feature extraction, construction and selection*: Springer, 1998, pp. 117-136.
3. F. Hafiz, A. Swain, N. Patel, and C. Naik, "A two-dimensional (2-D) learning framework for particle swarm based feature selection," *Pattern Recognition*, vol. 76, pp. 416-433, 2018.
4. H. Faris *et al.*, "An efficient binary salp swarm algorithm with crossover scheme for feature selection problems," *Knowledge-Based Systems*, vol. 154, pp. 43-67, 2018.
5. K. Price, R. M. Storn, and J. A. Lampinen, *Differential evolution: a practical approach to global optimization*. Springer Science & Business Media, 2006.
6. J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, 1995, vol. 4, pp. 1942-1948 vol.4.
7. O. T. Chong, W. H. Lim, N. A. M. Isa, K. M. Ang, S. S. Tiang, and C. K. Ang, "A Teaching-Learning-Based Optimization with Modified Learning Phases for Continuous Optimization," in *Science and Information Conference*, 2020, pp. 103-124: Springer.
8. Z. C. Choi *et al.*, "Hybridized Metaheuristic Search Algorithm with Modified Initialization Scheme for Global Optimization," *Advances in Robotics, Automation and Data Analytics: Selected Papers from ICITES 2020*, vol. 1350, p. 172, 2021.
9. W. S. Koh *et al.*, "Multi-objective Particle Swarm Optimization with Alternate Learning Strategies," Singapore, 2022, pp. 15-25: Springer Singapore.
10. K. M. Ang *et al.*, "Modified Particle Swarm Optimization with Unique Self-cognitive Learning for Global Optimization Problems," Singapore, 2022, pp. 263-274: Springer Singapore.
11. K. M. Ang, W. H. Lim, S. S. Tiang, C. K. Ang, E. Natarajan, and M. Ahamed Khan, "Optimal Training of Feedforward Neural Networks Using Teaching-Learning-Based Optimization with Modified Learning Phases," in *Proceedings of the 12th National Technical Seminar on Unmanned System Technology 2020*, 2022, pp. 867-887: Springer.
12. K. Ang, Z. Yeap, C. Chow, W. Cheng, and W. Lim, "Modified Particle Swarm Optimization with Chaotic Initialization Scheme for Unconstrained Optimization Problems," *MEKATRONIKA*, vol. 3, no. 1, pp. 35-43, 2021.
13. K. M. Ang *et al.*, "A Constrained Teaching-learning-based Optimization with Modified Learning Phases for Constrained Optimization," *Journal of Advanced Research in Dynamical and Control Systems*, vol. 12, no. 04, p. 15, 2020.
14. K. M. Ang, W. H. Lim, N. A. M. Isa, S. S. Tiang, and C. H. Wong, "A constrained multi-swarm particle swarm optimization without velocity for constrained optimization problems," *Expert Systems with Applications*, vol. 140, p. 112882, 2020.
15. X.-S. Yang, "Metaheuristic optimization: algorithm analysis and open problems," in *International Symposium on Experimental Algorithms*, 2011, pp. 21-32: Springer.
16. D. H. Wolpert and W. G. Macready, "No free lunch theorems for optimization," *IEEE transactions on evolutionary computation*, vol. 1, no. 1, pp. 67-82, 1997.
17. E. Emary, H. M. Zawbaa, and A. E. Hassanien, "Binary grey wolf optimization approaches for feature selection," *Neurocomputing*, vol. 172, pp. 371-381, 2016.
18. M. Lichman. UCI Machine Learning Repository [Online]. Available: http://archive.ics.uci.edu/ml/index.php

Authors Introduction

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. 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. 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. Ang Chun Kit

He is an 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.

Ms. Eryana Eiyda Hussin

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. She is also a lecturer from the Department of Electrical and Electronics of UCSI University, Malaysia.

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.

Mr. Ting Hui Chong

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 swarm intelligence and feature selection.

Implementation of LoRa in River Water Quality Monitoring

Syarifah Nabilah Syed Taha@Tahir, Mohamad Sofian Abu Talip, Mahazani Mohamad, Mohamadariff Othman, Tengku Faiz Tengku Mohmed Noor Izam, Mohd Faiz Mohd Salleh
Department of Electrical Engineering, University of Malaya, Kuala Lumpur
Wilayah Persukutuan Kuala Lumpur, 50603, Malaysia[*]
Zati Hakim Azizul Hasan
Department of Artificial Intelligence, University of Malaya, Kuala Lumpur
Wilayah Persukutuan Kuala Lumpur, 50603, Malaysia[†]
Zeeda Fatimah Mohamad
Department of Science and Technology Study, University of Malaya, Kuala Lumpur
Wilayah Persukutuan Kuala Lumpur, 50603, Malaysia[‡]
Amir Feisal Merican Aljunid Merican
Department of Biological Science, University of Malaya, Kuala Lumpur
Wilayah Persukutuan Kuala Lumpur, 50603, Malaysia[§]
E-mail: ananabilah2261@gmail.com, sofian_abutalip@um.edu.my, mahazani@um.edu.my, zati@um.edu.my, zeeda21@um.edu.my, merican@um.edu.my, mohamadariff@um.edu.my, tengkufaiz@um.edu.my, faizsalleh@um.edu.my
www.um.edu.my

Abstract

Emergence of Long Range (LoRa) in network technologies become game changer for Internet of Things (IoT) application. Deployment of LoRa enable IoT application of environment monitoring to cover wide area while maintain at low energy and low cost. Water quality monitoring program was developed to maintain and protect quality of water resources for daily purpose. Also, to prevent pollution and disease epidemic peculiarly during Covid19. This research aimed to build autonomous water quality monitoring prototype implemented with LoRa network for support decision system. The Wireless Sensor Nodes (WSN) that embedded with five type of water quality sensors of pH, turbidity, total dissolved solid (TDS), dissolved oxygen (DO) and temperature linked to single gateway. Water environmentalist able to view the result of timely water quality from mobile application dashboard. Though the performance not severely affected, acquired results revealed non-line of sight condition, transmission power and Spread Factor (SF) value influenced LoRa performance in urban environment. In conclusion, a few improvements on the system grant LoRa high capabilities to be integrated with IoT environment application in urban environment.

Keywords: river conservation, IoT, smart river monitoring, LoRa, CSS.

1. Introduction

Cultivation of IoT from industrial revolution 4.0 contribute in improvement of productivity and efficiency of human life. IoT allow connection digital and physical dimension by exchange data from real gadget over the internet. The system reduces human involvement and able to enhance the accuracy of data collection, processing and analytics (1). The communication between two realms executed by wireless network. Implementation of IoT in industry allow people easily to operate things remotely. It is mostly recognized as a solution for monitoring purposes in Smart Earth technologies. Parallel with Sustainable Development Goals (SDG) 6: Clean Water and Sanitation, target 6.3 aim to protect heath of ecosystem and human from harm of pollution and effect of climate changes in water bodies (2). Hence, it is crucial to monitor the ambience of water status.

River is the primary water resources at Malaysia. It maintains human activities and balancing water ecosystem with its freshwater and nutrient (3). Malaysia practiced Integrated Water Resources Management (IWRM) and Integrated River Basin Management (IRBM) program in water resources management. However, due to poor management, rapid development and population increase demand on sufficient clean water. Moreover, clean water also become critical during pandemic Covid19 to reduce the virus epidemic with frequent sanitization (4). However, there were 47% of river in Malaysia classified polluted. Within year 2020 and 2021, the water bodies at Sungai Gong and Sungai Semenyih were contaminated with hazardous chemical and agriculture sewage forced water treatment plant at Selangor to ceased operation. Thus, unscheduled water supply resulted water shortage to a lot of districts during decontamination of pollution (5).

Implementation of water quality monitoring agenda provides data of qualitative analysis on physical, chemical and biological of water characteristics based on standard of Water Quality Index (WQI) and National Water Quality Standards (NWQS) for Malaysia. Autonomous water quality monitor was invented to support manual water quality monitoring. Lack of the autonomous system can reduce the accuracy of data of water properties that can vary over time (6). Embedding IoT feature in water quality monitoring system can reduce human labor as well become a quick decision system for environmentalist for its real time data.

Addressing function of the system and type of data is important to match with the suitable wireless network. LoRa is a long range network derived from Chirp Spread Spectrum (CSS) technology using air interface. It able to transmit low rate data type within geographical area with its vast radio coverage at low power and low cost (7). This research intent to implement LoRa as the wireless network for continuous water quality monitoring system. This paper contributes on providing evaluation of LoRa network performance to be applied in quick decision system to provide water quality status to environmentalist (8).

2. Literature Review

2.1 *IoT Application in Continuous Water Quality Monitoring with Different Wireless Network*

IWRM program utilize water quality monitoring program at operational level for water resources conservation management. It verifies water bodies status and analyzed the environment impact on water ecosystem. Development of continuous water quality monitoring system completing the observed water characteristics result with manual method (9). Benefited from wireless network, the efficiency of the continuous water quality monitoring system can be improved.

Wireless network acts as communication interface between devices and internet platform over information exchange. Cloud act as data storage for IoT platform processing and analyzing data for view purposes for user. The outline process of the autonomous water quality monitoring displayed as Fig. 2 below.

Fig. 1. IoT Structure for Water Quality Monitoring System

Every type of wireless network has its own capability for its service significant to the applied system. Understand type of network of its capacity, limitation and benefit are essential to meet the requirement of the applied system. Previous researcher had investigated and tested various type of wireless network with water quality monitoring system.

Jiang et al. and Li et al. experimenting water quality monitoring system with PAN type of network, Zigbee as radio frequency model. The telecommunication range was extended from GSM version using GPRS. Reflecting water quality data type in low rate, it unnecessarily consumed higher energy and require external SIM card for communication activation (10). However, this method is suitable with application that employ large bandwidth within short distance. Next, Dilshad et al and Pandi et al and Bhisekar et al deployed the most common practical network used in IoT application, WiFi in water quality monitoring system by. As it extracted large amount of data, it able to improve data accuracy. Though, it utilized higher power for low rate data type due to low latency (11).

Joseph et.al conquer the challenge the high cost on extension subscriptions of GSM network, and send data through Short Message Service (SMS) from its autopilot water quality monitor (12). Collaboration Ericsson between with city officials and university researchers at Stockholm operated NB-IOT to replace 4G LTE network for water quality monitoring system. The experiment success in increase range network scale and proficiency of water sensors (13). Nevertheless, the cost was not minimized as the deployment require expensive reserved frequency or channel.

2.2 *Introduction of LoRa Technology*

LoRa technology is a compilation of two different layer. The physical layer (PHY) employs CSS modulation technique that manipulated allocated bandwidth during signal transmission able to improve LoRa strength against noise and any frequency degradation mechanisms. This technique implements modulation of orthogonal spreading factor (SF) to optimize data rates and each nodes power level. Hence, it increases the efficacy of receiver to decode signals even affected by low signal-to-noise-ratio (SNR) while reduces energy consumption.

Build system LoRaWAN created by LoRa Alliance performed as Media Access Control (MAC), network and application layer at top of LoRa architecture. It generates the communication between nodes and gateway operated in star topology from LoRa PHY modulation to exchange data with host layers. The PHY layer control radio frequency signals within transmission line. SF, bandwidth and coding rate adapt the working environment used to constructs LoRa PHY modulation (14).

This paper referred study of LoRa performance in various IoT application accomplished by previous researchers.

2.3 *LoRa Performance Evaluation from Previous Study*

Utilization of SF integrated with bandwidth and coding rate to ensure the transmission services from concluded data rate by LoRa PHY modulation. The higher SF able to increase transmission distance, processing gain and receiving sensitivity but reduce data rate. Coding rate is utilized as the reliable unit in LoRa to control the amount of Forward Error Correction (FEC) within payload data.

LoRa manage the radio frequency signal within transmission line. Verification of wireless medium characterization on path loss, shadowing and multipath fading are required to reflect condition of non-line of sight (NLOS) condition and multipath propagation signal within urban environment. Hence, it is crucial for signal values gain at IoT platform to be attuned with the sensitivity level of transceivers.

The received power, Prx can be defined as in Eq. (1). The value of Prx is influenced with fusion of the transmitted power, Ptx, Gsystem and Lsystem which both are gains and losses related cables and antennas respectively. It also influenced by losses related at transmission channel, Lchannel that affected by surroundings and the fading margin, M (15).

A few of equation models were derived from various of experimental data to manage the problem of non-specific path loss model. It estimated Lsystem for LoRa in different scenario.

Previous researcher study LoRa network performance on different network properties. Researcher from Switzerland implemented LoRa network for flood early detection and notify user at Twitter. The result showed there were no lose packet between receiver node and water level sensor node with distance of 500 m away with each other (16). While Wotherspoon et al. monitoring wildlife at distance of 5.5. km that powered with 20 dBm and obtained frame loss up to 2% (17).

J. Lousado, S. Antunes, proposed IoT concept for monitor and support elderly people that live by themselves. They found the limitation of the system of LoRa to repeatedly send large amount of data repetitively (18). Buyukakkaslar et al. found Frame Air Times and moderated duty cycle are crucial for real time health basis data tracking (19).

Different application and situation differentiate LoRa performance. Referred to previous study, various of network parameter can be used to analyzed its performance depend on the objective of the application.

3. Methodology

The system process involves from selecting water quality sensor, testing send data to LoRa server, design of water station points, Wireless Sensor Network (WSN) of LoRa nodes properties setup and data analysis of water quality parameter and LoRa performance. The flow of the experiment as summarized in flow chart below

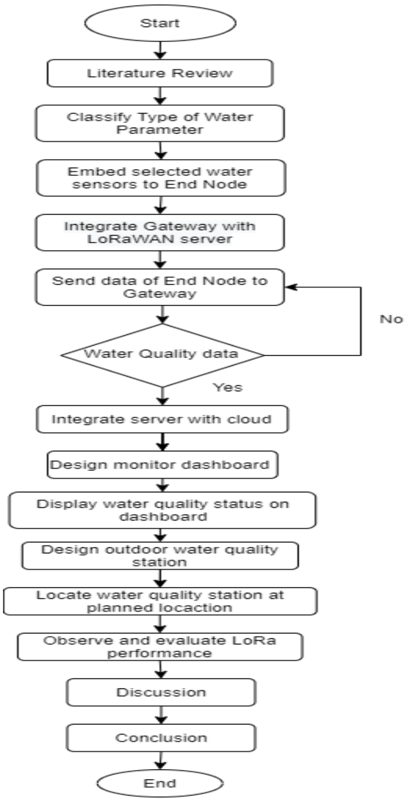


Fig. 2. Process of Water Quality Monitoring Prototype Testing

3.1. *Water Quality Monitor Station Design*

The prototype of water quality monitoring was built with three LoRa nodes at different location of fish pond. LoRa nodes positioned at different height and distance from single LoRa Gateway. The stations supplied up with 5V solar panel. Water sensors were selected based on water characteristics following specified regulations of World Health Organization (WHO). Also, its durability to operate in outdoor environment and datasheet supplied by DFRobot Distributor. Water sensors as listed in Table 1.

Table 1. Water Quality Properties

Water Sensor	Function
pH	Acidification
Turbidity	Water opacity
Temperature	Variable for other parameters
Conductivity	Salinity and total dissolved solid
Dissolved Oxygen	Examine amount of oxygen suitable for aquatic organism

As receiver antenna with gain of 1dB is not suitable for outdoor environment, it was located within roofed balcony at residential area. Following Asia region LoRa frequency band, the receiver operated with 920-923 MHz.

Fig. 3 display distances and location of P1 and P3 were positioned within line of sight with the gateway, Rx while P2 located with obstacle in between.

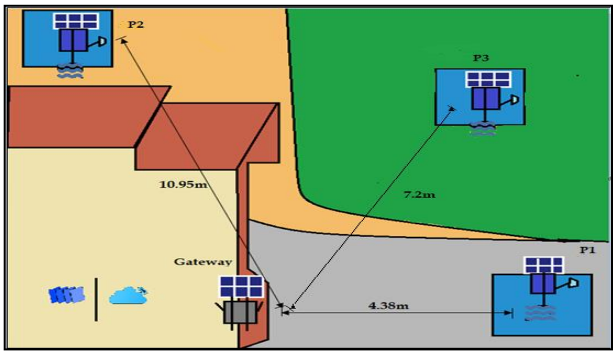


Fig. 3. Water Station Location from Gateway

3.2. *Embedded Transmitter System Parameter*

ESP32-LoRa32 used as WSN with operation frequency of 868M/915 MHz and transmission power up to 20dBm. The device employed Frequency Shift Keying (FSK) modulation mode with data rate from 1.2 Kbps till 300 Kbps. The end nodes transfer data to gateway through LoRaWAN network server, The Thing Stack V3. The data transmission also regulated for every 30 seconds to reduce energy consumption. Below Table 2 and Table 3 show the configurations LoRa properties for all transmitters using two different transmitted power simulated with two different SF.

Table 2. LoRa Parameter Setup 1

Parameter	Tx1	Tx2
Spread Factor	7	9
Bandwidth (kHz)	125	
Frequency Plan	AS920-923 Class A	
Transmitted Power (dBm)	14	
Antenna Gain (dB)	2.5	
Data Transfer Rate (kbps)	5.469	1.758

Table 3. LoRa Parameter Setup 1

Parameter	Tx1	Tx2
Spread Factor	7	9
Bandwidth (kHz)	125	
Frequency Plan	AS920-923 Class A	
Transmitted Power (dBm)	5	
Antenna Gain (dB)	2.5	
Data Transfer Rate (kbps)	5.469	1.758

Deliberated to the total number of end nodes and one gateway, star topology was practiced as preferred in LoRa network technology to prolong the lifetime of the devices (20).

3.3. *Water Environmental Dashboard Design*

The dashboard integrated with LoRaWAN server through LoRaWAN version of MAC V1.0.2 at The Things Stack v3.14.2 and LoRa Cloud. The developed dashboard in apps visualizes data of water quality.

3.4. *Data Analysis*

The properties of water quality data collected through deployed WSN were evaluated with the ideal clean water quality value. The average of data shown in developed dashboard manipulated to determine the state of water and categorized the water class.

LoRa performance on RSSI and SNR observed in The Thing Stack server. Data presented in the server used as the reference to verify packet loss data. Furthermore, the acquired value able to determine received power using equation (1). LoRa performance for each end nodes at different location were recorded and analyzed.

4. Result and Discussion

Movement Control Order (MCO) was executed due to extreme increment of Covid19 epidemic cases, the experiment was deployed at different location from original propose place. The outcome of analysis from collected water quality data were utilized to determine the water quality status.

4.1. *Water Quality Properties Observation and Analysis*

As mentioned in Table I, each of water quality properties cited different information of water status. The result unable to determinate class of water due to absence of Chemical Oxygen Demand (COD), Biochemical Oxygen Demand (BOD) and Ammoniacal Nitrogen that are essential for WQI determination. Those properties can only be gain through manual laboratory test and calibration for higher accuracy. Nonetheless, it able provide real time data to allow environmentalist keep surveillance on water state.

Water properties data collected from water station were processed and analyzed in cloud collected from water station. Fig. 4 below demonstrated the comparison of water quality properties at both water station. The result of water status can be easily view in water quality apps dashboard as visualized in Fig.5.

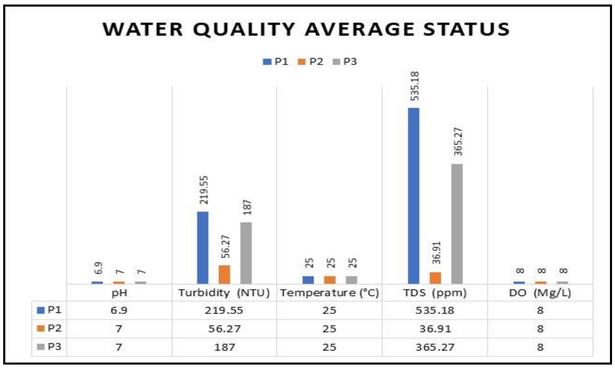


Fig. 4. The Average of Water Quality Parameter Obtained Water Stations

Fig. 5. Visual Board of Water Quality Monitoring Station 1

As the temperature remains at 25°C, a slight conclusion indicated that the water bodies did not contain any hazardous substances. However, several parameters of water properties were affected by microscopic objects influenced by aquatic living and weather.

PH parameter effortlessly affected by external factors but it will remain within acceptable range for unpolluted water. Value of pH analysis result were between 6.9 and 7. The acquired pH value can be classified as Class I and Class II respectively. While turbidity properties inspect excess volume of suspended substances that can disturb water ecosystem productivity. The desire range of turbidity within 0 and 5 NTU. As presented, the turbidity value obtained were 219.55 NTU, 56.27 NTU and 187 NTU respectively. Hence, all water bodies descend to Class III for turbidity features as referred to NWQS.

Total Dissolved Solid (TDS) sensor measure organic and inorganic substances in water. The higher the value of TDS, the higher the salinity. Furthermore, it can determine drinkable water and act as indicator on presence of excessive chemical contaminant. The accepted value of TDS for daily utilization are within range of 1000 ppm and for drinking purpose are between 50 ppm and 500 ppm. Thus, only water station 3 was good for drinking purpose as graphed in Fig. 4. Nonetheless, all water bodies were good for utilization purposes.

The last water properties examined in this research was Dissolved Oxygen (DO). DO is important to indicate the amount of oxygen for aquatic organism survival. Average of DO for all water stations gain 8 mg/L thus the water bodies that classified as Class II allow growth of aquatic organism in the ecosystem. Harmful effect on water organism from oxygen reduction can affect water environment.

In a conclusion, water at Station P2 was purer compared to other two station. However, the nutrient in the water bodies was at small scale so it was not suitable for drinking purposes. Nonetheless, the provided result operated as Support Decision System allow water environmentalist monitor the status of water in timely manner.

4.2. *LoRa Performance Evaluation*

The static test was done within dual environment. The receiver with 1 dB of antenna gain located indoor and three LoRa nodes placed outdoor at fish pond. The test was important to evaluate LoRa performance for autonomous water quality monitoring application. All transmitted nodes, Tx, located at similar height of 0.62 m from the ground but different with height of the receiver, Rx at 4.13 m. The distance of node P1 P2 and P3 from the receiver were 4.3 m, 10.95 m and 7.2 m respectively. Total 10 payload were transmitted to receiver by each of end nodes.

The value of RSSI, SNR and loss packet were recorded for both P1, P2 and P3. The acceptable value of RSSI is above -120 dB while good SNR value for LoRa must be more than -20 dB. Implying the Empirical Okumura-Hata Module, the path loss was determined for each water station.

4.2.1. Result of network setup 1 and 2 at P1

Table 4 below summarize the result of network characteristic on RSSI, SNR, loss packet and received power.

Table 4. LoRa Parameter Performance at P1

Setup	Transmission Power	SF	SNR	RSSI	Loss Packet %	Received Power
1	14	7	11.5	-44	0	-16.5
		9	13.75	-63	0	-16.13
2	5	7	10	-65	0	-25.5
		9	13	-71	0	-25.31

P1 was positioned within Line of Sight (LOS) and nearest with Rx, hence the transmission signal able to operate at its optimal capabilities. It can be verified by value of RSSI obtained during transmission. Range of the RSSI of P1 were within usable RSSI for both setup. Furthermore, one of the RSSI value with transmission power of 14 obtained desirable value of RSSI at -44. The value of SNR also showed its good performance as it operated more than +10 dB. There was also no loss packet recorded for P1. The path loss for P1 decrease as value of transmission power and SF increase. The result of path loss for P1 as presented in Fig. 6.

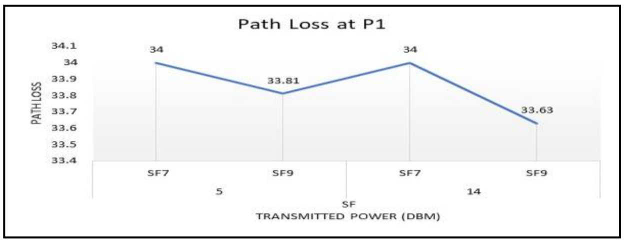


Fig. 6. Path Loss at P1

4.2.2. Result of network setup 1 and 2 at P2

The result of network performance for P2 was summarized as in Table 5.

Table 5. LoRa Performance at P2

Setup	Transmission Power	SF	SNR	RSSI	Loss Packet %	Received Power
1	14	7	9.5	-77	0	-34.91
		9	12.25	-87	0	-33.76
2	5	7	8.25	-89	<1	-45.21
		9	9.75	-95	0	-43.91

P2 was located at NLOS from the Rx and further than P1 location. Based on Table V, value of RSSI were still within acceptable RSSI range though undesirable. The value of RSSI increase as SF is low and transmission power increase. SNR value at P2 show good performance for LoRa network even there were obstacles between end node and gateway. There was notable loss packet for transmission power of 5 and SF7 but the value was still below 1%. Hence, it did not extremely affect packet of data travel. The path loss for P2 is illustrated as Fig.7 below.

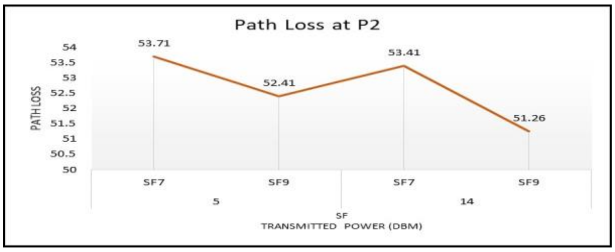


Fig. 7. Path Loss at P2

4.2.3. Result of network setup 1 and 2 at P3

Table 6 portrayed result network characteristic at P3.

Table 6. LoRa Performance at P3

Setup	Transmission Power	SF	SNR	RSSI	Loss Packet %	Received Power
1	14	7	11	-47	0	-19.82
		9	12.25	-65	0	-19.23
2	5	7	9.75	-70	0	-30.78
		9	11.75	-75	0	-29.9

P3 placed within LOS of Rx with present of trees along the transmission path. Though, one of RSSI in setup 1 with SF 7 stay in desirable range while others remain within practical range. SNR obtained for this setup also show good performance with the lowest at 9.75 dB. Also, no loss packet recorded and its path loss result can be viewed in Fig. 8 below.

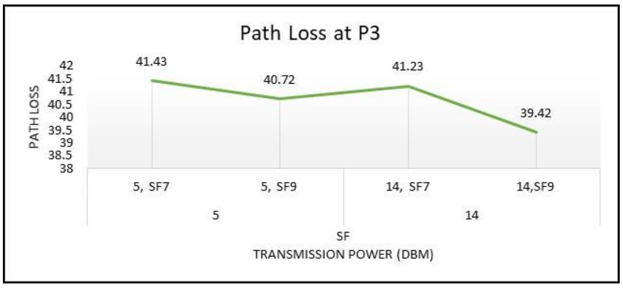


Fig. 8. Path Loss at P3

Analyzing Fig. 6, Fig, 7, and Fig. 8, value of path loss increased with the distance increment of node from gateway and NLOS effect. The pattern of path loss for all location were similar depend on its transmission power and SF. Reduce transmission power save energy consumption but reduce LoRa performance. Furthermore, it is still applicable for low rate data type like water quality data because the range of network properties were still within acceptable range.

5. Conclusion

This research able to reach its objective engaging LoRa in continuous water quality monitoring system. This smart system has potential to reduce the possibility of pollution outbreak as environmentalist able to process quick decision system in real time. Increase SF value can increase the network coverage but it reduced RSSI value and bit rate. Moreover, it increased the energy consumption that can affecting the life of the devices. This research also found that inconsistent weather from sunny to rainy at Malaysia did not influencing LoRa performance. Henceforth, it is suitable to be applied at outdoor environment. The water quality properties value collected from WSN shown in the dashboard verified LoRa ability as wireless network for Smart Environment System at low cost and energy respectively. Though, there were no loss packet recorded in this study, the location of LoRa antenna require LOS and NLOS study effect to reduce undesirable effect on LoRa performance.

In this study, the antenna of gateway used was not suitable with outdoor environment. Nonetheless, LoRa able to connecting gateway and end nodes in semi-outdoor environment. LoRa abilities on SNR, RSSI, packet data loss, the receive power and path loss in this study were evaluated good. Increase the distance of end nodes from each other can reduce redundancy of packet loss data. Moreover, modifying several network parameters of LoRa can improve the system efficiency. As an example, employ suitable antenna for outdoor environment. It is also recommended to operate the system in mesh topology for wide area application to improve the system capabilities.

This paper had exhibited the abilities LoRa within residential environment as experimented in semi-outdoor contemplated with type of antenna used.

Acknowledgements

A special thanks to team of Collaborative IR4.0 Strategies for Lake and River Conservation Management using Innovative Sensing Platforms, the team of sub-program for river conservation using LoRa network and Water Warrior team from UM for idea, knowledge and material support. The financial support with under research grant of IIRG006B-2019 contributed by UM.

References

1. Atwell, C. (2019). Everything You Need to Know about LoRa and the IoT.
2. rget 6.3 – Water quality and wastewater - sd“Tag6monitoring”.
3. K.Kailasam, (2006). Community Water Quality Monitoring Programme in Malaysia, pp1-6.
4. Neal, M. COVID-19 and water resources management: reframing our priorities as a water sector. *Water International*, *45*(5), 435-440, 2020.

5. Berita Harian. 2020. Water Resources Polluted Forced 4 Water Treatment Plant Stop. Berita Harian. 2020. *Sumber Air Mentah Tercemar, 4 LRA Henti Tugas | Berita Harian*. (in Malay).
6. Abdullah, S., Chand, F., Zakaria, S., Loganathan, P., Transforming the Water Sector :National Integrated Water Resources Management Plan (Strategies and Road Map) VOLUME 1 - MAIN REPORT (pp. 1-29). Malaysia. 2016
7. Bartram, J., & Ballance, R., Water quality monitoring:A Practical Guide to the Design and Implementation of Freshwater Quality Studies and Monitoring Programmes (Chapter 2: pp. 2,3). London: Taylor & Francis. 2007.
8. Park, J., Kim, K., & Lee, W., Recent Advances in Information and Communications Technology (ICT) and Sensor Technology for Monitoring Water Quality. *Water*, *12*(2), 510. 2020.
9. Secretariat of the World Meteorological Organization. *Manual on water-quality monitoring* (pp. 4-9). Geneva, Switzerland. 1989
10. Jiang, P., Xia, H., He, Z., & Wang, Z., Design of a Water Environment Monitoring System Based on Wireless Sensor Networks. *Sensors*, *9*(8), 6411-6434. 2009.
11. Bhisekar, C., Meghare, H., Parate, S., Prajapti, S., Meshram, B., & Bobade, P., An IoT based Water Monitoring System for Smart City. *International Research Journal Of Engineering And Technology*, *05*(04), 1352-1354. 2018
12. Joseph Bryan G., I., Meo Vincent C., C., Jen, A., Magno Christian Lemuel, S., Villaruel, K., Villeza, S., & Zaliman, S., Water Quality Monitoring System Using 3G Network. *Journal Of Telecommunication, Electronic And Computer Engineering*, *10*(1-13), 15-18. 2018
13. *BAFWAC-Ericsson, Using smart technology to monitor Stockholm's water systems [Ebook] (pp. 3-5). Stockholm.*. 2017.
14. *RP002-1.0.0 LoRaWAN® Regional 49 Parameters*. Fremont, California: LoRa Alliance, pp. 21-22, 83-86. 2019.
15. Paredes, M., Bertoldo, S., Carosso, L., Lucianaz, C., Marchetta, E., Allegretti, M., & Savi, P., Propagation measurements for a LoRa network in an urban environment. *Journal Of Electromagnetic Waves And Applications*, *33*(15), 2022-2036.
16. E. Leon, C. Alberoni, M. Wister and J. Hernández-Nolasco, "Flood Early Warning System by Twitter Using LoRa", *Proceedings*, vol. 2, no. 19, p. 1213, 2018.
17. Wotherspoon, J.; Wolhuter, R.; Niesler, T. Choosing an integrated radio-frequency module for a wildlifemonitoring wireless sensor network. In Proceedings of the IEEE AFRICON, Cape Town, South Africa, pp. 314–319. 18–20 September 2017.
18. J. Lousado and S. Antunes, "Monitoring and Support for Elderly People Using LoRa Communication Technologies: IoT Concepts and Applications", *Future Internet*, vol. 12, no. 11, p. 206, 2020.
19. Buyukakkaslar, M.T.; Erturk, M.A.; Aydin, M.A.; Vollero, L. LoRaWAN as an e-Health CommunicationTechnology. In Proceedings of the 2017 IEEE 41st Annual Computer Software and Applications Conference(COMPSAC), Turin, Italy, Volume 2; pp. 4–8 July 2017.
20. A. J. Wixted, P. Kinnaird, H. Larijani, A. Tait, A. Ahmadinia, and N. Strachan, "Evaluation of LoRa and LoRaWAN for wireless sensor networks," IEEE SENSORS, pp. 1–3. 2016

Authors Introduction

Ms. Syarifah Nabilah Syed Taha@Tahir

She received her D. Eng. degree in Mechatronic Engineering at University of Malaysia Perlis on 2018. She is currently a Master Degree student of Master of Engineering at University of Malaya in Malaysia.

Dr. Mohamad Sofian Abu Talip

He is a senior lecturer at Department of Electrical Engineering at University of Malaya in Malaysia. He graduated from System Information, Keio University in Japan. He received his Master Degree in Master of Science, Computer Engineering at International Islamic University, Malaysia. He received his D. Eng. degree in Electronic Engineering University of Science Malaysia. His research interest is embedded system.

Mr. Mahazani Mohamad

He is a lecturer at Department of Electrical Engineering at University of Malaya in Malaysia. He received his Master Degree in Computer System Engineering at Queensland University, Australia. He received his D. Eng. degree in Information System Engineering from Kyoto University, Japan. His research interest is computer system.

Dr. Zati Azizul Hakim Hasan

She is a senior lecturer at Department of Artificial Intelligence at University of Malaya in Malaysia. She graduated from Artificial Intelligent and Robotics, University of Auckland in New Zealand. She received both Master Degree and D. Comp. Sc. degree in Computer Science from University of Malaya, Malaysia. Her research interest is artificial intelligence and robotics.

Associate Prof. Zeeda Fatimah Mohamad

She is an Associate Professor at Department of Science and Technology Studies at University of Malaya in Malaysia. She graduated from Science and Technology, Sussex University in UK. She received her Master Degree in Science and Technology Management Policy from Lund University, Sweden. She received her D. Eco. degree in Ecology at University of Malaya. Her research interest is environmental protection.

Prof. Dr. Amir Feisal Merican Aljunid Merican

He is a professor at Institute of Biological Sceience at University of Malaya in Malaysia. He graduated from Computational Biology, Cambridge University, UK. He received his Master Degree in Biological Chemical at Oxford University, UK. He received his degree in Microbiology at University of Malaya, Malaysia. His research interest is environmental biotechnology.

Dr. Mohamadariff Othman

He is a senior lecturer at Department of Electrical Engineering at University of Malaya in Malaysia. He graduated from Engineering in Antenna and Propagation at University of Science, Malaysia. He received his Master Degree in Telecommunications Electronic at University of Science Malaysia. He received his D. Eng. degree in Computer Electronic at Malaysia Multimedia University. His research interest is antenna.

Dr. Tengku Faiz Tengku Mohmed Noor Izam

He is a senior lecturer at Department of Electrical Engineering at University of Malaya in Malaysia. He graduated from Electronic Engineering at Surrey University, UK. He received his Master Degree in Engineering at Sheffield University, UK. He received his D. Eng. degree in Electrical Engineering at Purdue University, US. His research interest is wireless technology.

Dr. Mohd Faiz Mohd Salleh

He is a senior lecturer at Department of Electrical Engineering at University of Malaya in Malaysia. He received all Doctoral, Master Degree and D. Eng. degree in Engineering at Shizuoka University, Japan. His research interest is renewable energy.

Wideband Antenna with UHF Sensor Applicability for HV Equipment in Smart Grid

S. M. Kayser Azam*
Department of Electrical Engineering, Universiti Malaya
Kuala Lumpur 50603, Malaysia†

Mohamadariff Bin Othman
Department of Electrical Engineering, Universiti Malaya
Kuala Lumpur 50603, Malaysia

Tarik Abdul Latef
Department of Electrical Engineering, Universiti Malaya
Kuala Lumpur 50603, Malaysia

H. A. Illias
Department of Electrical Engineering, Universiti Malaya
Kuala Lumpur 50603, Malaysia

Mohd Fadzil Ain
School of Electrical and Electronic Engineering, Universiti Sains Malaysia
Nibong Tebal 14300, Malaysia

Yazeed Qasaymeh
Department of Electrical Engineering, Majmaah University
Al-Majmaah, 11952, Saudi Arabia
E-mail: kaysereee2k8@gmail.com, mohamadariff@um.edu.my, tariqlatef@um.edu.my, h.illias@um.edu.my, eemfadzil@usm.my, y.qasaymeh@mu.edu.sa
www.um.edu.my

Abstract

Wideband antenna is proposed as an ultra-high frequency (UHF) sensor for high voltage equipment in smart grid. First, an antenna is designed in UHF range with wideband characteristics. Next, another antenna is designed from the first one to further improve its sensor applicability. Both antennas offer sufficient levels of realized gain and total efficiency while radiating with omni-directional patterns. Proposed antennas are compact and suitable to be applied as UHF sensors especially for the faulty-insulation detection in high voltage switchgears and power transformers.

Keywords: Wideband antenna, UHF sensor, partial discharge detection, smart grid.

1. Introduction

Antennas, apart from being the most fundamentally significant components in wireless communication systems, have emerging demands in many applications of next generation technologies. Other than communications, in the 5G-based systems, antennas are utilized for energy harvesting applications as per Ref. 1 and Ref. 2. Recently, antennas are having especial attentions from many researchers for their sensor applications on detecting the insulation faults in high voltage equipment e.g., cables, switchgears, power transformers etc[3-4]. Among the insulation faults, partial discharge (PD) is one of the most potentially dangerous phenomena in high voltage equipment. PD can be detected by many techniques that are mostly based on invasive and direct electrical connections. Since PD signals also appear as electromagnetic signals in the ultra-high frequency (UHF) range, antennas are employed as PD sensors that can detect the PD wirelessly which is highly required for the next generation smart grid system. As a result, instead of using antennas only for communication purposes, UHF antennas are being researched to improve their sensor applicability for detecting the PD[5] in high voltage equipment.

UHF antennas, as sensors for detecting the PD, are designed in a way so that they cover the UHF range as wide as possible and provide sufficient gain. In fact, an UHF antenna should provide more than 2 dBi gain so that weak PD signals are significantly detected[4]. For example, a bio-inspired monopole antenna is developed[5] to detect PD through wideband characteristics with sufficiently high gain. However, the antenna is quite large in physical dimensions which is a major limitation since UHF sensors for PD detection should be compact due to their expected portability or remote integration. A wideband antenna is proposed in Ref. 6 and compared with high frequency current transformer method for the PD detection. Nevertheless, since the antenna gain is not considered by Ref. 6, the antenna used a transmission line method i.e., direct electrical connections which does not fulfil the contactless method of PD detection. Recently, a bare patch antenna and its integrated design with filters are proposed[3] for PD detection. Being a microstrip patch type, antennas in Ref. 3 inherently have no wideband characteristics which is highly required for PD detection in the UHF range. Furthermore, Ref. 3 has not revealed antenna gain information in useful details while antenna radiation patterns in fact are worsened after the filter-antenna technique is applied. As a result, implications found from the Ref. 3 are not much convincing to detect PD sources on practical sites. Another recent work is found in Ref. 4 on wideband antenna design for PD detection. Although authors claimed that the antenna is a microstrip type with wideband and high gain characteristics, it is obvious from the highly distorted monopolar radiation patterns in higher frequency regions that the antenna actually works in the lower frequency region of its bandwidth. Consequently, the entire bandwidth may not accurately detect PD signals in higher frequency regions.

To overcome the major challenges described so far, wideband antenna is proposed by this work for detecting partial discharge in high voltage equipment. Two novel antennas have been introduced with wideband characteristics where both antennas have sufficiently high realized gain throughout their corresponding bandwidth. Purposes of designing two antennas are to first integrate with smart grid technology and then to separate communication bands from the UHF range. Next section of this paper is going to discuss design techniques of proposed antennas. Results obtained from the analysis will be discussed on section 3 before leaving a conclusion on section 4.

2. Design Techniques

Both the 1st antenna and 2nd antenna are designed on the commercial Rogers RT5880 substrate. Computer Simulation Technology (CST) software with 2021 version is utilized for design and simulation purposes. Partial ground plane (PGP) is method adopted for both antennas by applying the microstrip feeding technique[7],

$$w_f = \frac{7.48 \times h}{e^{\left(Z_0 \frac{\sqrt{\varepsilon_r + 1.41}}{87}\right)}} - 1.25 \times t. \tag{1}$$

Here, w_f is the feedline width, h is the substrate height (dielectric), t is the conductor thickness, Z_0 is the characteristic impedance, and ε_r is the relative permittivity of the dielectric substrate.

2.1. *1st Antenna*

Design of the 1st antenna is inspired from the four-leaf clover shape which is a unique variation in the nature.
The step-by-step design technique of the 1st antenna is illustrated in Fig. 1 where four conductive patch elements of circular shapes are utilized to mimic the four-leaf clover shape. Each circle has radius of 15 mm while connecting other three circles at its center to complete the radiating patch design. The substrate has a width (W) of 70 mm and a length (L) of 100 mm. The feedline of a length (l_f) of 50 mm connects the radiating patch from the bottom edge on the top plane. In the PGP, width is same as the substrate width, but the length (l_g) is 49.9 mm from the bottom edge of the ground plane.

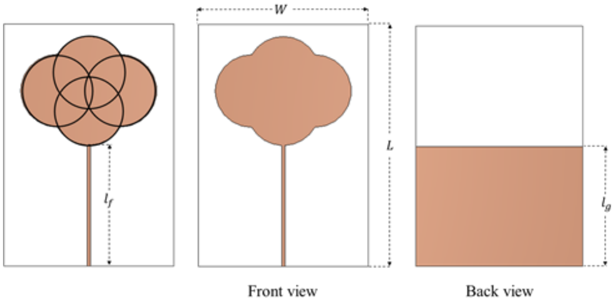


Fig. 1. Design technique (1st antenna).

Based on the design of the 1st antenna, a simulation is performed to observe its S-parameter (return loss) as shown in Fig. 2. Bandwidth ($S_{11} < -10$ dB) of the antenna is found wide due to its unconventional shaped radiating patch over the PGP. A sharp resonance is also found at 1.41 GHz with a return loss of 32.55 dB which is caused by the closest circle near the feedline.

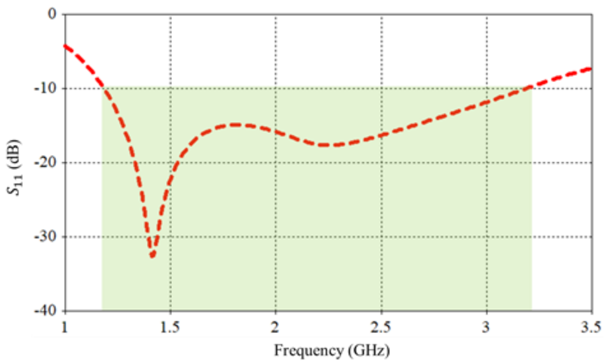


Fig. 2. S-parameter (1st antenna).

Since the S-parameter observation of the 1st antenna reveals that the antenna has a wide bandwidth by mostly covering the UHF range, it obtains the primary requirement of being used as an UHF sensor for PD detection. In Fig. 3, surface current distribution on the 1st antenna is observed where it is found that the closest circle to the feedline accumulates more current at lower frequencies of the UHF range.

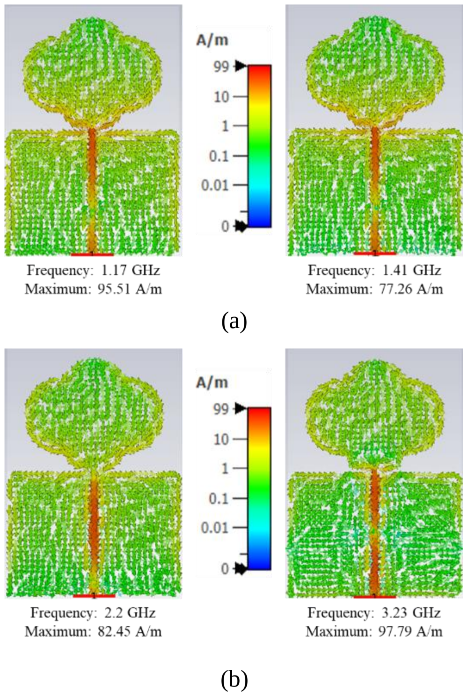


Fig. 3. Surface current distribution (1st antenna).

2.2. 2nd Antenna

The 2nd antenna is designed with the same dimensions as the first one except the PGP is truncated for separating the higher frequencies from the original antenna bandwidth. This is to separate the available communication bands i.e., GSM, 3G, Bluetooth, WLAN etc. that can highly affect the incoming PD signals. Fig. 4 shows the design technique of the 2nd antenna.

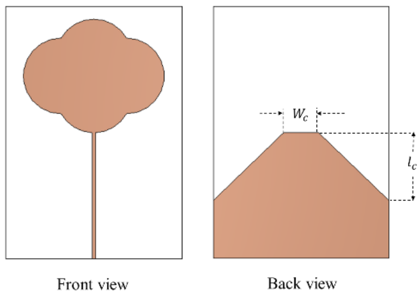


Fig. 4. Design technique (2nd antenna).

From the upper edge on the PGP, at a distance (l_c) of 27 mm, the ground plane is truncated as shown in Fig. 4 by leaving a width (W_c) of 14 mm on the upper edge. This truncated PGP technique greatly affects the bandwidth of the original antenna. As seen from Fig. 5, gradual slopes on both side of the truncated PGP gradually separates the higher frequencies the original bandwidth by leaving a decently wide bandwidth near between 1 GHz to 1.5 GHz. Because there are no commonly available communication bands within this frequency range, this bandwidth adds an extra applicability to the 2nd antenna of being utilized as a UHF sensor for PD detection. In fact, most of the PDs as electromagnetic signals appear within this range as well.

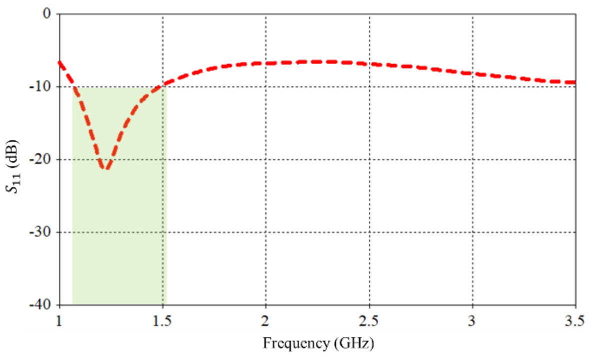


Fig. 5. S-parameter (2nd antenna).

In Fig. 6, surface current distribution of the 2nd antenna is presented, and it is seen that the truncated PGP prevents the current to be accumulated on upper parts of the patch.

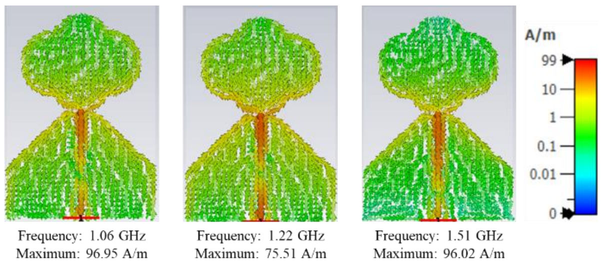


Fig. 6. Surface current distribution (2nd antenna).

3. Results and Discussion

3.1. 1st Antenna

Based on the electromagnetic simulation on CST software with Hexahedral Transmission Line Method for the time domain solver, results are obtained for both antennas. The simulated results are presented for the VSWR (voltage standing wave ratio), realized gain, radiation pattern, and efficiency of the antenna. In Fig. 7, VSWR is presented for the 1st antenna. It is found that the commercial bandwidth (VSWR<2) of antenna is 2060 MHz from 1.17 GHz to 3.23 GHz frequencies. This largely covers the UHF range for sensing the PD in HV equipment.

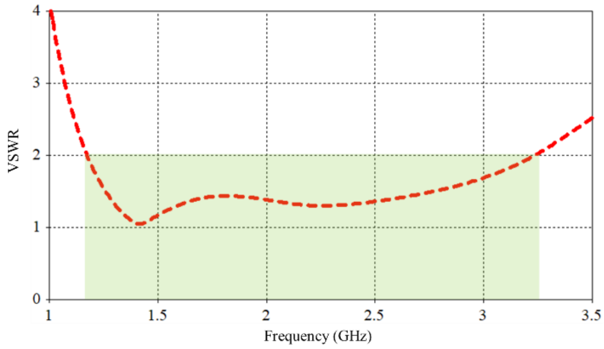
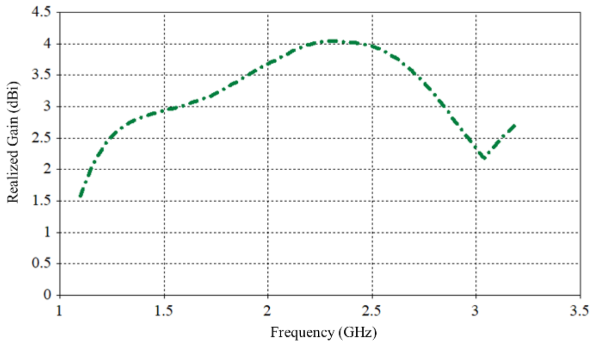
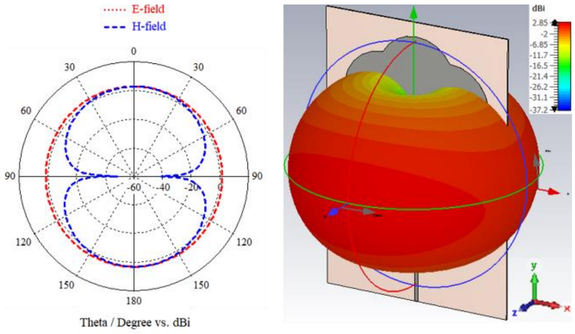


Fig. 7. VSWR (1st antenna).

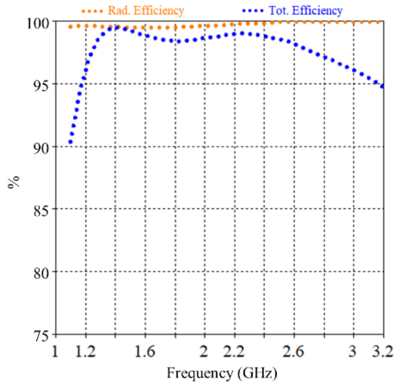
Realized gain is the most significant parameter to be observed prior to detecting PD with UHF antenna since it determines how much the amplitude of the received will be. Figure 8 illustrates the realized gain over frequency, radiation pattern, radiation efficiency, and total efficiency of the 1st antenna. For the 1st antenna, the average value of the realized gain is about 3.21 dBi throughout the entire bandwidth of 2060 MHz from 1.17 GHz to 3.23 GHz. It is clear that the average value of the realized gain for the 1st antenna is much above the expected 2 dBi gain to detect PD signals. The maximum realized gain obtained by the 1st antenna is 4.04 dBi at 2.31 GHz while minimum realized gain is obtained as 1.65 dBi at 1.06 GHz.



(a)



(b)



(c)

Fig. 8. 1st Antenna radiation analysis (a) realized gain over frequency (b) 2D and 3D radiation patterns (c) efficiency.

The antenna radiates with a typical monopolar omnidirectional pattern as shown in Fig. 8(b) where both the 2D and 3D plotted patterns are illustrated. Such a pattern is highly required for detecting PD in HV equipment since it can receive PD signal from all directions. Figure 8(c) shows the radiation efficiency and the total efficiency of the 1st antenna. It is found that the antenna, on average, has more than 99.5% radiation efficiency. The total efficiency finds the maximum value as 98.8% and the lowest value as 90%. This indicates that the antenna can receive PD signals with high accuracy.

3.2. *2nd Antenna*

The 2nd antenna is simulated and observed for its corresponding parameter in the same manner as before. The VSWR is presented in Fig. 9, and it is found that the 2nd antenna has a commercial bandwidth (VSWR<2) of 450 MHz from 1.06 GHz to 1.51 GHz. This is the prominent frequency range that a PD source is detected with its electromagnetic signals. Furthermore, there are no common band for communication systems within this range. Hence, the 2nd antenna achieves more precise applicability of being used as an UHF sensor for PD detection.

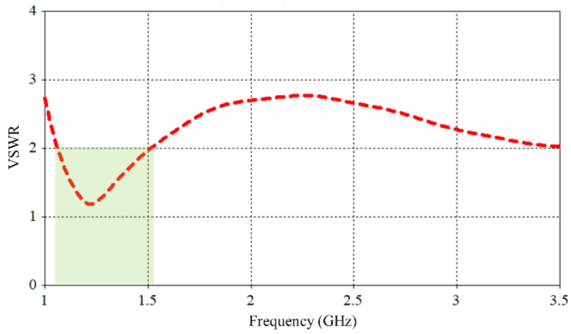
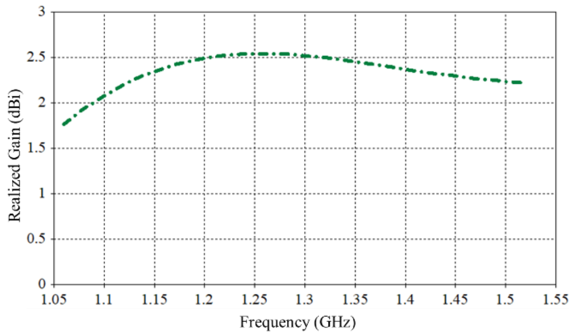
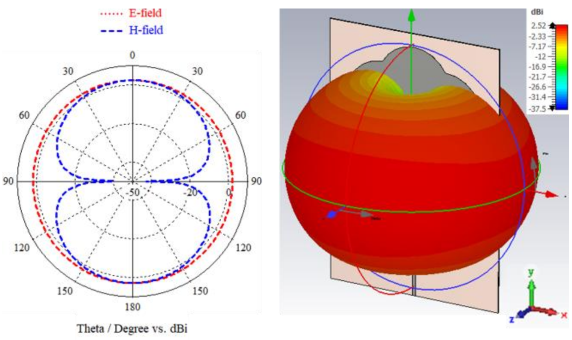


Fig. 9. S-parameter (2nd antenna).

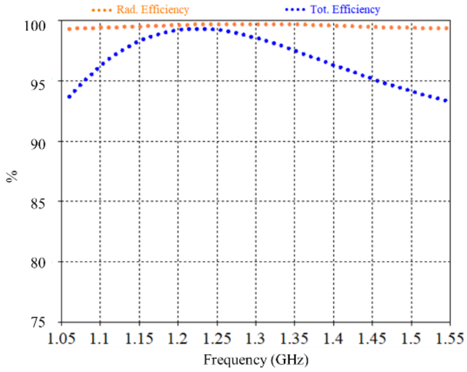
Figure 10 presents the realized gain over frequencies, radiation patterns, radiation efficiency, and total efficiency of the 2nd antenna. It is found that the 2nd antenna has an average realized gain of 2.33 dBi while having a maximum realized gain of 2.54 dBi at 1.26 GHz. This indicates that the 2nd antenna, throughout its entire bandwidth, maintains a consistently higher realized gain than the level of realized gain required for PD detection.



(a)



(b)



(c)

Fig. 10. 2nd antenna radiation analysis (a) realized gain over frequency (b) 2D and 3D radiation patterns (c) efficiency.

Radiation pattern in 2D and 3D plots are shown in Fig. 10 (b), and it is observed that the 2nd antenna radiates with a better monopolar omnidirectional pattern which is required for PD detection from all directions of HV equipment. As seen from Fig. 10(c), the average radiation efficiency for the 2nd antenna is found as 99.4%. The maximum value of the total efficiency is 98.6% and the minimum value of the total efficiency is 87.5% which enables the 2nd antenna to detect PD signals with higher accuracy as well. Table 1 presents the summarized results of both antennas and their comparison with recent works. It is found that both antennas are compact and compatible to be utilized as UHF sensors for partial discharge detection in high voltage equipment of smart grid.

Table 1. Proposed antenna results compared to recent works

Antenna	f_0 (GHz)	FBW (%)	$W \times L$ (mm^2)	Ga_{avg} (dBi)	ηR_{avg} (%)
in Ref. 3	1.55	12.85	–	–	–
in Ref. 4	2.09	68.00	80×100	–	–
1st	2.22	92.80	70×100	3.21	99.5
2nd	1.28	35.02	70×100	2.33	99.3

4. Conclusion

In this paper, two wideband antennas have been proposed for the applications of partial discharge detection in high voltage equipment. Partial ground technique is applied for designing both antennas whereas a truncated ground technique is applied only for the second antenna to separate the communication bands from the antenna bandwidth in terms of partial discharge detection requirements through UHF sensing method. From the analysis of results, it is obvious that both antennas have applicability of being utilized as UHF sensors in detecting partial discharge because of their consistently high realized gain over the wideband.

5. References

1. S. M. K. Azam, M. S. Islam, A. K. M. Z. Hossain, and M. Othman, “Monopole Antenna on Transparent Substrate and Rectifier for Energy Harvesting Applications in 5G,” International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 11, No. 8, pp. 84-89, September 2020.
2. A. K. M. Z. Hossain, N. B. Hassim, S. M. K. Azam, M. S. Islam, and M. K. Hasan, “A Planar Antenna on Flexible Substrate for Future 5G Energy Harvesting in Malaysia,” International Journal of Advanced Computer Science and Applications (IJACSA), vol. 11, no. 10, pp. 151-155, November 2020.
3. Z. H. Bohari, M. Isa, A. Z. Abdullah, P. J. Soh, M. F. Sulaima, and S. N. M. Arshad, “Performance Assessment for Bare Patch Antenna and Integrated Filter-antenna Microstrip Patch Resonator Based for Partial Discharge Application in HV Equipment,” 8th International Conference on Computer and Communication Engineering (ICCCE), pp. 294-298, 2021.
4. J. P. Uwiringiyimana, Suwarno, and U. Khayam, “Design of an Ultra-Wide Band Microstrip Patch Antenna for Partial Discharge Detection on Power Transformer,” IEEE International Conference on the Properties and Applications of Dielectric Materials (ICPADM), pp. 242-245, 2021.
5. J. Cruz, A. Serres, A. de Oliveira, G. Xavier, C. de Albuquerque, E. da Costa, and R. Freire, “Bio-inspired Printed Monopole Antenna Applied to Partial Discharge Detection,” Sensors, Vol. 19, No. 3, pp. 628, February 2019.
6. Z. Cui, S. Park, H. Choo, and K.-Y. Jung, “Wideband UHF Antenna for Partial Discharge Detection,” Applied Sciences, Vol. 10, No. 5, pp. 1698, March 2020.

Acknowledgements

This work was supported by a university grant with the grant number of GPF032A-2019 from the Universiti Malaya, Malaysia.

Authors Introduction

S. M. Kayser Azam

He received his B.Sc. degree in electrical and electronic engineering from the Khulna University of Engineering and Technology, Bangladesh, in 2013. He obtained his M.Sc. research degree in electronics engineering from the International Islamic University Malaysia (IIUM), Malaysia, in 2019. He is currently a Ph.D. student and a Graduate Research Assistant in the Faculty of Engineering at the University of Malaya, Malaysia.

Dr. Mohamadariff Bin Othman

He received his B.Eng. (Hons.) degree in electronic engineering from Multimedia University, Malaysia, in 2006. He obtained his M.Sc. degree in RF and microwave, and the Ph.D. degree in antenna and propagation from Universiti Sains Malaysia (USM), Malaysia, in 2008 and 2015, respectively. He is now working at the Department of Electrical Engineering, Universiti Malaya, Malaysia, as a Senior Lecturer, since 2016. Prior to that, he served as a lecturer at the UCSI University, Malaysia for almost 1 and half year. His research interests include 5G antennas, dielectric characterization, dielectric resonator antenna design, and antenna design optimization. He is also a member in the Board of Engineers Malaysia (BEM).

Dr. Tarik Abdul Latef

He received his B.Sc. degree in electrical and electronic engineering from the University of Oita, Japan, in 1997, the M.Sc. degree in electrical engineering from the University of Leeds, and the Ph.D. degree from The University of Sheffield, U.K., in 2005 and 2011, respectively. Since 2011, he has been a Senior Lecturer with the Department of Electrical Engineering, Faculty of Engineering, University of Malaya, Malaysia. His research interests include circularly polarized antennas, conformal antennas, frequency independent antennas, array design, dielectric resonator antennas, superconducting metamaterial antennas, and RCS reduction mechanisms.

Assoc. Prof. Dr. H. A. Illias

He received the bachelor's degree in electrical engineering from the University of Malaya, Kuala Lumpur, Malaysia, in 2006, and the Ph.D. degree in electrical engineering from the University of Southampton, Southampton, U.K., in 2011.,He joined Freescale Semiconductors (M) Sdn. Bhd., as a Product Engineer in 2008. Since 2011, he has been an Academic Staff with the Department of Electrical Engineering, University of Malaya. He was the Secretary of the IEEE Dielectrics and Electrical Insulation Society (DEIS) Malaysia Chapter from 2015 to 2019. Since 2018, he has been the Counselor of the IEEE University of Malaya Student Branch. He is currently a member of Institution of Engineering and Technology, Corporate Member of Institution of Engineers Malaysia, a registered Chartered Engineer (C.Eng U.K.), and a Professional Engineer (P.Eng Mal).

Prof. Dr. Mohd Fadzil Ain

He received his B.S. degree in electronic engineering from Universiti Teknologi Malaysia, Johor Bahru, Malaysia, in 1997, the M.S. degree in radio frequency and microwave from Universiti Sains Malaysia (USM), Penang, Malaysia, in 1999, and the Ph.D. degree in radio frequency and microwave from the University of Birmingham, Birmingham, U.K., in 2003. Since 2003, he has been with the School of Electrical and Electronic Engineering, USM, Malaysia. He is also the Dean of the School of Electrical and Electronic Engineering, USM. His current research interests include wireless circuit design, low-temperature-cofired-ceramics-based antenna in package, rain propagation, microwave links, and dielectric resonator antennas.

Dr. Yazeed Qasaymeh

He is affiliated to Department of Electrical Engineering, University of Majmaah. He is currently providing services as Assiatant Profesor. He has published numerous publications in various national and international peer-reviewed journals and conference papers across the world. Because of the active association with different societies and academies as well as the contributions, he is been recognized by the subject experts around the world. His research interests include microstrip and dielectric resonator antennas, wireless communications hardware etc.

New Particle Swarm Optimization Variant with Modified Neighborhood Structure

Koon Meng Ang[1], Mohd Rizon Mohamad Juhari[1], Wy-Liang Cheng[1], Wei Hong Lim[1,*], Sew Sun Tiang[1], Chin Hong Wong[2], Hameedur Rahman[2], Li Pan[1]

[1]*Faculty of Engineering, Technology and Built Environment, UCSI University, Jalan Menara Gading, UCSI Heights, 56000 Cheras, Kuala Lumpur, Malaysia*

[2]*Maynooth International Engineering College, Fuzhou University, 350108 Fuzhou, Fujian, China*

[3]*Faculty of Computing and Artificial Intelligence, Air University, Main Campus PAF Complex E-9 Islamabad, Islamabad Capital Territory 44000, Pakistan*

E-mail: 1001436889@ucsiuniversity.edu.my, mohdrizon@ucisuniversity.edu.my, 1001437466@ucsiuniversity.edu.my, limwh@ucsiuniversity.edu.my, tiangss@ucsiuniversity.edu.my, chinhong.wong@mu.ie, rhameedur@mail.au.edu.pk, 1002060534@ucsiuniversity.edu.my

Abstract

Numerous particle swarm optimization (PSO) variants were proposed in past decades to tackle different types optimization problems more robustly. Nevertheless, the imbalance of explorative and exploitative search behaviors remains as an on-going research challenge that can restrict the performance of PSO. In this paper, a new variant known as PSO with time-varying topology connectivity (PSO-TVTC) is proposed. A time-varying topology connectivity (TVTC) module is designed to achieve the proper regulation on explorative and exploitative behaviors of PSO via dynamic modifications of particle's topology connectivity throughout the optimization process. Experimental results reveal that the proposed PSO-TVTC has exhibited prominent performance among its competitors by producing 7 best mean fitness out of 8 benchmark functions.

Keywords: global optimization, modified neighborhood structure, metaheuristic, particle swarm optimization

1. Introduction

Various real-world engineering problems such as energy management[1], material machining[2], pattern recognition[3] etc. can be formulated as complex optimization problems that are not feasible to solve with conventional methods. Recently, metaheuristic search algorithms (MSAs) inspired by different natural phenomena have emerged as the popular approaches to solve these complex problems.

Motivated by bird flocking behaviors to locate food sources, particle swarm optimization (PSO) is a popular MSAs used to solve various optimization problems due to its desirable characteristics such as the simplicity in implementation and fast convergence speed. Similar with most MSAs, original PSO tends to suffer with premature convergence when solving complex problems.[4] Numerous methods such as parameter adaptation[5,6], modified neighborhood structure[4,7,8], modified learning strategies[9,10,11,12] etc. were introduced to enhance the performance of PSO. Despite of improvement achieved, some new PSO variants might suffer with undesirable drawbacks such as high complexity and manual tuning of new parameters. Appropriate strategies used to balance the exploration and exploitation searches of PSO for solving different types of optimization problems effectively remain as an on-going research challenge.

In this paper, a PSO with time-varying topology connectively (PSO-TVTC) is designed by leveraging the capability of a novel TVTC module to adjust the exploration and exploitation strengths of particles via dynamic modification of their neighborhood structures. This is because each particle behaves more explorative when connected with less neighbors and vice versa.[4] The Shuffling and perturbation mechanisms are also designed for PSO-TVTC to handle premature convergence issue.

2. Basic PSO

Each PSO particle serves as a potential solution for an optimization problem with *D*-dimensional size. The current state of each i-th particle is associated with the velocity and position vectors of $V_i = [V_{i,1}, \ldots, V_{i,d}, \ldots, V_{i,D}]$ and $X_i = [X_{i,1}, \ldots, X_{i,d}, \ldots, X_{i,D}]$, respectively, where $d = 1,\ldots, D$ refers to a dimension index. Each i-th particle is also able to memorize the best solutions ever achieved by itself and population in the personal best position of $P_i = [P_{i,1}, \ldots, P_{i,d}, \ldots, P_{i,D}]$ and global best position of $P_g = [P_{g,1}, \ldots, P_{g,d}, \ldots, P_{g,D}]$, respectively. In the $(t+1)$-th iteration of search process, the velocity and position of each i-th particle in d-th dimension can be updated as:

$$V_{i,d}(t+1) = \omega V_{i,d}(t) + c_1 r_1 (P_{i,d}(t) - X_{i,d}(t)) + c_2 r_2 (P_{g,d}(t) - X_{i,d}(t)) \quad (1)$$

$$X_{i,d}(t+1) = X_{i,d}(t) + V_{i,d}(t+1) \quad (2)$$

where $i = 1,\ldots,S$ and S is the population size; ω is inertia weight used to control the influence of previous velocity; r_1 and r_2 are two random numbers generated between 0 to 1; c_1 and c_2 are acceleration coefficients.

3. Proposed PSO-TVTC

3.1. *TVTC Module*

Previous studies reported that PSO with smaller topology connectivity can perform well in multimodal problems with numerous local optima due to its more explorative nature, whereas PSO with large connectivity has stronger exploitation strength and able to exhibit fast convergence speed in solving unimodal problems[4]. Therefore, a TVTC module is proposed in this paper to regulate the exploration and exploitation strengths of each particle via the dynamic adjustment of its topology connectivity with time. At the beginning stage of optimization process, the topology connectivity of each i-th particle is set as $TC_i = 1$ to connect with a randomly selected particle. The connection between particles is not bidirectional as shown in Fig. 1. For instance, if the i-th particle has chosen to connect with j-th particle, the latter particle might select the other k-th particle as its neighbor instead.

Suppose that $TC_{min} = 1$ and $TC_{max} = S - 1$ refers to the minimum and maximum topology connectivity assigned to each particle, respectively, whereas FE_{max} is the maximum fitness evaluation numbers. The TVTC module is designed to increase the topology connectivity TC_i of each i-th particle linearly with the current fitness evaluation numbers (FEs) of k as follow:

$$TC_i = \lfloor TC_{\min} + TC_{\max}[(k-1)/(FE_{\max}-1)] \rfloor \quad (3)$$

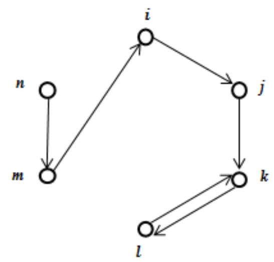


Fig. 1. Initial topology connectivity of PSO-TVTC particles.

This topology modification strategy of TVTC module cam gradually increase exploitation strengths of particles for fine tuning the promising solution regions found.

A shuffling mechanism is further incorporated into TVTC module to address the premature convergence issue. When the i-th particle is unable to improve the global best fitness $f(P_g)$ for M successive FEs, it is trapped into local optima. The TVTC module is executed to discard all existing neighbors of i-th particle and then randomly reassign another TC_i new neighbors as shown in Fig. 2. The new topology information is expected to offer new directional information for stagnated particle and guide it to escape from local optima. A random perturbation process is also performed on the P_g to provide additional momentum for it to jump out of local optima. For the randomly selected d-th dimension of P_g, i.e., $P_{g,d}$, the perturbed component $P_{g,d}^{per}$ is produced as:

$$P_{g,d}^{per} = r_3 P_{g,d} + (1 - r_3)(P_{x,d} - P_{y,d}) \quad (4)$$

where r_3 is a randomly generated number between 0 to 1; $P_{x,d}$ and $P_{y,d}$ are personal best positions of two randomly selected particles. Current P_g is replaced by its perturbed counterpart P_g^{per} if the latter particle is more superior.

The overall mechanisms of TVTC module are shown in Fig. 3. For each i-th particle, a flag variable $fc(i)$ is used to record the number of successive FEs for i-th particle fails to update $f(P_g)$. When $fc(i) > M$, both shuffling and perturbation mechanisms are triggered. The parameter

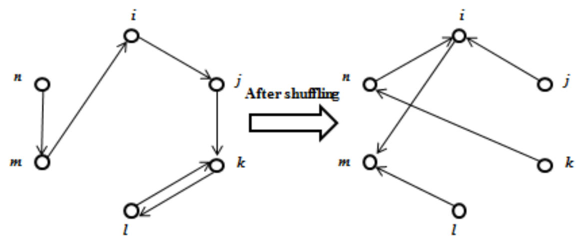


Fig. 2. Shuffling mechanisms triggered by TVTC module if the i-th particle fails to update the $f(P_g)$ for M successive FEs.

```
TVTC_module(fc(i),mode,P_g, k, TC_i)
TC_old = TC_i
Calculate current TC_i using Eq. (3);
if current TC ≠ TC_old then
    Randomly select TC_i - TC_old particles as neighbor;
else
    if fc(i) > M then
        Randomly select TC particles as neighbors;
        Perform perturbation on P_g using Eq. (4);
        Update P_g if P_g^per has better fitness;
        k = k + 1;
        fc(i) = 0;
    end if
end if
```

Fig.3. Implementation of TVTC module.

value of M = 5 is set to achieve better tradeoff in term of solution accuracy and convergence speed.

3.2. *Complete PSO-TVTC*

The complete framework of PSO-TVTC is shown in Fig. 4. Unlike basic PSO that relies on both P_i and P_g, the proposed PSO-TVTC leverages the promising direction information of P_i and $P_{n,i}$ (i.e., best neighbor of i-th particle) in updating the velocity of each i-th particle with better diversity for $d = 1,\ldots,D$ as:

$$V_{i,d}(t+1) = \omega V_{i,d}(t) + c_1 r_1 (P_{i,d}(t) - X_{i,d}(t)) + c_2 r_2 (P_{n,d}(t) - X_{i,d}(t)) \tag{5}$$

The search process of PSO-TVTC using Eqs. (5) and (2) along with TVTC module is iterated until the termination condition of $k > FE_{max}$ is satisfied.

4. Performance Evaluation

4.1. *Simulation Settings*

Eight scalable benchmark functions with dimensional sizes of D = 50 are used for performance evaluation. These include F1 Sphere with range of $[-100, 100]^D$, F2 Step with range of $[-100, 100]^D$, F3 Rosenbrock with range of $[-2.048, 2.048]^D$, F4 Rastrigin with range of $[-5.12, 5.12]^D$, F5 Noncontinuous Rastrigin with range of $[-5.12, 5.12]^D$, F6 Grienwank with range of $[-600, 600]^D$, F7 Ackley with range of $[-32, 32]^D$ and F8 Salomon with range of $[-100, 100]^D$. Five well-established PSO variants known as BPSO[6], FlexiPSO[10], FIPSO[9], FLPSO-QIW[5] and FPSO[8] are used to compare with PSO-TVTC. All algorithms are simulated with S = 30 and FE_{max} = 30,0000 for fair comparisons.

```
PSO-TVTC
Initialize population of PSO-TVTC
TC = 1, k = 0 ;
for each particle i do
    randomly select 1 particle as neighbor;
    fc(i) = 0;
end for
while k < FE_max do
    for each particle i do
        Perform TVTC_module (Refer Fig. 3);
        Identified the best neighbor P_n(i);
        Update V_i with Eq. (5) and X_i with Eq. (2);
        Evaluate the fitness of updated X_i;
        k = k + 1;
        Update the P_i and P_g;
        if P_g is improved then
            fc(i) = 0;
        else
            fc(i) = fc(i) + 1;
        end if
    end for
end while
```

Fig.4. Complete framework of PSO-TVTC.

4.2. *Comparisons of PSO Variants*

The search accuracies of all PSO variants are evaluated in terms of the mean fitness value (F_{mean}) and standard deviation (SD). Smaller F_{mean} and SD values are more desirable because it indicates the errors between P_g and actual global optimum are consistently smaller. The F_{mean} and SD values produced by the compared PSO variants when solving all benchmark functions are shown in Table 1 and the best results are represented in bold fonts. From Table 1, PSO-TVTC can outperforms its peers with large margin when solving majority of tested problems, implying the excellent search accuracy of the proposed algorithm. PSO-TVTC produces the best (i.e., smallest) F_{mean} values for all benchmark functions except for F3. PSO-TVTC is also the only PSO variant that is able to locate the global or near global optima of F1, F2, and F4 to F8. Majority of compared PSO variants exhibit relatively poor performance in F3 because the global optimum of Rosenbrock function locates in a valley with long narrow parabolic shape. While most algorithms can find the valley regions, it is extremely challenging for them to converge towards the global optimum, hence leading to poor F_{mean} values.

5. Conclusions

A new variant of PSO-TVTC is proposed in this paper, where a TVTC module is designed to attain better

Table 1. Simulation results of F_{mean} and SD for solving 50-D benchmark functions

f	Metric	FLPSO-QIW	FlexiPSO	FPSO	FIPSO	BPSO	PSO-TVTC
F1	F_{min}	2.90E-81	1.78E-04	7.02E+01	2.96E-01	4.67E+03	**0.00E+00**
	SD	5.97E-81	5.23E-05	6.98E+01	8.06E-01	7.30E+03	0.00E+00
F2	F_{min}	3.33E-02	8.50E+03	8.48E+01	3.03E+00	3.33E+03	**0.00E+00**
	SD	1.83E-01	3.09E+03	6.97E+01	1.14E+01	5.47E+03	0.00E+00
F3	F_{min}	**4.22E+01**	4.48E+01	5.68E+01	4.77E+01	2.10E+02	4.74E+01
	SD	2.39E-01	1.04E+00	7.08E+00	8.44E-01	4.34E+02	5.79E-01
F4	F_{min}	2.60E+00	2.12E-04	1.85E+01	1.57E+00	1.15E+02	**1.68E-10**
	SD	1.52E+00	6.24E-05	1.02E+01	3.71E+00	7.78E+01	2.72E-10
F5	F_{min}	5.58E+00	2.07E-04	1.60E+01	5.70E-01	1.14E+02	**1.68E-10**
	SD	2.36E+00	7.51E-05	9.56E+00	8.65E-01	5.81E+01	2.01E-10
F6	F_{min}	5.75E-04	8.34E-03	1.86E+00	1.93E-01	3.92E+01	**0.00E+00**
	SD	2.21E-03	9.48E-03	9.28E-01	3.47E-01	7.00E+01	0.00E+00
F7	F_{min}	3.43E-14	3.55E-03	1.80E+00	1.70E-01	1.21E+01	**0.00E+00**
	SD	1.07E-14	5.36E-04	1.10E+00	3.38E-01	5.99E+00	0.00E+00
F8	F_{min}	2.33E-01	1.22E+00	1.32E+00	3.33E-02	5.43E+00	**0.00E+00**
	SD	4.79E-02	1.51E-01	8.01E-01	1.30E-01	5.65E+00	0.00E+00

balancing of exploration and exploitation searches via the dynamic adjustment of particles' topology connectivity. Shuffling and perturbation mechanisms are also used by TVTC module to reduce the likelihood of premature convergence. Simulation results show that PSO-TVTC can outperforms its peers in term of search accuracy.

Acknowledgements

This research is supported by Ministry of Higher Education Malaysia (MOHE) under the Fundamental Research Grant Scheme with the project codes of Proj-FRGS/1/2019/TK04/UCSI/02/1 and Proj-FRGS/1/2020/ TK0/UCSI/02/4.

References

1. L. Yao and W. H. Lim, "Optimal purchase strategy for demand bidding," IEEE Transactions on Power Systems, vol. 33, no. 3, pp. 2754-2762, 2017.
2. S. Suresh, N. Elango, K. Venkatesan, W. H. Lim, K. Palanikumar, and S. Rajesh, "Sustainable friction stir spot welding of 6061-T6 aluminium alloy using improved non-dominated sorting teaching learning algorithm," *Journal of Materials Research and Technology*, vol. 9, no. 5, pp. 11650-11674, 2020.
3. M. Alrifaey, W. H. Lim, and C. K. Ang, "A Novel Deep Learning Framework Based RNN-SAE for Fault Detection of Electrical Gas Generator," *IEEE Access*, vol. 9, pp. 21433-21442, 2021.
4. W. H. Lim *et al.*, "A self-adaptive topologically connected-based particle swarm optimization," IEEE *Access*, vol. 6, pp. 65347-65366, 2018.
5. Y. Tang, Z. Wang, and J.-a. Fang, "Feedback learning particle swarm optimization," Applied Soft Computing, vol. 11, no. 8, pp. 4713-4725, 2011.
6. Y. Shi and R. Eberhart, "A modified particle swarm optimizer," in 1998 IEEE international conference on evolutionary computation proceedings. IEEE world congress on computational intelligence (Cat. No. 98TH8360), 1998, pp. 69-73: IEEE.
7. A. A. Karim, N. A. M. Isa, and W. H. Lim, "Hovering Swarm Particle Swarm Optimization," IEEE Access, 2021.
8. M. A. M. De Oca, T. Stutzle, M. Birattari, and M. Dorigo, "Frankenstein's PSO: a composite particle swarm optimization algorithm," IEEE Transactions on Evolutionary Computation, vol. 13, no. 5, pp. 1120-1132, 2009.
9. R. Mendes, J. Kennedy, and J. Neves, "The fully informed particle swarm: simpler, maybe better," *IEEE transactions on evolutionary computation*, vol. 8, no. 3, pp. 204-210, 2004.
10. M. Kathrada, "The flexi-PSO: Towards a more flexible particle swarm optimizer," Opsearch, vol. 46, no. 1, pp. 52-68, 2009.
11. A. A. Karim, N. A. M. Isa, and W. H. Lim, "Modified particle swarm optimization with effective guides," IEEE *Access*, vol. 8, pp. 188699-188725, 2020.
12. K. M. Ang *et al.*, "Modified Particle Swarm Optimization with Unique Self-cognitive Learning for Global Optimization Problems," Singapore, 2022, pp. 263-274: Springer Singapore.

Authors Introduction

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.

Mr. Wy-Liang Cheng

He received the B.Eng. degree in Mechatroic Engineering with Hnours from UCSI University, Malaysia in 2020. He is currently purhsing Master of Philosophy in Engineering in UCSI University, Malaysia. His research interests are swarm intelligence and feature selection.

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. 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. 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. Hameedur Rahman

He is an Assistant 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.

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.

N-Switch and All-Path Test Coverage Criterion for Extended Finite State Machine

Tomohiko Takagi
Department of Engineering and Design, Faculty of Engineering and Design, Kagawa University
2217-20 Hayashi-cho, Takamatsu-shi, Kagawa 761-0396, Japan

Koichiro Sakata
Division of Reliability-based Information Systems Engineering, Graduate School of Engineering, Kagawa University
2217-20 Hayashi-cho, Takamatsu-shi, Kagawa 761-0396, Japan

Kouichi Akiyama
Japan WillTech Solution Co., Ltd.
14F Nihonbashi-Muromachi Mitsui Tower, 2-1, Nihonbashi-Muromachi 3-chome, Chuo-ku, Tokyo 103-0022, Japan
E-mail: takagi@eng.kagawa-u.ac.jp, s20g465@stu.kagawa-u.ac.jp

Abstract

This paper shows a new test coverage criterion for extended finite state machine to comprehensively test the combination of state transitions and accompanying actions in software. Our criterion requires that (i) test cases cover all the successive state transition sequences of specified length, and also (ii) the test cases cover all the paths on control flow graphs of actions that accompany each of the successive state transition sequences. (i) and (ii) are the characteristics of N-switch and all-path test coverage criterion, respectively. Its definition, example and effectiveness are discussed in this paper.

Keywords: model-based testing, finite state machine, test case, test coverage criterion

1. Introduction

One of the well-used software modeling languages in model-based testing[1] is EFSM (Extended Finite State Machine)[2,3]. It enables engineers to define the expected behavior of software from the aspect of not only state transitions but also data processing. The part of state transitions, which corresponds to FSM, is drawn in the form of simple table or directed graph. On the other hand, the part of data processing, which corresponds to the actions on state transitions, is written in natural languages or TBFML (Text-Based Formal Modeling Language). Test cases are usually created from EFSM models so as to satisfy a test coverage criterion called *N*-switch[4], that is, cover all the successive state transition sequences of length N+1 ($N \geq 0$). However, *N*-switch is originally designed for FSM, and actions are not taken into account in it. The behavior of software is determined by the combination of state transitions and actions in EFSM models, and thus it should be comprehensively tested.

To address this problem, we propose a new test coverage criterion for EFSM, named N-SAP$_L$ (*N*-Switch for state transitions and All-Path with loop frequencies L for actions) in this paper. It is a combination of *N*-switch and AP (All-Path test coverage criterion). AP is usually used in structural testing[5] in order to cover all the paths on control flow graphs of programs under test. N-SAP$_L$ requires that (i) test cases cover all the successive state transition sequences of length N+1, and also (ii) the test cases cover all the paths on control flow graphs of actions that accompany each of the successive state transition sequences. (i) and (ii) are the characteristics of *N*-switch and AP, respectively. Note that L expresses a set of loop frequencies to be taken into account in the control flow graphs of actions. When there are any loop structures in the actions, L makes the number of paths that should be covered finite. To avoid the ambiguity and allow the code coverage analysis for AP, all the actions in our EFSM models are written only in TBFML, such as VDM++[6,7].

This paper is organized as follows. Section 2 shows the overview of traditional model-based testing using EFSM. In section 3, we describe the details of N-SAP$_L$ and a simple example. Section 4 gives discussion based on preliminary experiments. Finally, section 5 shows conclusion and our future work.

2. Traditional Model-Based Testing Using EFSM

Fig. 1 shows an example of an abstracted EFSM model. An EFSM model in this study mainly consists of states, events, and actions. It illustrates possible state transitions with data processing in software, and each of them is identified by a tuple in the form of <*from-state*, *event*, *action*, *to-state*>. For example, the state transition β in Fig. 1 is identified by $<s_2, e_1, a_2, s_4>$. Each state transition can optionally have event parameters and a guard. Each of actions consists of codes written in TBFML, and can define and refer variables that characterize the behavior of software. Some actions may be shared by different state transitions. For example, a_2 is shared by β and δ in Fig. 1.

Test cases are created from an EFSM model in the form of sequences of successive state transitions and expected values of variables. The quality of the test cases is usually evaluated by N-switch, and they are created so as to satisfy it. Software implemented based on the EFSM model can be closely tested by a larger value of N. For example, a technique using orthogonal arrays is used to reduce testing effort[8].

In general, coverage of given test cases is calculated by $|M'|/|M|$. M expresses a set of measuring objects that should be covered, and is determined by a selected test coverage criterion. For example, M for 1-switch is a set of all the successive state transition sequences of length 2. In N-switch, a larger value of N results in M that consists of a larger number of measuring objects. On the other hand, M' expresses a set of measuring objects that have actually been covered by the given test cases, and satisfies $M' \subseteq M$. When M' is equal to M, the test coverage criterion is satisfied. Therefore, a larger value of N results in a larger number of test cases to satisfy N-switch.

3. N-SAP$_L$ Test Coverage Criterion for EFSM

As is discussed above, test coverage criteria are used to determine measuring objects. In this section, we propose a procedure to systematically get measuring objects for N-SAP$_L$ from a given EFSM model. The procedure

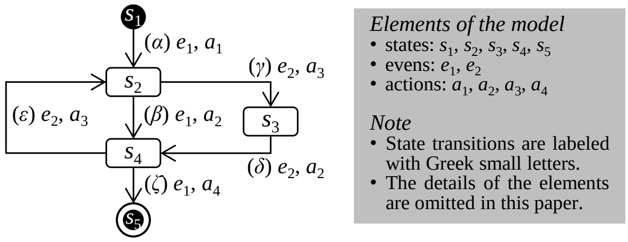


Fig. 1. Example of an abstracted EFSM model.

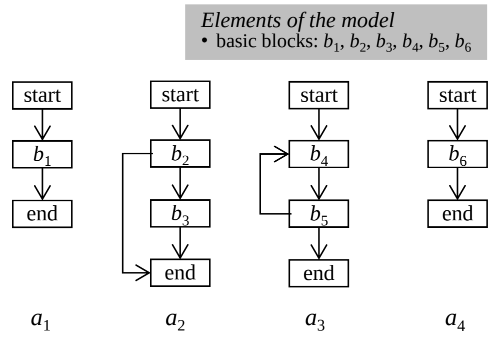


Fig. 2. Abstracted control flow graphs of actions in Fig. 1.

consists of the following five steps. At least Step 2, 3, 4 and a part of Step 5 can be automated.

Step 1. N and L are determined by test engineers according to a given test plan. The former is for the length of successive state transition sequences to be covered, and the latter is the set of loop frequencies to be taken into account in actions.

Step 2. A set of all the measuring objects for N-switch is generated by using a common graph search algorithm. The set is hereinafter referred to as MS_N. The measuring objects as the elements of MS_N are successive state transition sequences of length N+1, and each of them is hereinafter referred to as sts_x ($1 \leq x \leq |MS_N|$). For example, MS_1 for Fig. 1 is {$\alpha\beta$, $\alpha\gamma$, $\beta\varepsilon$, $\beta\zeta$, $\gamma\delta$, $\delta\varepsilon$, $\delta\zeta$, $\varepsilon\beta$, $\varepsilon\gamma$}.

Additionally, MS_N should include special measuring objects that satisfy all of the following conditions:

- Their length are less than N+1.
- They start with the initial state, and end with a final state.

For example, MS_3 for Fig. 1 should include $\alpha\beta\zeta$. The special measuring objects ensure that the satisfaction of N-SAP$_L$ results in the satisfaction of N'-SAP$_L$ (N'<N).

Step 3. A set of all the paths in each action is derived by using common path analysis based on control flow graphs. When there are any loop structures in actions, only the loops of frequencies specified by L are taken into account. Fig. 2 shows abstracted control flow graphs of the actions in Fig. 1. For example, when L is {0, 1}, the sets for a_1, a_2, a_3 and a_4 are $\{b_1\}$, $\{b_2, b_2b_3\}$, $\{b_4b_5, b_4b_5b_4b_5\}$ and $\{b_6\}$, respectively. Note that in a_3 there are infinite paths but L makes the number of paths that should be tested finite.

Step 4. A set of all the paths in the actions that accompany sts_x, which is hereinafter referred to as $PSTS_x$, is derived by the following equation:

$$PSTS_x = \prod_{y=1}^{\#sts_x} PST_{x,y} \tag{1}$$

$\#sts_x$ means the length of sts_x, and $PST_{x,y}$ expresses a set of all the paths in the action that accompanies the yth state transition in sts_x. For example, when L is {0, 1}, $PSTS$ for $\beta\varepsilon$ in Fig. 1 is $\{<b_2, b_4b_5>, <b_2, b_4b_5b_4b_5>, <b_2b_3, b_4b_5>, <b_2b_3, b_4b_5b_4b_5>\}$.

Step 5. A set of all the measuring objects for N-SAP$_L$, which is hereinafter referred to as $MSAP_{N,L}$, is derived by the following equation:

$$MSAP_{N,L} = \bigcup_{x=1}^{|MS_N|} F(\{sts_x\} \times PSTS_x) \tag{2}$$

$F(S)$ is a feasibility evaluation function, and it removes infeasible elements from a given set S. The feasibility in EFSM models can be evaluated by using metaheuristics[3], symbolic execution, and manpower. For example, $|MSAP_{1,\{0,1\}}|$ for Fig. 1 is 28 as shown in Table 1, if there are no infeasible ones. When given test cases cover 14 elements in the $MSAP_{1,\{0,1\}}$, 1-SAP$_{\{0,1\}}$ test coverage achieves 50% (14/28). Note that the same paths often appear in different state transition sequences, but they are distinguished in $MSAP_{N,L}$. For example, $<\beta\zeta, <b_2b_3, b_6>>$ and $<\delta\zeta, <b_2b_3, b_6>>$ are strictly distinguished. Of course, the same state transition sequences with different paths are also distinguished in $MSAP_{N,L}$. For example, $<\alpha\beta, <b_1, b_2>>$ and $<\alpha\beta, <b_1, b_2b_3>>$ are strictly distinguished. Additionally, the same basic blocks that appear in different state transitions are also distinguished. For example, $<\varepsilon\gamma, <b_4b_5, b_4b_5b_4b_5>>$ and $<\varepsilon\gamma, <b_4b_5b_4b_5, b_4b_5>>$ are strictly distinguished.

4. Discussion

We developed a tool for preliminary experiments, and tried to automatically generate the measuring objects of N-SAP$_L$ from Fig. 1. The overview of its result is shown in Table 2. Note that we added special measuring objects in Step 2, and assumed that there are no infeasible ones in Step 5. This result indicates that the number of measuring objects rapidly becomes larger when a larger number and set are given to N and L respectively. For example, the following ways can be adopted to address this problem:

- Remove redundant measuring objects according to the results of fault-proneness prediction using bug-fixing record[9], identifiers in source code[10], and so on.
- Limit L to typical loop frequencies that should be tested, such as 0, 1, and multiple times.
- Automate the generation and execution of test cases that satisfy N-SAP$_L$.

Table 1. Measuring objects of 1-SAP$_{\{0,1\}}$ in Fig. 1.

state trans. seq.	**all paths in actions that accompany the state trans. seq.**
$\alpha\beta$	$<b_1, b_2>$, $<b_1, b_2b_3>$
$\alpha\gamma$	$<b_1, b_4b_5>$, $<b_1, b_4b_5b_4b_5>$
$\beta\varepsilon$	$<b_2, b_4b_5>$, $<b_2, b_4b_5b_4b_5>$, $<b_2b_3, b_4b_5>$, $<b_2b_3, b_4b_5b_4b_5>$
$\beta\zeta$	$<b_2, b_6>$, $<b_2b_3, b_6>$
$\gamma\delta$	$<b_4b_5, b_2>$, $<b_4b_5b_4b_5, b_2>$, $<b_4b_5, b_2b_3>$, $<b_4b_5b_4b_5, b_2b_3>$
$\delta\varepsilon$	$<b_2, b_4b_5>$, $<b_2, b_4b_5b_4b_5>$, $<b_2b_3, b_4b_5>$, $<b_2b_3, b_4b_5b_4b_5>$
$\delta\zeta$	$<b_2, b_6>$, $<b_2b_3, b_6>$
$\varepsilon\beta$	$<b_4b_5, b_2>$, $<b_4b_5b_4b_5, b_2>$, $<b_4b_5, b_2b_3>$, $<b_4b_5b_4b_5, b_2b_3>$
$\varepsilon\gamma$	$<b_4b_5, b_4b_5>$, $<b_4b_5, b_4b_5b_4b_5>$, $<b_4b_5b_4b_5, b_4b_5>$, $<b_4b_5b_4b_5, b_4b_5b_4b_5>$

Table 2. Number of measuring objects for N-SAP$_L$ in Fig. 1.

N	***L***			
	{0}	**{0, 1}**	**{0, 1, 2}**	**{0, 1, 2, 3}**
0	8	10	12	14
1	16	28	42	58
2	28	74	140	226
3	52	198	476	922
4	92	526	1604	3674
5	164	1390	5384	14618

Test engineers will need a tool especially to generate the test cases from their EFSM models. The difficulty of constructing the tool is in (a) solving feasibility problems on state transitions and actions, and (b) minimizing the number of the test cases. One of effective techniques for addressing (a) and (b) is metaheuristics[3]. For example, the following algorithm will enable to generate test cases that satisfy N-SAP$_L$:

Step 1. Search (execute) a given EFSM model randomly to find test case candidates. This step ensures the feasibility of test cases.

Step 2. Evaluate each candidate by using a fitness function. Good candidates include many new measuring objects without redundancy. If there are not good ones, modify the search policy as necessary, and then go to Step 1.

Step 3. Select the best candidates as a subset of final test cases. If the set of final test cases does not cover enough measuring objects, modify the search policy, and then go to Step 1.

5. Conclusion and Future Work

In this paper, we have proposed N-SAP$_L$ test coverage criterion for EFSM to comprehensively test the combination of state transitions and accompanying actions in software. A characteristic of this study is to blend a functional state-based testing technique and a structural path testing technique. The latter plays an important role in testing of the actions.

The satisfaction of N-SAP$_L$ results in the satisfaction of N'-SAP$_{L'}$, if $N' < N \wedge L' \subset L$ is satisfied. When test engineers give a larger number to N and a larger set to L in order to achieve higher software reliability, they will need to create a larger number of test cases to cover the measuring objects of N-SAP$_L$. It will not be so easy for test engineers to manually calculate N-SAP$_L$ coverage and create test cases that satisfy N-SAP$_L$. Therefore we will plan to develop a prototype tool to support such tasks.

One of challenges in future work is to develop a technique to automatically find infeasible measuring objects in EFSM models, which will make it possible to get precise N-SAP$_L$ coverage.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number JP17K00103.

References

1. M. Utting, A. Pretschner, B. Legeard, "A taxonomy of model-based testing approaches", Journal of Software Testing, Verification and Reliability, Vol.22, No.5, pp.297-312, Aug. 2012.
2. Object Management Group, "Unified Modeling Language", https://www.uml.org/.
3. A. Kalaji, R.M. Hierons, S. Swift, "Generating Feasible Transition Paths for Testing from an Extended Finite State Machine (EFSM)", Proc. of International Conference on Software Testing Verification and Validation, IEEE, Denver, CO, United States, pp.230-239, April 2009.
4. T.S. Chow, "Testing Software Design Modeled by Finite-State Machines", IEEE Transactions on Software Engineering, Vol.SE-4, No.3, pp.178-187, May 1978.
5. B. Beizer, "Software Testing Techniques", 2nd edition, Van Nostrand Reinhold, New York, NY, United States, 1990.
6. J. Fitzgerald, P.G. Larsen, P. Mukherjee, N. Plat, M. Verhoef, "Validated Designs for Object-Oriented Systems", Springer-Verlag London, 2005.
7. T. Takagi, K. Sakata, "Test-First for Abstracted Behavior of Software Using Extended Finite State Machine", Proc. of International Symposium on Software Reliability Engineering, IEEE, Coimbra, Portugal, pp.159-160, Oct. 2020.
8. K. Akiyama, "A Problem of N-Switch Coverage Testing and Its Solution", Proc. of Software Symposium, Gifu, Japan, 7 pages, July 2013 (in Japanese).
9. J.A. Whittaker, J. Arbon, J. Carollo, "How Google Tests Software", Addison-Wesley Professional, Boston, MA, United States, 2012.
10. O. Mizuno, N. Kawashima, K. Kawamoto, "Fault-Prone Module Prediction Approaches Using Identifiers in Source Code", International Journal of Software Innovation, Vol.3, Issue 1, pp.36-49, 2015.

Authors Introduction

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, particularly software testing.

Mr. Koichiro Sakata

He received the B.S. degree from Kagawa University in 2020. He is a master's student in the Graduate School of Engineering at Kagawa University. His research interests are in software engineering, particularly software testing.

Dr. Kouichi Akiyama

He received the Ph.D. degree from Kagawa University in 2013. Since 2021 he has been an internal appraiser of CMMI. His research interests are in software engineering, particularly software testing.

Proposal of a Method to Generate Classes and Instance Variable Definitions in the VDM++ Specification from Natural Language Specification

Kensuke Suga[*], Tetsuro Katayama[*], Yoshihiro Kita†,
Hisaaki Yamaba[*], Kentaro Aburada[*], Naonobu Okazaki[*]
[*] *Department of Computer Science and Systems Engineering, Faculty of Engineering, University of Miyazaki,*
1-1 Gakuen-kibanadai nishi, Miyazaki, 889-2192 Japan
[†]*Department of Information Security, Faculty of Information Systems, Siebold Campus, University of Nagasaki*
1-1-1 Manabino, Nagayo-cho, Nishi-Sonogi-gun, Nagasaki, 851-2195 Japan
E-mail: suga@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
Writing VDM++ specifications is difficult. The existing method can automatically generate type and constant definitions in VDM++ specification from natural language specification using machine learning. This paper proposes a method to generate classes and instance variable definitions in the VDM++ specification from natural language specification to improve the usefulness of the existing method. From the evaluation experiment by using F-values, it has been confirmed that the proposed method can improve the usefulness of the existing method.

Keywords: natural language specification, machine learning, VDM++ specification, automatic generation.

1. *Introduction*

The importance of software in society is increasing, and software bugs have a huge impact on our society. One of the causes of software bugs is the use of natural language in the upstream process of software development. Due to natural language containing ambiguity, the programmer might embed some bugs in the program. One way to solve this problem is to design software using formal methods in the upstream process. The development of software using formal methods is written by a formal specification description language based on mathematical logic. This allows writing specifications without the ambiguity of natural language.

VDM (Vienna Development Method) ++[1] is a formal specification language that can handle object-oriented modeling. A formal specification description language such as VDM++ is difficult to write because it has a strict syntax and requires writing data types and system invariant conditions. Traditionally, this task has depended on the experience of each programmer and has the problem of high dependency. For this reason, we proposed a method for automatically generating VDM++ specifications using machine learning by focusing on words in natural language specifications[2,3]. The existing method can classify words that are extracted from natural language specifications, into type definitions and constant definitions in VDM++ specifications. However, the existing method is unable to classify words into classes or other block definitions, so the generated VDM++ specification can only output type definitions and constant definitions. Therefore, the existing method is less useful.

In this paper, we propose a method to generate classes and instance variable definitions and apply it to the existing method in order to improve their usefulness. Here, this study focuses on specifications written in Japanese language.

2. *Existing Method*

Fig. 1 shows the flow of the method in this research. The existing method automatically generates a VDM++

specification from a natural language specification. The steps of the existing method are shown below.

1. The morphological analyzer morphologically analyzes each sentence of the natural language specification and generates a chained list that contains each sentence after analysis.
2. The converter focuses on the words of the sentences in the chained list and adds the parameters necessary for machine learning to each word. In addition, it generates a word list that contains the words and a numeric list that contains the words that are numbers.
3. The machine learning part classifies the words in the word list and generates a judgment list containing the results of the classification.
4. VDM++ generator generates a VDM++ specification using the numerical list generated in Step 2 and the words classified in Step 3.

The existing method uses morphological analysis of natural language and machine learning classification to extract the necessary words for the VDM++ specification and can automatically generate the VDM++ specification. However, there is a problem with the existing method. The existing method can only classify words extracted from natural language specifications into type definitions and constant definitions and cannot output class and other block definitions. Therefore, the existing method is less useful. In order to improve the usefulness of existing method, this paper proposes a method to generate classes and instance variable definitions in the VDM++ specification and apply it to existing method. First, we output a word list adding new parameters by extending the converter. Next, the machine learning part classifies the words in the word list into three categories.

- Words that are not necessary for the VDM++ specification.
- Words that are necessary for the VDM++ specification but are not candidates for classes.
- Words that are necessary for the VDM++ specification and are candidates for classes.

From now on, we will refer to the above words as Word A, Word B, and Word C, respectively. Finally, the machine learning part outputs a judgment list containing the classification results.

3. *Proposal Method*

The proposed method improves the functions and

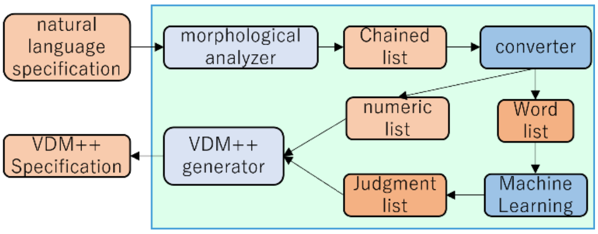


Fig. 1. Flow of the method in this research

intermediate data shown in dark colors in Fig. 1. The proposed method supports not only type definitions and constant definitions, but also classes and instance variable definitions, and automatically generates the VDM++ specification. We adopt WordNet[4] to extract candidate words for classes in the VDM++ specification from the natural language specification. WordNet is a dictionary created based on the semantic relationships between nouns, such as synonyms, superlatives, and subordinates. The steps of our proposed method are shown below.

1. The converter uses WordNet to generate a tree structure of words that are semantically related to words. In addition, the number of nodes and the root depth of the tree structure are used to calculate the concept level, which is newly defined in this research, and added to each word as a parameter.
2. The machine learning part extracts words and classifies them into Word A, Word B, and Word C.
3. In the machine learning part, the words extracted in Step 2 that are Word B are classified into words that are necessary for the VDM++ specification and are candidates for classes.
4. The machine learning part extracts words that are instance variables based on the relationship between the words classified in Step 3 and the words that are Word C in the natural language specification.

In this paper, we focus on the above steps 1-2. The classification of each word into classes and the dealing with instance variable definitions in the VDM++ specification in the steps 3-4 are future works.

3.1. *Concept Level Calculation*

The existing method outputs word list after adding four parameters to each word in the converter: TFI-DF value, number of occurrences, priority value, and number of concatenations. We extend the word list by adding a concept level for each word as a new parameter. In calculating the concept level, we use WordNet to

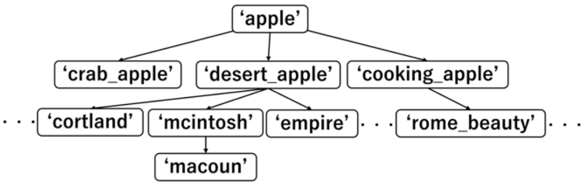


Fig. 2. Example of a tree structure of a words

generate a tree structure of words that are semantically related to the word. Fig 2 shows an example of a tree structure of words. Eq.1 shows the formula for calculating the concept level.

$$concepLevel = \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{1}{n} \tag{1}$$

$$\begin{pmatrix} m: Nodes\ with\ the\ same\ root\ depth \\ n: Depth\ of\ roots \end{pmatrix}$$

In order to classify each word into word A, word B, or word C, the proposed method adds a concept level value as a parameter to each word.

It is found that the concept level value and words in the natural language specifications had the below characteristics.

- Words with too large a concept level value (data, object, etc.) are likely to be words that are not necessary for the VDM++ specification.
- Words with too small a concept level value ('number', 'ID', etc.) are likely to be words that are not necessary for the VDM++ specification. However, if it is connected to a word that is a candidate for a class, it is most likely to be an instance variable that the class has.
- Among the words that have a concept level value between too large and too small, words with a larger concept level value are more likely to be candidates for the class.
- Otherwise, words are more likely not candidates for the class.

Based on the above features, The proposed method classifies each word into word A, word B, or word C.

3.2. *Word Classification*

The existing method uses a logistic regression model for the machine learning part to classify words into two categories: words necessary for the VDM++ specification and words not necessary for the VDM++ specification, and to output a judgment list. In the proposed method, we improve the judgment list by performing multi-class classification, which includes the judgment of words that are necessary for the VDM++ specification and words that are candidates for classes.

ユーザ認証:教員は、ユーザIDとパスワードでユーザ認証を行う。学生登録:教員は、システムを利用する学生の情報を登録できる。登録する情報は、学籍番号、氏名とする。企業登録:教員は、インターンシップを提供する企業を登録できる。登録する情報は、企業名であり、登録後、企業IDを発行する。エントリ登録:教員は、インターンシップに参加を希望する学生のエントリを登録することができる。ユーザ認証:企業担当者は、企業担当者IDとパスワードでユーザ認証を行う。インターンシップ登録:企業担当者は、インターンシップ情報を登録することができる。登録する情報は、インターンシップ名、実施日、実施日数とする。ユーザ認証:学生は、学生IDとパスワードでユーザ認証を行う。インターンシップ情報閲覧:学生は、インターンシップ情報を確認することができる。確認する項目は、インターンシップID、インターンシップ名、企業名、実施開始日、実施終了日、実施日数とする。

User authentication: Teachers can authenticate themselves with a user ID and password. Student Registration: Teachers can register the information of students who use the system. The information is to be registered in the student number and name. Company Registration: Teachers can register their companies that offer internships. The information to be registered is the company name. Teachers will be issued a company ID after registering the company name. Entry Registration: Teachers can register the entry of students who wish to participate in the internship. User authentication: Company staff can authenticate themselves with their company ID and password. Internship Registration: Company staff can register their internship information. The information to be registered in the name of the internship, the date of the internship, and the number of days of the internship. User authentication: Students can authenticate themselves with their student ID and password. Viewing Internship information: Students can check the internship information. The items to check are internship ID, internship name, company name, start date, end date, and a number of days of implementation.

Fig. 3. Japanese specification and its English translation

Word	Judgment result	Probability of Word A	Probability of Word B	Probability of Word C
教員(teacher)	Word C	0.239892	0.297224	0.462883
パスワード(password)	Word B	0.303368	0.367307	0.329323
企業(company)	Word C	0.287615	0.176724	0.53566
企業id(company id)	Word B	0.39726	0.43816	0.164578
システム(system)	Word A	0.43733	0.413029	0.14964
学生(student)	Word C	0.195058	0.204736	0.600204
学生id(student id)	Word B	0.396771	0.444731	0.158497
利用(use)	Word A	0.351601	0.39905	0.249338

Fig. 4. Word list

Word	TF-IDF value	number of occurances	Preferred value	Number of connections	concept level
教員(teacher)	0.31718	4	1	0	24
パスワード(password)	0.35217	3	2.2	0	0
企業(company)	0.47214	1	1	7	136.5
企業id(company id)	0.43043	1	1.8	0	68.2
システム(system)	0.46335	1	2	0	323.5
学生(student)	0.39363	4	1	2	25
学生id(student id)	0.44688	1	1.8	0	14.1
利用(use)	0.43692	1	1	0	39.1

Fig. 5. Judgement list

4. *Application Example*

In this paper, we extend the converter and machine learning part of the existing method and improve the output word list and judgment list. The specifications used in the application of the proposed method and its English translations are shown in Fig. 3, and part of the word list and judgment list output by the converter and machine learning part is shown in Fig. 4 and Fig. 5, respectively.

We can see that we have been able to add the concept level as a new parameter to the word list in Fig. 4. The results shown in Fig. 5 show that the nouns in the specification of Fig. 3, such as “teacher”, “company”, and “student”, can be classified as necessary and candidate class words for the VDM++ specification.

From Fig. 3 to Fig. 5, we can confirm that the proposed method is able to classify words in natural language specification properly into Word A, Word B, or Word C.

5. *Evaluation Experiment*

In order to evaluate the improvement of the usefulness of the proposed method, we experiment on the classification accuracy of words that are necessary for the VDM++ specification and are candidates for classes, using two specifications: the Internship Online Submission System Specification and the ET Robot Contest 2020 competition Rules[5]. From now on, we will refer to the two specifications as Specification A and Specification B. In the evaluation experiment, the machine learning part builds a trained model using Specification A. We evaluate the model by using F-values for the judgment lists generated from each specification. The experimental results are shown in Table 1.

Table 1 shows that the proposed method achieves a high F-value in classifying words that are necessary for the VDM++ specification and are candidates for classes, with an F-value of 0.8 for Specification A and an F-value of 0.71 for Specification B. Therefore, the proposed method can classify words that are necessary for the VDM++ specification and are candidates for classes, in addition to the existing method and achieve the improvement of the usefulness of the existing method.

6. *Conclusion*

In this paper, in order to improve the usefulness of the existing method for automatically generating VDM++ specifications from natural language specifications, we have proposed a method to generate classes in addition to type definitions and constant definitions and applied it to

Table 1. The experimental results

specification	precision	recall	F-value
Specification A	0.8	0.8	0.8
Specification B	0.6	0.86	0.71

the existing method. This corresponds to steps 1-2 of the proposed method in chapter 3.

As a result of evaluation experiments using natural language specifications, the proposed method can classify words that are necessary for the VDM++ specification and are candidates for classes with an accuracy of F-value 0.8 for Specification A and F-value 0.71 for Specification B. Therefore, it can be said that the proposed method achieves the improvement of the usefulness of the existing method.

Our future tasks are shown below.

- Classification of words necessary for the VDM++ specification into extracted classes.
- Dealing with instance variable definitions.

7. *Reference*

1. International Organization for Standardization, “ISO/IEC 13817-1:1996, Information technology - Programming languages, their environments and system software interfaces –Vienna Development Method – Specification Language -Part 1: Base language”, 1996.
2. Tetsuro Katayama and Yasuhiro Shigyo et al, “Proposal of an Algorithm to Generate VDM++ Specification Based on its Grammar by Using Word Lists Extracted from the Natural Language Specification” Journal of Robotics, Networking and Artificial Life, vol7(3), pp. 165-169, 2020.
3. Yasuhiro Shigyo and Tetsuro Katayama, "Proposal of an Approach to Generate VDM++ Specifications from Natural Language Specification by Machine Learning," 2020 IEEE 9th Global Conference on Consumer Electronics (GCCE), 2020, pp. 292-296, doi: 10.1109/GCCE50665.2020.9292047.
4. “Japanese wordnet” http://compling.hss.ntu.edu.sg/wnja/index.en.html/ (Accessed 2021-12-14)
5. “ET Robocon 2020 Simulator Competition Rules” (in Japanese) https://docs.etrobo.jp/rules/2020/ETRC2020_rules(sim)_1.0.1.pdf (Accessed 2021-12-14)

Authors Introduction

Kensuke Suga

Kensuke Suga received the Bachelor's degree in engineering (computer science and systems engineering) from the University of Miyazaki, Japan in 2021. He is currently a Master's student in Graduate School of Engineering at the University of Miyazaki, Japan. His research interests natural language processing, machine learning, and formal specification.

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.

Expansion of Application Scope and Addition of a Function for Operations into BWDM to Generate Test Cases from VDM++ Specification

Takafumi Muto*, Tetsuro Katayama*, Yoshihiro Kita†,
Hisaaki Yamaba*, Kentaro Aburada*, Naonobu Okazaki*
** Department of Computer Science and Systems Engineering, Faculty of Engineering, University of Miyazaki,*
1-1 Gakuen-kibanadai nishi, Miyazaki, 889-2192 Japan
†Department of Information Security, Faculty of Information Systems, Siebold Campus, University of Nagasaki
1-1-1 Manabino, Nagayo-cho, Nishi-Sonogi-gun, Nagasaki, 851-2195 Japan
E-mail: muto@earth.cs.miyazaki-u.ac.jp, kat@cs.miyazaki-u.ac.jp, kita@sun.ac.jp,
yamaba@cs.miyazaki-u.ac.jp, aburada@cs.miyazaki-u.ac.jp, oka@cs.miyazaki-u.ac.jp

Abstract

Generating test cases from the VDM++ specification to eliminate ambiguity in the specification is labor-intensive and time-consuming. Therefore, our laboratory developed BWDM, which is an automatic test case generation tool for VDM++ specifications. However, BWDM is not very useful because it has three problems about its narrow scope of application. This paper extends BWDM to solve three problems. In addition, we conducted a comparison experiment with manual test case generation and confirmed a time saving of about 17 minutes.

Keywords: software testing, formal methods, test cases, VDM++, automatic generation.

1. Introduction

One of the methods to eliminate the ambiguity of specifications in software design is to use formal methods for software design. One of the formal specification description languages is VDM++[1].

On the other hand, software testing is also necessary in design using formal methods, but manually generating test cases is labor-intensive and time-consuming. Therefore, we have developed BWDM, which is an automatic test case generation tool for VDM++ specifications, in our laboratory[2,3]. BWDM automatically generates test cases that can be used to perform boundary value testing, domain analysis testing, and testing based on structure recognition of if-then-else expressions.

However, BWDM is not very useful because it has the following three problems about its narrow scope of application.

- It does not support conditional expressions for invariant conditions and pre-conditions and post-conditions.
- It does not support type definition blocks.
- It is not able to generate test cases for operation definitions that manipulate the object state.

Therefore, in order to improve the usefulness of BWDM, this paper extends BWDM to solve the above three problems.

2. The Extended BWDM

The structure of the extended BWDM is shown in Fig. 1.

2.1. *Analysis and Evaluation of Each Conditional Expressions in the Definition*

To solve the problem that the existing BWDM does not support conditional expressions for invariant conditions, pre-conditions, and post-conditions set in the definition, we extend the Syntax Analyzer and Test Suite Generator of BWDM.

The Syntax Analyzer is modified to obtain inputConditions, which store the conditional expressions

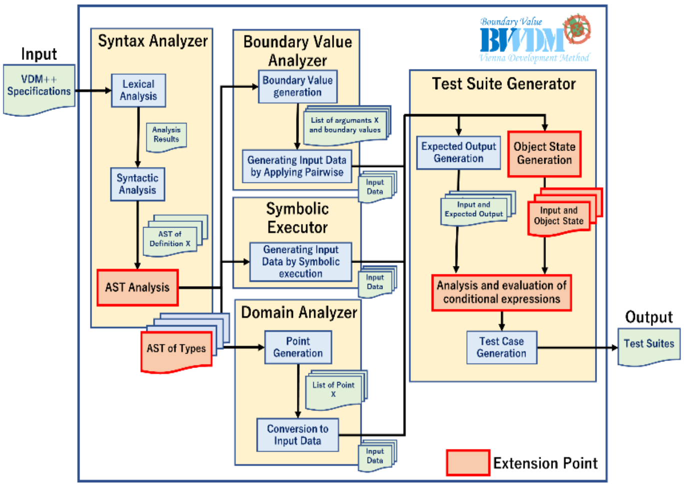


Fig. 1. The structure of the extended BWDM

needed to determine the input data, and outputConditions, which are the conditional expressions needed to determine the expected output. inputConditions store pre-conditions and invariant conditions of argument types, and outputConditions store post-conditions. The inputConditions are passed to the Boundary Value Analyzer to obtain the input data.

The Test Suite Generator adds a process to evaluate the conditional expressions of inputConditions and outputConditions during the process of generating the expected output. If the conditional expression is false, it means that the conditional expression set in the definition is not satisfied at the beginning or end of the process, so "Undefined Action" is set as an expected output. If the conditional expression is true, the expected output is the same as the existing BWDM.

2.2. *Support Type Definition Blocks*

To solve the problem that the existing BWDM does not support typedef blocks, the Syntax Analyzer is modified.

In the extended BWDM, the Syntax Analyzer keeps an abstract syntax tree of type definitions when it performs abstract parsing. When a type definition is used in each definition block, the type definition is converted to the actual type based on this abstract syntax tree. Furthermore, during conversion, conditional expressions for invariant conditions are added to inputConditions if the type to be converted is an argument type of an operation definition or function definition, or to outputConditions if it is an instance variable type.

Table 1. The expected states and the conditions corresponding to the states

expected state	conditions
Normal	All conditionals are true
Failure	The invariant condition of the instance variables definition is false.
	Post-condition is false
	The value of the instance variable after the operation is outside the range of the type.
Undefined Action	Pre-condition is false
	The invariant condition of the types definition is false
	The value is outside the range of the argument type

2.3. *Addition of a Function to Generate Test Cases for Object States*

To solve the problem that the existing BWDM cannot generate test cases for operation definitions that manipulate object states, we add a function to generate test cases for object states.

The VDM++ specification includes invariant conditions for classes and types, and pre-conditions and post-conditions for operation and function definitions. The VDM++ specification includes invariant conditions for classes and types, and pre-conditions and post-conditions set in operation and function definitions. The test cases for object states to be generated in this paper are test cases that use these conditions to generate the expected output of whether the state of the object after the operation is Normal or Failure, or whether there is an error in the input. If there is an error in the input, the expected output is set to "Undefined Action" as in the existing test case generation. Table 1 shows the expected states and the conditions corresponding to the states.

In generating the test case, the Test Suite Generator obtains the object state after the operation by using an arithmetic expression to be assigned to the instance variable and the value of the instance variable.

3. Application Example

In this chapter, we confirm that the extended BWDM works correctly by using application examples. An example of the VDM++ specification is shown in List 1, and the output of applying it as input to the extended BWDM is shown in List 2.

List 1. Example of VDM++ specification

```
class Payment
types
  public yen = nat;
values
  cardUsageLimit: yen = 100000;
instance variables
  coupon: nat := 8;
  cardUsageAmount: yen := 0;
  inv cardUsageAmount <= cardUsageLimit;
operations
  PayWithCardsAndCoupons: yen * nat ==> ()
    PayWithCardsAndCoupons(amount, tickets) ==
      (cardUsageAmount :=
              cardUsageAmount + (amount - amount *
(tickets * 0.1));
      coupon := coupon - tickets)
  pre tickets <= 10
  post coupon~ = coupon + tickets;
functions
end Payment
```

List 2. Output when List 1 is applied to the extended BWDM

```
Function Name : PayWithCardsAndCoupons
Argument Type : amount:nat tickets:nat
Return Type : ()
Number of Test Cases : 24 cases

Boundary Values for Each Argument
amount : 4294967295 4294967294 0 -1
tickets : 4294967295 4294967294 0 -1 10 11

Test Cases for Object States
No.1 : 4294967295 4294967295 -> Undefinde Action
(- Omission -)
No.11 : 0 0 -> Normal
No.12 : -1 0 -> Undefinde Action
No.13 : 4294967295 -1 -> Undefinde Action
No.14 : 4294967294 -1 -> Undefinde Action
No.15 : 0 -1 -> Undefinde Action
No.16 : -1 -1 -> Undefinde Action
No.17 : 4294967295 10 -> Undefinde Action
No.18 : 4294967294 10 -> Failure
No.19 : 0 10 -> Failure
No.20 : -1 10 -> Undefinde Action
No.21 : 4294967295 11 -> Undefinde Action
No.22 : 4294967294 11 -> Undefinde Action
No.23 : 0 11 -> Undefinde Action
No.24 : -1 11 -> Undefinde Action
```

3.1. *Confirmation of the Analysis and Evaluation of Each Conditional Expression in the Definition.*

In List 2, we have generated a test case that inputs 10 and 11 for the argument tickets. From this, we can confirm that we can obtain the boundary value from the pre-condition "tickets <= 10" in List 1. The output is "Undefined Action" when the pre-condition is not satisfied, "Failure" when the invariant and post-conditions are not satisfied, and "Normal" when all the conditions are satisfied. We can confirm that each conditional expression is evaluated correctly.

3.2. *Confirmation of Support Type Definition Blocks*

In the "PayWithCardsAndCoupons" operation shown in List 1, the argument type of "amount" is "yen", which is defined in the type definition. In List 2, we have confirmed that we can generate a test case for the boundary value of the nat type, which is the actual type of "yen".

3.3. *Confirmation of Addition of a Function to Generate Test Cases for Object States*

In the "PayWithCardsAndCoupon" operation of List 1, if 0 and 10 are used as inputs, the value of the instance variable "coupon" after the operation will be -2. Since "coupon" is of type nat and the conditional expression "0 <= coupon" stored in outputConditions is false, the expected state of the object is "Failure". In the test case No. 19 in List 2, the expected object state is "Failure" with 0 and 10 as inputs, so it can be confirmed that the test cases for the object state can be generated appropriately.

4. Discussion

4.1. *Evaluation on the Analysis and Evaluation of Each Conditional Expression in the Definition.*

We have confirmed that the extended BWDM can generate test cases corresponding to the conditional expressions in the invariant, pre-condition, and post-condition. This enables the generation of test cases corresponding to the conditional expressions set in the definitions, thus extending the range of applications of BWDM. Therefore, we can say that the usefulness of BWDM has been improved.

4.2. *Evaluation on Support Type Definition Blocks*

We have confirmed that the extended BWDM can generate test cases for VDM++ specifications using type definitions. As a result, the extended BWDM can generate test cases corresponding to the definitions described in the type definition block, and the application range of BWDM has been extended. Therefore, we can say that the usefulness of BWDM has been improved.

4.3. *Evaluation on Addition of a Function to Generate Test Cases for Object States*

We have confirmed that the extended BWDM can generate test cases for object states. As a result, the extended BWDM can generate test cases for operation definitions that manipulate the state of objects, which the existing BWDM cannot generate. Therefore, we can say that the usefulness of BWDM has been improved by adding the ability to generate test cases for object manipulation to BWDM.

4.4. *Comparison and Verification with Manual Test Case Generation*

We compared and verified the test case generation time for the object states of the extended BWDM with that of the manual case. The target VDM++ specification is the specification in List 1. The results of the comparative verification are shown in Table 2.

Manual verification was conducted by a total of five people, two graduate students, and three fourth-year undergraduates, and the time required to finish writing all the necessary test cases was measured. If the test cases were inaccurate, we pointed out the mistakes, and the time measurement ended when the subjects wrote the correct test cases.

As shown in Table 2, the time required to generate test cases using the extended BWDM was reduced by about 17 minutes compared to generating test cases manually. In addition, human error was observed in the manual test case generation.

In the function to generate test cases for object states added in this paper, it was confirmed that the time required for test case generation, which was a feature of the existing BWDM, could be reduced and that human errors could be eliminated. Therefore, we can say that the usefulness of BWDM has been improved.

Table 2. Comparison of test case generation time by the extended BWDM and manual test case generation time for the specification in List 1

	Time
Average of 5 subjects	17m19s
BWDM	1.4s

5. Conclusion

In this paper, to improve the usefulness of BWDM, it has been extended to solve the three problems.

An example of the application to the extended BWDM is shown, and it is confirmed that the above three problems have been solved.

Furthermore, as a result of comparing and verifying the time required to generate test cases manually, we were able to confirm that the time required to generate test cases using the extended BWDM was reduced by about 17 minutes. In addition, human errors were observed in the manual test case generation, and it was confirmed that human errors could be eliminated in the test case generation by using the extended BWDM.

From the above, the BWDM extended in this paper can be said to have improved its usefulness.

The following is a list of future tasks.

- Support for types other than integer types
- Support for conditional expressions that refer to the value after the operation

6. References

1. Overture Project. Manuals. http://overturetool.org/documentation/manuals.html. Accessed: 2021-12-13.
2. H. Tachiyama, T. Katayama, T. Oda. Automated Generation of Decision Table and Boundary values from VDM++ Specification. The 15th Overture Workshop: New Capa-bilities and Applications for Model-based Systems Engineering Technical Report Series, No. CS-TR-1513-2017, pp. 89-103, 2017.
3. T. Katayama, F. Hirakoba, Y. Kita, H. Yamaba, K. Aburada, and N. Okazaki. Application of Pirwise Tsting into BWDM which is a Test Case Generation tool for the VDM++ Specification. Journal of Robotics, Networking and Artificial Life, Vol.6, No.3, pp.143-147, 2019.

Authors Introduction

Takafumi Muto

Takafumi Muto received the Bachelor's degree in engineering (computer science and systems engineering) from the University of Miyazaki, Japan in 2021. He is currently a Master's student in Graduate School of Engineering at the University of Miyazaki, Japan. His research interests software testing, software quality, and formal method.

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.

Proposal of Gamma Which Is a Spatial Data Sharing Distributed MQTT System

Takahiro Ueda[*], Tetsuro Katayama[*], Yoshihiro Kita†,
Hisaaki Yamaba[*], Kentaro Aburada[*], Naonobu Okazaki[*]
[*] *Department of Computer Science and Systems Engineering, Faculty of Engineering, University of Miyazaki,*
1-1 Gakuen-kibanadai nishi, Miyazaki, 889-2192 Japan
[†]*Department of Information Security, Faculty of Information Systems, Siebold Campus, University of Nagasaki*
1-1-1 Manabino, Nagayo-cho, Nishi-Sonogi-gun, Nagasaki, 851-2195 Japan
E-mail: uedat@earth.cs.miyazaki-u.ac.jp, kat@cs.miyazaki-u.ac.jp, kita@sun.ac.jp,
yamaba@cs.miyazaki-u.ac.jp, aburada@cs.miyazaki-u.ac.jp, oka@cs.miyazaki-u.ac.jp

Abstract

This paper proposes Gamma, which is a new distributed MQTT system, to improve the usefulness of distributed MQTT systems for sharing spatial data. In the experiment, it has been found that Gamma can process more messages than a single MQTT broker. It is confirmed that Gamma achieves scalability by increasing the number of Gateways and distributed MQTT brokers. Furthermore, it is confirmed that the efficiency of the distributed MQTT system can be improved by setting the responsibility area of Gateways.

Keywords: MQTT, Spatial Data, Load management, Information management, S2 Geometry

1. Introduction

Spatial data is "data that has as its elements both location information and semantic information (state, shape, size, etc.) of objects". For example, data such as "a bicycle is traveling at 30 km/h at point A."[1] In recent years, with improvements in the processing power of smartphones, the performance of cameras and other sensors, and improvements in image processing technology, it has become possible to acquire spatial data such as the type and location of objects captured by smartphone cameras in real time[2].

MQTT (Message Queuing Telemetry Transport)[3] is a well-known application layer communication protocol for the IoT (Internet of Things). Since the basic MQTT defines one system with one broker to mediate messages, the number of Publishers to send messages and Subscribers to receive messages is limited. Therefore, there are distributed MQTT implementations that work together among multiple brokers and perform load balancing.

A previous work on distributed MQTT for sharing spatial data is the work of Ryo Kawaguchi et al[4]. They introduced a topic structure that is suitable for handling spatial data. They also introduced a distributed MQTT broker and Gateway to reduce the load on the broker and to support different types of brokers. By numerical calculations, they confirmed that each broker received fewer messages per unit time than the existing distributed MQTT system, especially when the number of Subscribers was large or when Subscriber's topic was changed frequently.

However, there are two problems as below when their proposed distributed MQTT system is used actually.

- The configuration of the distributed MQTT broker cannot be changed while the system is running because IP address and port number of the distributed MQTT broker must be registered with the Gateway in advance.

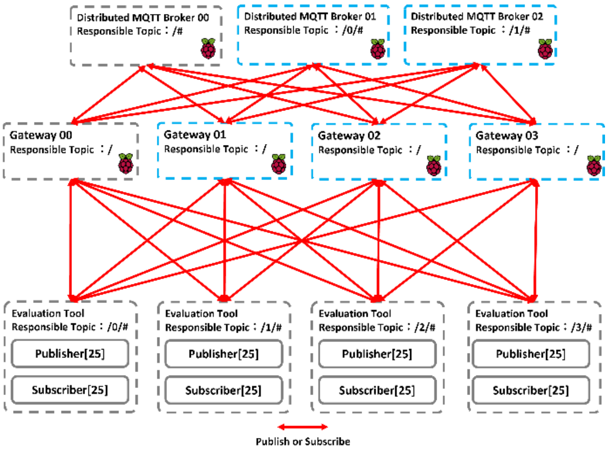


Fig. 1. System structure example (without Gateway's topic setting)

- Messages cannot be shared efficiently between the distributed MQTT broker and the Gateway because the Gateway is not configured with information on the topic it is responsible for.

Therefore, this paper proposes and implements a new distributed MQTT system, Gamma, to improve the usefulness of distributed MQTT systems for sharing spatial data. Specifically, to solve the above two problems, this paper adds Manager to the previous study[4] and implements five functions.

2. Gamma functions

In this chapter, we describe the functionality of Gamma and its client library. The Gamma implemented in this research can add a distributed MQTT broker and Gateway when the load on the system becomes large. This makes it possible to handle a larger number of messages. The followings describe the five functions implemented in this paper.

- Function to add a distributed MQTT broker while the system is running:
  This function allows system administrators to add distributed MQTT brokers while the system is running. This function is mainly performed by the Manager which is newly added in Gamma. The Manager also performs the process of updating the connection information of the distributed MQTT broker held by Gateway when a distributed MQTT broker is newly added.
- Function to add a Gateway while the system is running:

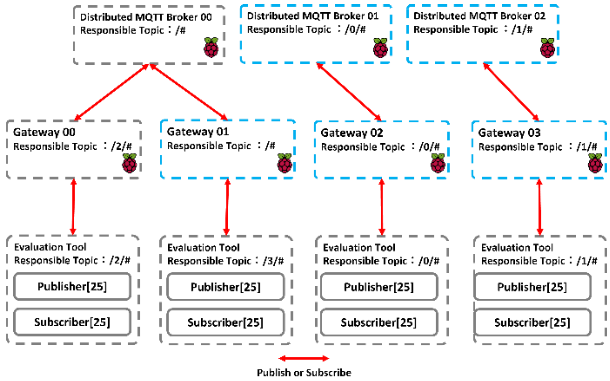


Fig. 2. System structure example (with Gateway's topic setting)

  This function allows the system administrator to add a Gateway while the system is running. The Manager receives the connection information of the newly added Gateway from it Gateway after the system starts and manages it. It is also responsible for notifying the newly added Gateway of the connection information of the distributed MQTT broker.
- Function to set the Gateway's assigned topic:
  This function allows the system administrator to set the topic for which the Gateway is responsible (topic is synonymous with area). This function is mainly performed by the Manager newly added in Gamma, which accepts the request from the system administrator to set the topic for which the Gateway is responsible and sets it. It is possible to set multiple topics for one Gateway.
- Function to select Gateway to connect based on location information:
  This function selects the Gateway to be connected based on the location information of the spatial data published/subscribed by the Client. The Client connects to the Manager before connecting to the Gateway and obtains the information about the connection to the Gateway and the topic that each Gateway is in charge of. Therefore, the Client only needs to know the connection information to the Manager. In addition, the Client can select and connect to the most suitable Gateway by comparing the topic of responsibility of each Gateway obtained from the Manager and the topic to be published/subscribed by the Client itself.

Table 1. Checking the scalability gained by increasing the number of distributed MQTT brokers

	Number of distributed MQTT brokers	1	2	3
Without responsibility area	Average delay time [ms]	14925.22	1163.05	118.92
With responsibility area	Average delay time [ms]	294.35	127.05	46.75

Table 2. Checking the scalability gained by increasing the number of Gateways

	Number of Gateways	1	2	3	4
Without responsibility area	Average delay time [ms]	11423.46	233.15	118.73	42.65
With responsibility area	Average delay time [ms]	11423.46	116.14	57.30	15.96

- Function to publish/subscribe to spatial data based on location information:
  This function allows the client to publish/subscribe to spatial data based on location information. The client calculates a topic to publish based on the latitude and longitude of the spatial data, and can publish the spatial data. It can also calculate topics for subscribing to spatial data in that range from the given latitude, longitude, and radius, and subscribe to it. This function uses S2 Geometry[5] for these calculation.

3. Evaluation of Gamma

We evaluate the usefulness of Gamma implemented in this paper.

Firstly, we compare it with the single MQTT broker. As a result, we confirmed that Gamma is capable of processing more messages than the single MQTT broker.

Next, we will conduct an experiment to evaluate the scalability that can be obtained by increasing the number of Gateways. Fig. 1 shows the configuration of the system when no topic is assigned to the Gateway, and Fig. 2 shows the configuration of the system when a topic is assigned to the Gateway.

The experiments shown below were conducted under the following conditions: the length of the transmitted message was 150 characters, the number of Publishers was 100, the number of Subscribers was 100, the publish interval for each Publisher was 100 milliseconds, and the measurement time was 100 seconds. In addition, Publishers and Subscribers were equally placed in four areas.

3.1. *Experiments and Discussion on Scalability Gained by Increasing the Number of Distributed MQTT Brokers*

Check the scalability by varying the number of distributed MQTT brokers; the number of Gateways is fixed to four.

Here, when the upper limit of the CPU utilization of the distributed MQTT broker is viewed in the entire command line tool called cputool[6] during the experiment, the percentage of the load occupied by the Gateway is higher than that of the distributed MQTT broker. In addition, there were only four Gateways that could be prepared in this paper. However, even when there are four Gateways for only one distributed MQTT broker, the problem happens that the processing capacity of the Gateway became insufficient before that of the distributed MQTT broker. Therefore, in this experiment, we will intentionally reduce the processing capacity of the distributed MQTT broker using cputool and confirm the scalability obtained by increasing the number of distributed MQTT brokers.

From the experimental results in Table 1, we can see that the average delay time decreases with the number of distributed MQTT brokers. Because of this, we can see that increasing the number of distributed MQTT brokers improves the processing capacity of the entire system. Therefore, we were able to confirm that scalability is ensured by increasing the number of distributed MQTT brokers.

In addition, Table 2 shows that the average delay time is smaller in the case of "with topic setting" than in the case of "without topic setting," regardless of the number of Gateways. From this, it can be said that the setting of the topic to the Gateway, which is one of the functions of

Gamma implemented in this paper, has improved the efficiency of the entire system.

3.2. *Experiments and Discussion on Scalability Obtained by Increasing the Number of Gateways*

Check the scalability by varying the number of Gateways. The number of distributed MQTT brokers is fixed to one.

Table 2 shows the results of the experiment when the number of Gateways is varied, and it can be confirmed from Table 2 that the average delay time decreases with the number of Gateways regardless of whether the topic for which the Gateway is responsible is set or not. In other words, by increasing the number of Gateways, the processing capacity of the entire system can be improved. It can be confirmed that Gamma implemented in this paper ensures scalability by increasing the number of Gateways.

From Table 2, we can see that the average delay time is smaller in the case of "with topic setting" as shown in Fig. 2 than in the case of "without topic setting" as shown in Fig. 1. Therefore, it can be said that the setting of the assigned topic in the Gateway, one of the functions of Gamma implemented in this paper, has improved the efficiency of the entire system.

4. Conclusion

This paper has proposed and implemented a new distributed MQTT system, Gamma, to improve the usefulness of distributed MQTT systems for sharing spatial data.

A previous study of distributed MQTT for sharing spatial data is the work of Ryo Kawaguchi et al[4]. This previous research has two problems, which are described in Chapter 1.

In order to solve these two problems, Gamma adds Manager to the previous study[4], and is implemented five functions. We have confirmed that the five functions of Gamma properly are run. This confirms that Gamma solves the problem of not being able to change the configuration of the distributed MQTT broker while the system is running.

The experiment also showed that Gamma can process more messages than a stand-alone MQTT broker, and it has been confirmed that scalability can be ensured by increasing the number of Gateways and distributed MQTT brokers. Furthermore, it has been confirmed that the efficiency of the entire system can be improved by setting the topic of responsibility of the Gateway so that the load is not unevenly distributed to a specific Gateway. From this, it has been confirmed that the problem of not being able to efficiently share messages between the distributed MQTT broker and the Gateway because the information on the topic of responsibility is not set in the Gateway could be solved.

From the above, Gamma realized the following two points and solved the two problems of the previous studies.

- The ability to add a distributed MQTT broker while the system is running
- The ability to notify the Client of the information of multiple Gateways and the topics for which the Gateway is responsible.

In summary, we have confirmed that Gamma proposed and implemented in this paper is useful.

Future issues are as follows.

- Support for automatic addition of Gateways and distributed MQTT brokers based on load.
- Add the ability to monitor the activity and death of Gateway and distributed MQTT brokers
- Support for single-level wildcards

References

1. “Utilizing Spatial Data: The Future Visible through Objects and Locations” (in Japanese), https://www.nttdata.com/jp/ja/data-insight/2020/033001/ (Accessed 2021-12-14)
2. “DRiVR”, https://drivr.app/ (Accessed 2021-12-14)
3. “MQTT Version 3.1.1”, http://docs.oasis-open.org/mqtt/mqtt/v3.1.1/os/mqtt-v3.1.1-os.html (Accessed 2021-12-14)
4. Ryo Kawaguchi and Masaki Bandai, “A Distributed MQTT Broker System for Location-based IoT Applications”, 2019 IEEE International Conference on Consumer Electronics (ICCE), pp.1-4, 2019.
5. “S2 Geometry”, https://s2geometry.io/ (Accessed 2021-12-14)
6. “cputool”, http://manpages.ubuntu.com/manpages/bionic/man8/cputool.8.html (Accessed 2021-12-14)

Authors Introduction

Takahiro Ueda

Takahiro Ueda received the Bachelor's degree in engineering (computer science and systems engineering) from the University of Miyazaki, Japan in 2021. He is currently a Master's student in Graduate School of Engineering at the University of Miyazaki, Japan. His research interests software development support method and networking.

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.

A Three-Dimensional Design of the Multi-material Joint System to Realize a Structural Spring-Damper Compliant Mechanism with Versatility in Engineering Fields

Pancho Dachkinov[1]
[1]Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu-Ku,
Kitakyushu, 808-0196, Japan

Anirudha Bhattacharjee[2]
[2]Indian Institute of Technology Kanpur, Kanpur-208016, India

Bishakh Bhattacharya[2]
[2]Indian Institute of Technology Kanpur, Kanpur-208016, India

Hiroaki Wagatsuma[1, 3]
[1]Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu-Ku,
Kitakyushu, 808-0196, Japan
[3]RIKEN CBS, Japan
E-mail: : dachkinov.pancho-nikolaev608@mail.kyutech.jp, anirub@iitk.ac.in, bishakh@iitk.ac.in, waga@brain.kyutech.ac.jp

Abstract

Design of a 3D printed cross-spring compliant joint is an emerging topic for its multipurpose applications in various fields due to its realization from a combination of flexible materials with different mechanical properties. It performs the motion by deformation in the elastic region and is suitable for precision engineering applications and instruments. The proposed concept is a modification of a traditional cross-spring pivot, which effectively provides frictionless and wear free in-plane motion. The joint's behavior is analyzed based on a non-linear FEA simulation and the properties were investigated with various loading conditions. Compliant joints are envisaged to bring paradigm shift in the design of high-precision actuators and robotic manipulators.

Keywords cross-spring pivot, compliant mechanisms, FEA simulation, multi-material joint

1. Introduction

Compliant mechanisms are mechanisms that rely on deformation caused by an external loading to perform a motion. In the recent years there is an increased interest in the implementation of compliant mechanisms due to the capabilities of 3D Printing as an appropriate technology for fabricating complex geometry [1]. The advancements in the Additive Manufacturing lead not only to realization of complex geometry but also to the development to variety of new filaments with different behavior and properties. One of the most commonly used materials is Poly Lactic Acid (PLA). This a relatively low cost and environmentally friendly filament used in the FDM (Fused Deposition Modeling) 3D Printing. The mechanical properties of this material are explored at a tensile and flexural test by [2] and [3]. PLA has been compared with other popular 3D Printing filaments such as ABS (Acrylonitrile Butadiene Styrene) in [4]. Other works investigate the relative density of flexible cellular structures fabricated from (thermoplastic polyurethane) [5]. TPU models are flexible and provide larger deformations in the elastic region. Combining the above factors with the unique capability to control the geometry of the model from inside – its orientation and density unlike any traditional technology, make 3D printing suitable for application in the compliant mechanisms allowing for greater range of motion due to the larger deformations. However, there are certain challenges that must be considered while developing compliant mechanisms:

- when large displacements occur, there are nonlinearities in the geometry and the material.
- depending on the orientation of the load, the stiffness of the part can vary in orders of magnitude.
- finding the balance between stiffness and flexibility is a complex design problem.
- to prevent the mechanism from fatigue failure of the material, high stress concentrations should be avoided.
- determining the desired motion parameters is a difficult task.

In rigidly articulated joints the clearances between the components causes backlash in the assemblies. Further, there is friction due to relative motion leading to wear of the parts and increasing the clearances and generating heat in the joint [6]. All of the above disadvantages result in poor accuracy and performance.

On the other hand, to address these issues researchers in the last years have focused on the compliance of the joints rather than restricting such deformations. Compliant joints provide adaptive and monolithic motion, avoiding entirely the assembly process.

Flexural hinges have been widely studied in the literature and several applications in machine design which harness the advantages of flexure-based design elements have been realized. Being monolithic in structure and with the advent in Additive manufacturing techniques, it is even simpler nowadays to design, prototype, test and verify the performance of these hinges in various applications, especially in the field of Precision Engineering. The most popular type of flexural hinge is the Cross-spring pivot, commercially known as Free-flex joints.

Primitive flexural joints primarily consist of leaf-type or notch-type configurations, but they have their own inherent limitations. Notch-type hinges have high stiffness in transverse directions and the rotation capability is limited by stress concentrations. However, the leaf-type flexures provide large rotations but at expense of drifting rotation centers. In-order to mitigate these issues several complex flexural hinges have been proposed like the cartwheel hinges [7], butterfly hinges [8]. While these hinges provide greater benefits over conventional leaf or notch type, Cross-spring pivots are relatively simpler and versatile in applications. The diversity of these flexures stems from the fact that they were extensively explored in the literature for decades. They were first studied analytically in [9] and more recently in [10]. Goncalves provides a thorough theoretical formulation and laser based optical experiments to characterize the performance of Cross-spring pivots.

Such accuracy provided by compliant pivots is useful for small displacement in precise mechanisms and instruments [11]. Relaying on material deformation, flexure joints have become major component in development of precise instruments with high resolution of positioning.

Th other sections of this paper are structured as follows: Section 2 explains methodology of the design and the mathematical modelling of the multi-material joint; Section 3 focuses on the results; Section 4 gives the discussion and Section 5 concludes this paper.

2. Methodology

2.1. *Design of the Compliant Joint*

As discussed in the previous section, compliant mechanisms and joints are becoming a replacement of the traditional equivalents in certain applications regarding their benefits. In the current study, a new compliant type of universal joint realized from two materials is proposed. Rigidly articulated universal joints have two perpendicular axes of rotation where two pin joints are connected. The motion occurs as two rotations in orthogonal planes.

The design of the compliant multi-material joint consists of two cross-spring pivots rotated at a right angle according to the Z-axis of the origin coordinate system as shown on Figure 1.

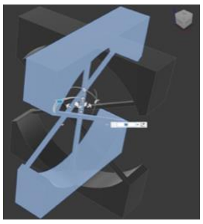


Fig. 1. Orientation of the Cross-Spring Pivots Compounding the Universal Compliant Joint

That type of orientation allows the joint to deform according to X- and Y-axis and provides stability for out of plane movements and moments around the Z-axis. Another advantage of the configuration is minimal axial drift of the center of the joint.

2.2. *Modelling and Physical Prototype*

As illustrated on Figure 2 below, in XY plane the model of the joint consists of three beams oriented at 60º angle (α) from each other (beams A, C and E shown in Figure 5). Another identical unit is attached to it at a perpendicular orientation around Z-axis. All bars of the joint have same dimensions: length (L), and square cross-section (a). The offset between beams from a single cross-spring joint is also equal to (a) and the beams do not intersect. A general view representation of the joint is shown on Figure 2. The joint has two platforms – top and bottom (1) and five beams (2).

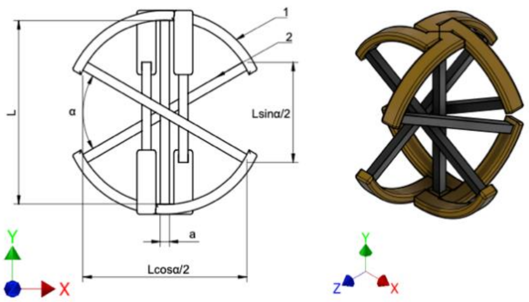


Fig. 2. General View of the Compliant Joint

On Figure 3 are shown a) a 3D rendered model and b) a physical 3D printed prototype of the compliant multi-material joint. The platforms are printed from a solid material PLA (polylactic acid) and the beams from a flexible filament – TPU (Thermoplastic Polyurethan). This multi-material combination provides unique capabilities. The platforms that attach to other components are rigid but the beams that execute the motion throughout deformation are flexible.

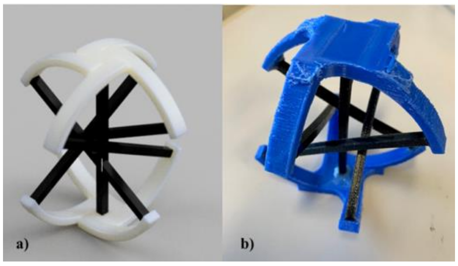


Fig. 3. A Model and 3D Printed Prototype of the Compliant Multi-material Joint

This two-material approach provides diversity for combining and experimenting with different materials pairs. However, this joint is not a single unit, the platforms and the beams have been 3D Printed separately and later assembled.

Technology, chosen for rapid prototyping of this specimen is Fused Deposition Modeling (FDM). The filament used for printing the model is PLA (polylactide) and TPU (Thermoplastic Polyurethan with diameter of 1.75 mm. The 3D printer used for fabricating the sample is Anycubic i3 Mega S, a desktop type machine. 3D printing process is described in Table 1. As discussed in [8], 3D printing parameters could impact the anisotropic behavior of the fabricated parts. Therefore, the configuration set up of the working process is essential for the properties of the physical prototypes. For configuring the printing process is used slicing software Ultimaker Cura. The 3D model was designed in Autodesk Fusion 360 CAD software. The flexible TPU filament is called Ninjaflex from the company Ninjatek and according to their recommendations the appropriate 3D Printing temperature of the nuzzle (extruder)is 230 ºC [12]. The infill pattern used for the beams is called Cross 3D described in the slicing software as suitable for 3D deformations and flexible material prints [13].

Table 1: 3D Printing Configurations

Parameter	***Value of the Printing Parameter***	
	Platforms	***Beams***
Filament	PLA	TPU
3D Printer	Anycubic i3 Mega s	
Slicin Software	Ultimaker Cura	
Layer Height	0.1 mm	0.1 mm
Infill Density, %	100	100
Infill Pattern	Grid	Cross 3D
Printing Temperature, ºC	210	230
Building Plate Temperature, ºC	60	60

Parameter	*Value of the Printing Parameter*	
	Platforms	***Beams***
Filament	PLA	TPU
3D Printer	Anycubic i3 Mega s	
Slicin Software	Ultimaker Cura	
Print Speed, mm/s	45	30
Support	yes	no

2.3. *Analysis of the Loading Conditions*

The performance of joint has been analyzed under various loading conditions illustrated alongside with the deformations on Figure 4 such as a) compression, b) tension, c) bending of the model. On the right side of the figure, the platforms are removed, and the direction of the forces is annotated for the different loading cases.

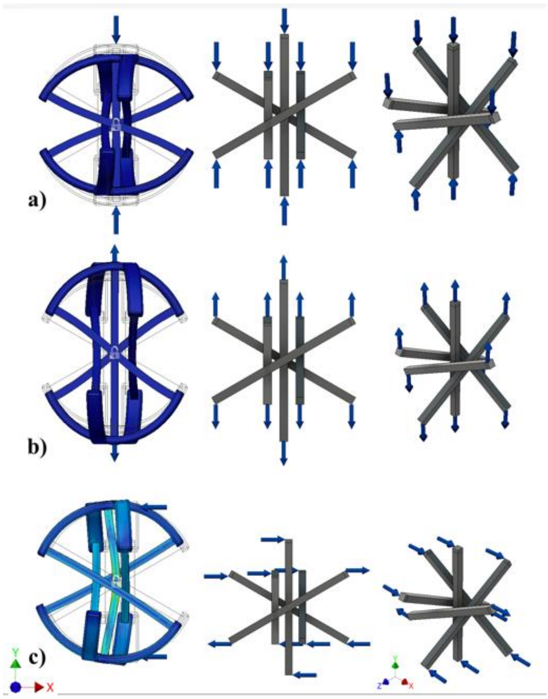


Fig. 4. Loading Conditions of the Joint

To better visualize the position and orientation of the joint components in the space, on Figure 5 the top and bottom platforms are removed, and the beams are color coded and annotated with letters. The coordinate systems in the different orientations are given as well.

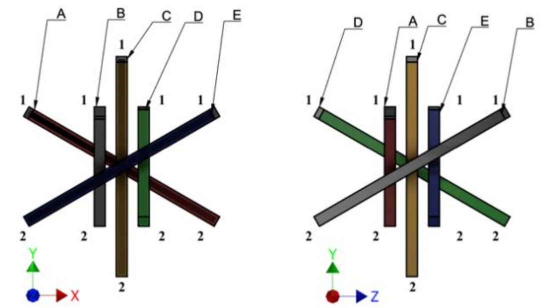


Fig. 5. Position and Orientation of the Beams

On Figure 6 are illustrated all unique loading condition of the beams taken from Figure 6 tension, compression and bending. There are three beams that have different force distribution and using the annotation of Figure 5 they are A, B and C. Beams D is symmetrical to B and E is sumetrical to A and therefore, forces with same magnitude and direction are acting on them.

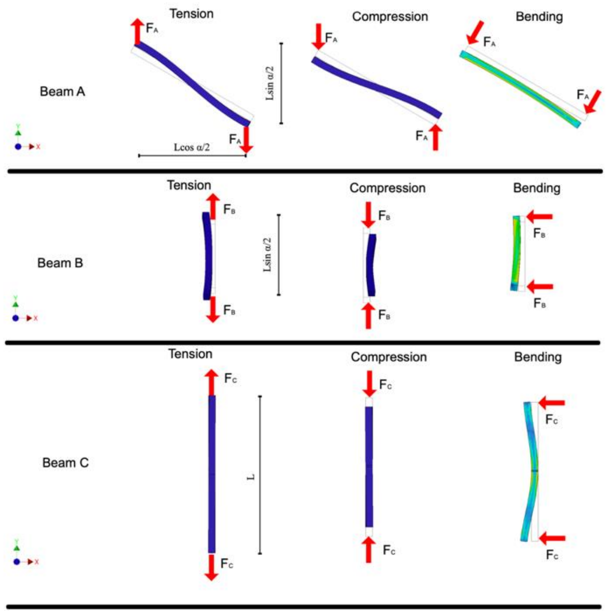


Fig. 6. Loads Acting on the Individual Beams

2.4. *Mathematical Modelling of the Cross-spring Pivots*

The proposed design can be modelled as an assembly of two Cross-Spring Pivots (CSPs) with free center of rotations. Behavior of CSP is widely studied in literature analytically [9,10] and experimentally [10, 11]. In figure 5 beams A and E corresponds to the leaf springs of the pivot 1 and beams B and D, rotated at 90 degrees along the Y-axis consists of pivot 2. The beam C at the center of the joint acts as a stiffener and provides additional rotational stability to the joint.

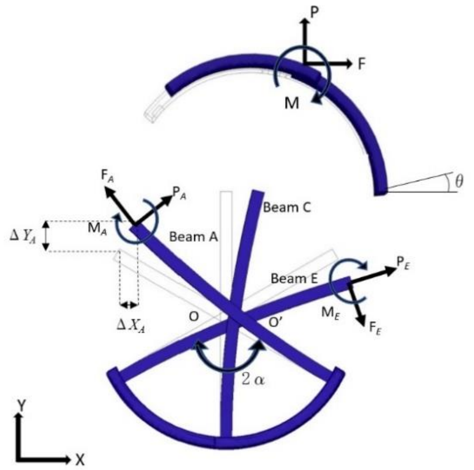


Fig. 7. Generalized Forces and Moments acting on the CSP

Figure 7 illustrates the generalized forces and moments acting on the CSPs and the behavior is described by 11 variables [14]. The load rotational relationship of a cross-spring pivot under a generalized loading condition M-couple, F-horizontal loads, P- vertical loads is characterized by 11 variables (θ , P_A , P_E , F_A , F_E , M_A , M_E , ΔY_A , ΔY_E , ΔX_A and ΔX_E). Similarly, the pivot 2 is characterized by the same set of variables corresponding to beams B and D. For each pivot a system of 11 equations is required to solve the problem. Five are obtained from the equilibrium conditions of forces in X and Y directions, the moment balance, and the compatibility at the edges of beams A and E. The remaining six are obtained from curvature equations for each beam as given in equation 1.

$$EI\frac{d^2y/_{dx^2}}{\left[1+\left(^{dy}/_{dx}\right)^2\right]^{\frac{3}{2}}} = P_{A(E)}y + M_{BA(E)} - F_{A(E)}x \qquad (1)$$

Where, subscripts A and E refers to the beams as shown in figure 7.

3. Results

Simulation studies were carried out in Comsol Multiphysics. The material properties used for the analysis are illustrated in Table 2.

Table 2: Materials Properties used in the FEA

Material	Young's Modulus	Poisson Ratio	Density
Polylactic Acid (PLA)	3.5 GPa	0.3	1.24 g/cm3
Thermoplastic Polyurethane (TPU)	15.6 MPa	0.346	1.23 g/cm3

On Figure 8 are shown the result from the studies for tension a) and b) and bending c) and d). As it can be observed, in b), there is a rotation of the top platform due to the tensile forces. The results for compression are similar to the tension with the reversed direction of the rotation.

The displacements corresponding to axial load ranging between 2 N and 20 N with the increment of 2 N are shown on the plots of Figure 9 a). On Figure 9 b) the displacements related to the transverse load are given, the range of which is 0.05 N to 0.2 N with the step size of 0.05 N.

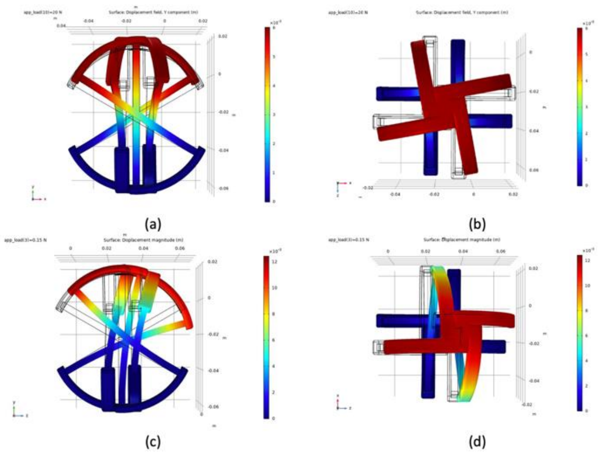


Fig. 8 – Illustrates of the Strains Induced due to Axial Loading (a, b) and Transverse Loading (c, d).

As illustrated a nonlinear behavior of the model can be observed when large deformations occur due to the nonlinear properties of the joint.

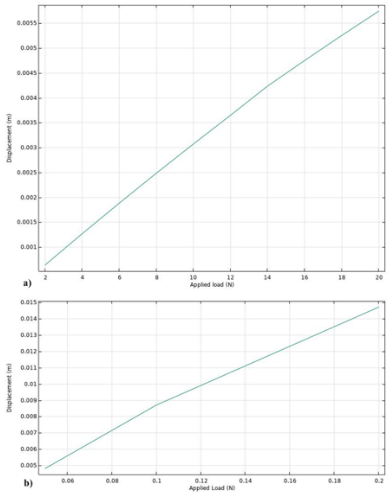


Fig. 9 – Displacement vs Load Plot a) Axial Loading (tension/compression); b) Transverse Loading (bending)

4. Discussion

The FEA simulation studies yields conforming results to the physical behavior of the design. The proposed configuration of the cross-spring pivot provides greater rotational stability due to the additional center leaf placed along the axis of the joint. However, in transverse plane when axial forces applied additional rotation can be observed about Z-axis. Designing such compliant joint for specific applications can be improved by parameterization of center-shift phenomenon and characterizing the mechanical properties of the flexible 3D printing materials.

The proposed design of a cross-spring joint has geometry similar to the ligaments of the human knee joint. The ligaments are tissue made from small fibers of collagen, twisted like a rope that attaches bones to other bones in the human body and provides stability of the joint. The application of the proposed compliant multi-material joint has been inspired by the knee ligament application since their geometrical similarities such as precision adjustments unbalanced forces in small displacement sensitive instruments. The proposed improved design of a cross-spring pivot is an ideal candidate owing to the degree of compliance it provides.

In addition, the joint can be 3D printed as one monolithic unit avoiding the assembly process as shown on Figure 10 (extracted from [14]). This prototype, made from PLA has the purpose to illustrate the model can be fabricated at once without assembling the parts. This can be especially applicable for 3D Printer using metal powder such as SLS (selective laser sintering) machines.

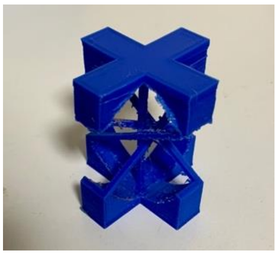


Fig. 10. Single 3D Print Prototype

The plans for future development of the current design extends to conduct experimental studies of a 3D printed prototype by 3D Laser Doppler Vibrometer (LDV) to validate its performance characteristics such as stress, displacements, rotational stiffness, natural frequency, etc. Both single and multi-material samples will be studied in this context as well. The frequency analysis will determine the vibration damping responses of the joint, to be used for stabilizing purposes. To optimize the stabilizing properties of the joint, 3D printing parameters - infill density, infill geometry, layer thickness, etc. - can be tuned (as mentioned in [15] and [16] the infill orientation, geometry and density could impact the stress performance of the printed samples).

5. Conclusion

In this paper a novel type of compliant multi-material joint is presented. The joint consists of two cross-spring compliant joints perpendicular to each other. The pivot has two rotations around its center according to X-axis and Y-axis. It has relatively small axial drift and stability in other planes of motion.

The design and methodology of creating the joint are described and 3D printing parameters are illustrated. The forces have been dissolved for all beams individually and depending on the various loading conditions as compression, tension and bending. A static FEM analysis has been conducted to demonstrate the behavior of the model.

Among the highlights of this model are its compactness – the rotational axis has matching center points. Its compliance leads to frictionless and wear free motion which preserves the reliability and durability in time. The multi-material pairing can be varied, and different material couple can be used depending on the requirements. Another advantage of its compliance is the adaptivity it has due to the flexibility of the elastic deformations.

There are several challenges that need to be considered in the design stage. Experiments need to be performed to determine the flexure from plastic buckling due to greater forces. Another important point is the material's fatigue. Further investigation needs to be conducted on fatigue testing of the joint. Lastly, it has relatively complex geometry which makes its fabrication by using traditional technologies with removing material nearly impossible task.

On the other hand, the rapid development of 3D Printing and Additive Manufacturing in general allows more creativity in the design stage and provides further future opportunities before the manufacturing in the context of the Fourth Industrial Revolution.

Acknowledgment

This work was supported in part by Bilateral Joint Research Project (DST-JSPS/India-Japan), EU Horizon 2020 MSCA RISE project "Cyber-Physical Systems for PEdagogical Rehabilitation in Special EDucation (CybSPEED)", JSPS KAKENHI (16H01616, 17H06383) and the New Energy and Industrial Technology Development Organization (NEDO), 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).

References

[1] Megaro, V., Zender, J., Bacher, M., Coros, S., Cross, M., Thomaszewski, B., A Computational Design Tool for Compliant Mechanisms, ACM Transactions on Graphics, Vol. 36, No. 4, Article 82, July 2017.

[2] Aveen K., P.,, Bhajathari F., V., Sudhakar Jambagi, S., C., 3D Printing & Mechanical Characterition of Polylactic Acid and Bronze Filled Polylactic Acid Components, IConMMEE 2018, IOP Conf. Series: Materials Science and Engineering 376(2018), doi:10.108

[3] Catana,D.;Pop,M.-A.; Brus, D.-I. Comparison between the Test and Simulation Results for PLA Structures 3D Printed, Bending Stressed. *Molecules* 2021, *26,* 3325. https://doi.org/10.3390/ molecules26113325.

[4] Ahmend M., Islam, Md., R., Vanhoose, J., Hewavitharana, L., Stanich, A., Hossain, M., Comparisons Of Bending Stiffness Of 3d Printed Samples Of Different Materials, Proceedings of the ASME 2016 International Mechanical Engineering Congress & Exposition, I

[5] Płatek, P., Rajkowski K.,¸ Cieplak, K., Sarzyn´ski, M., Małachowski, J., Woz´niak, R., Janiszewski, J., Deformation Process of 3D Printed Structures Made from Flexible Material with Different Values of Relative Density, Polymers 2020, 12, 2120; doi:10.

[6] Trease, B., Moon, Y,. Kota, S., Design of Large-Displacement Compliant Joints, Transactions of the ASME, Journal of Mechanical Design, Vol. 127/788-798, July 2005.

[7] Pei, X., Yu, J., Zong, G., Bi, S., Su, H., The modeling of cartwheel flexural hinges. Mech. Mach. Theory 44 (10), 1900–1909.,2009.

[8] Henein, S. , Spanoudakis, P. , Droz, S. , Myklebust, L.I. , Onillon, E., Flexure pivot for aerospace mechanisms. In: Proceedings of the 10th European Space Mechanisms and Tribology Symposium. San Sebastian, Spain, 2003.

[9] Haringx, J.A., The cross-spring pivot as a constructional element. Flow Turbul. Combust. 1 (1), 313–332, 1949 .

[10] L.A. Gonçalves Junior ,R. Theska, H.A. Lepikson , A.S. Ribeiro Junior, S. Linß, P. Gräser, Theoretical and experimental investigation of performance characteristics and design aspects of cross-spring pivots. International Journal of Solids and Structure

[11] Liu. L., Bi, S., Yang, Q., Wang, Y., Design and experiment of generalized triple-cross-spring flexure pivots applied to the ultra-precision instruments, AIP Publishing, Review of Scientific Instruments 85, 105102, 2014.

[12] Ninjatek, Ninjaflex, TPU 3D Printing Filament, Technical specifications, 2021/11/21, https://ninjatek.com/shop/ninjaflex/#print-guidelines.

[13] Ultimaker Cura, 3D Printing slicing software, Infill pattern guide, 2021/11/21. https://all3dp.com/2/cura-infill-patterns-all-you-need-to-know/

[14] Zelenika, S., Bona, F.D., Analytical and experimental characterisationof high- precision flexural pivots subjected to lateral loads. Precis. Eng.26 (4), 381–388, 2002.

[15] Dachkinov, P., Bhattacharjee A., Bhattacharya B., Wagatsuma, H.,, A Knee Joint Inspired Design of a Compliant Cross-Spring Pivot, International Symposium on Applied Engineering and Sciences 2020

[16] Firbman, D., McCoy, C., Materials Testing of 3D Printed Abs and PLA Samples to Guide Mechanical Design, MSEC2016-8668, July, 2016.

[17] Letcher, T., Waytashek, M., Material Property Testing of 3D-Printed Specimen in PLA on an Entry-Level 3D Printer, IMECE2014-39379, November 2014

Authors Introduction

Mr. Pancho Dachkinov

He received his Master's degree in Engineering in 2018 from the Technical University of Sofia, Bulgaria. He worked in the Institute of Robotics – Bulgarian Academy of Sciences 2016-2019. He is currently a doctoral student in Kyushu Institute of Technology, Japan.

Mr. Anirudha Bhattacharjee

He received his M.E. degree in Production Engineering in 2014 from the Faculty of Engineering and Technology, Jadavpur University in India. He is currently working as Research Associate at SMSS Laboratory in Department of Mechanical Engineering at Indian Institute of Technology Kanpur. His research interest is Robotics and Computational Geometry.

Dr. Bishakh Bhattacharya

He is a Professor, Department of Mechanical Engineering at IIT Kanpur, India. He is currently the HAL Chair for the period of Feb 2021 - Jan 2024. He received his Ph. D. in Aerospace Engineering from IISc Bangalore, India in 1998. His area of research encompasses: Active and Passive Vibration Control, Structural Health Management, Design of Energy Harvesting System, Intelligent System Development and Child-Robot Interaction Design.

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 Intelli- gence Systems. His research interests include non- linear dynamics and robotics. He is a member of IEEE.

Haptic Device that Presents Sensation Corresponding to Palm on Back of Hand for Teleoperation of Robot Hand
Report 5: Verification of development device specifications

Kyosuke Ushimaru*, Noritaka Sato, Yoshifumi Morita
Department of Electrical and Mechanical Engineering, Graduate School of Engineering
Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya, Aichi 466-8555, Japan
E-mail: k.ushimaru.884@nitech.jp

Abstract

In this research, we propose and develop a new tactile presentation device for tele-operation of a robot hand. The proposed device presents tactile sensation on the back of the operator's hand. The required specification for device development were determined by subject experiments in our previous studies. In this paper, we developed prototype device using suction stimulus with determined required specification and confirmed that the developed device can present tactile sensation correctly and discuss the results.

Keywords: *Rescue robot, Haptic Device, Teleoperation, Robot Hand*

1. Introduction

After the Great Hanshin-Awaji Earthquake in 1995, research and development of disaster response robots gained momentum [1]. In recent years, disaster response robots that can gather information and perform tasks are required [2]. Therefore, disaster response robots equipped with robot hands have been developed to work at disaster sites to achieve human-like characteristics, such as pulling out pipes and removing debris [3]. Disaster response robots are operated in unknown and extreme environments; therefore, they are controlled by teleoperation [4]. To make teleoperation more efficient with a robot hand, it is necessary to provide feedback to the operator on the contact state between the fingers and palms of the robot hand and the object to be grasped. In this study, the tactile sensation is substituted by the back of the hand instead of the palm, that is presented just behind the palm; therefore, we think that the operator can recognize the tactile state more accurately.

In the first report [5], the characteristics of tactile sensation on the back of the human hand were clarified to obtain the required specifications for creating a device. In the second report [6], we examined whether the tactile sensation can be presented as a surface instead of a point without changing the specifications. In the third report [7], we examined whether it was possible to present tactile sensations with suction stimulation instead of pressure stimulation. In the fourth report [8], we investigated relationship between the diameter of suction pad and the correct answer rate, and we found that the larger the diameter was, the higher the correct answer rate was. In this study, a prototype of the proposed device is fabricated based on the results obtained from previous experiments [8], and its performance is verified through subject experiments.

2. Substitute Tactile Presentation Device for Back of Hand

To correctly convey the state of contact between the robot hand and the grasping object to the operator ideally is to present the tactile sensation directly to the palm of the hand, as shown in Fig. 1. However, when the robot hand is directly tele-operated by the operator's hand, the haptic device may interfere with the work. In this study, we propose a method to present the tactile sensation on the back of the hand, as shown in Fig. 2.

2.1. *Parameter identification*

In designing such a back-of-the-hand substituting tactile presentation device, the following parameters are required:

- Interval between stimulus points (interval: *i*)
- Diameter of the stimulus point (diameter: *d*)
- Force of the stimulus point (force: *f*)

Human skin sensation has characteristics such as a two-point discrimination threshold. Too close stimulation points and very sparse stimulation points does not present an appropriate contact situation. Therefore, we conducted an experiment on subjects to confirm the characteristics of the tactile sensation on the back of the hand for our proposed device [5]. As a result, the combination (i, d, f) = (30 mm, 6 mm, 0.9 kg) had the highest rate of correct answers.

However to fabricate a tactile presentation device, suction stimulation is utilized instead of the pressure stimulation because the device will become very large with pressure stimulation. Therefore, we designed the device as shown in Fig. 3. In the third and fourth paper [7] [8], we checked how well the pressure stimulation position of the palm matched the suction stimulation position of the back of the hand. In addition, we found that the correct answer rate for suction stimulation was highest when (i, d, f) = (30mm, 15mm, 0.9kgf).

2.2. *Fabricated device*

The prototype device shown in Fig. 4 was fabricated based on the specification determined in our past studies. The device consists of a suction pad, a suction bracket, and a fixture created by a 3D printer. According to the determined parameters, the diameter of the suction pads is 15 mm and the interval between each pad is 30 mm.

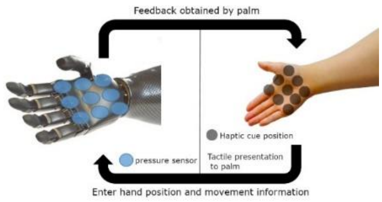


Fig. 1. Ideal method

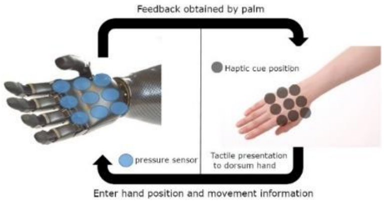


Fig. 2. Proposed method

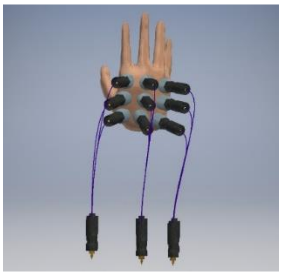


Fig. 3. Designed device with parameters obtained by experiment

2.3. *Substitutional Tactile Presentation System*

Fig. 5 shows the system configuration of the whole tactile presentation device for the back of the hand explained in Section 2.2. The device provides suction stimulation to the user. A vacuum pump is GCD-051X (ULVAC), and 9 regulators for vacuum pressure control are ITV2090 (SMC). The vacuum pressure can be controlled by the input voltage of the regulator, and a micro-computer (Arduino) outputs the input voltage to each regulator. 9 surface pressure sensors are used to convert the contact strengths to voltages, which are read by the Arduino.

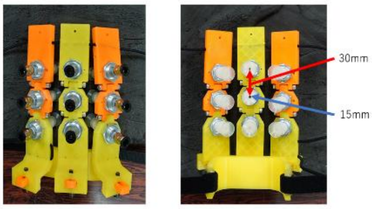


Fig. 4. Prototype device
(Left: Top view, Right: Bottom view)

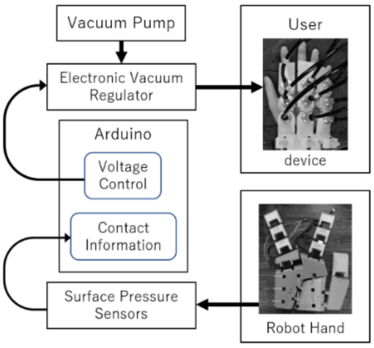


Fig. 5. Overview of the system for the experiment

3. Performance verification experiment of a tactile presentation device

3.1. *Experimental Methods*

In this study, we report the results (accuracy rates) of an experiment where two operators answer the position of a suction stimulation when it is applied to the back of the operator's hand using a prototype device. ten points on their backs of the hand are pressed one by one.

The scenes of the experiment are shown in Figs. 6 and 7. The stimulus presentation system for the experiment consisted of devices that provided stimuli to the back of the hand. Fig. 7 shows the devices of subject side. The prototype device, tubes, and a vacuum pump are used. The stimulation position is shown in Fig.8. There is a one-to-one relationship between the surface pressure sensor and the sucker. Suction stimulation was applied to the back of the hand in the pre-determined order, as shown in Table 1.

3.2. *Results and Discussion*

In this experiment, each subject was given 10 stimulation and the total number of collected data was 20. The graph of the results is shown in Fig. 9. Total accuracy rate was 80.0%. Accuracy rate of each subject was 80% and both subjects had 20% of wrong answers. However, we think that they recognized the position of the stimulus almost correctly, because the wrong answered position ware located to above or below position of the correct one. In addition, we measured response time from starting of stimulus to answering by subject. In most cases, the response was made within 5 seconds. Therefore, we think that the results are reliable because it was confirmed that the subjects answered without hesitation.

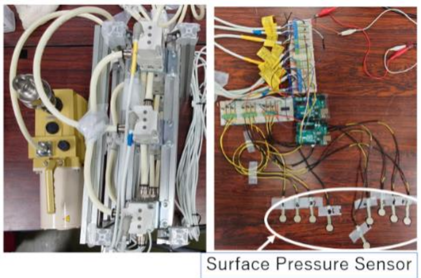


Fig. 6. Experimental scene
(Control side)

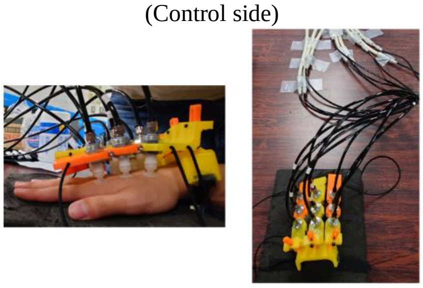


Fig. 7. Experimental scene
(Subject side)

Fig. 8. The number of stimulation point.

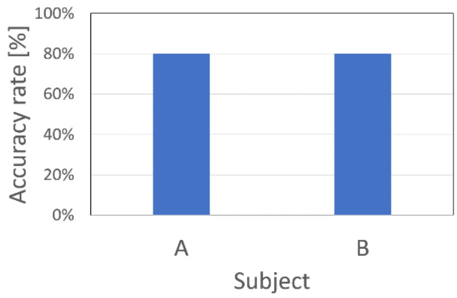


Fig. 9. Total accuracy rate.

Table 1. Order of Stimulus presentation.

Order of stimulation	Suction stimulation
1	5
2	7
3	8
4	3
5	6
6	4
7	5
8	2
9	1
10	9

4. Conclusion

In this study, we fabricated a device based on the required specification determined in our previous studies and performed subject experiments to verify the performance. As a result, we obtained 80% of the correct answers. Therefore, we think that this device has potential to be used in teleoperation of a robot hand.

In the future, at first, we will increase the number of subjects. After that, we will confirm if the users can recognize accurately when multiple stimulus points are given. Finally, the proposed device is verified in the teleoperation task.

Acknowledgements

This work was supported by a JSPS Grant-in-Aid for Scientific Research (JP19K04309).

References

1. S Tadokoro, et al., Rescue Robotics: DDT Project on Robots and Systems for Urban Search and Rescue, Springer-Verlag London.
2. S Tadokoro, Disaster robotics, Advances in Cooperative Robotics, pp. 3, 2017.
3. S. Tadokoro, Overview of the ImPACT Tough Robotics Challenge and Strategy for Disruptive Innovation in Safety and Security, Disaster Robotics, pp. 3-22, 2019.
4. N Sato, F Matsuno: "User Interface Technology for Rescue Robot Teleoperation," Journal of the Robots Society of Japan, Volume 28, Issue 2, Pages 156-159, 2010. (In Japanese).
5. N. Sato, K. Ushimaru, Y, Yoshifumi: Haptic Device that Presents Sensation Corresponding to Palm on Back of Hand for Teleoperation of Robot Hand Report 1: Clarification of Required Specifications by Subject Experiment", The 38th Annual Conference of the RSJ, 2I3-03, 2020. (In Japanese).
6. K. Ushimaru et al: Experimental Consideration on Requirements Specification of Haptic Device that Presents Sensation Corresponding to Palm on Back of Hand for Teleoperation of Robot Hand , ICAROB2021, GS6-3, 2021.
7. K. Ushimaru et al: Experimental Consideration on Requirements Specification of Haptic Device that Presents Sensation Corresponding to Palm on Back of Hand for Teleoperation of Robot Hand , Tokai-Section Joint Conference on Electrical, Electronics, Information, and Related Engineering, E2-2, 2021. (In Japanese),.
8. K. Ushimaru et al: Haptic Device that Presents Sensation Corresponding to Palm on Back of Hand for Teleoperation of Robot Hand, Report 4: Consideration of the Detection Error of the Stimulus Position between the Palm and the Back of the Hand, International Symposium on Socially and Technically Symbiotic Systems, STSS-35, 2021.

Authors Introduction

Mr. Kyosuke Ushimaru

He received the B.S. degree in Engineering in 2019 from Nagoya Institute of Technology. He is currently a Master course student in Nagoya Institute of Technology. His research interests include virtual reality, haptic interface, and rescue robots.

Dr. Noritaka Sato

He is an assistant professor of Nagoya Institute of Technology. He received the Ph.D degree of Engineering from The University of Electro-Communications in 2009. His research interests include human robot interaction for rescue robots, rehabilitation-assist robots

Prof. Yoshifumi Morita

He is a Professor of Department of Electrical and Mechanical Engineering, Graduate School of Engineering, Nagoya Institute of Technology, Japan. His research interests are robot motion control, support control for human-machine cooperative system, and rehabilitation support robot.

HBV Epidemic Control Using Time-Varying Sliding Mode Control Method

Arsit Boonyaprapasorn*
Department of Mechanical Engineering, Chulachomklao Royal Military Academy,
Suwanasorn Rd., Muang, Nakhon-Nayok, 26001, Thailand.

Suwat Kuntanapreeda
Department of Mechanical and Aerospace Engineering, Faculty of Engineering,
King Mongkut's University of Technology North Bangkok,
1518 Pracharat 1 Rd., Wongsawang, Bangsue, Bangkok, 10800, Thailand.

Parinya Sa Ngaimsunthorn
Department of Mathematics, Faculty of Science, King Mongkut's University of Technology Thonburi,
126 Pracha Uthit Rd., Bang Mod, Thung Khru, Bangkok, 10140 Thailand.

Thunyaseth Sethaput
Sirindhorn International Institute of Technology, Thammasat University,
99 Moo 18, Km. 41 on Paholyothin Highway Khlong Luang, Pathum Thani, 12120, Thailand

Tinnakorn Kumsaen
Department of Chemical Engineering, Faculty of Engineering, Khon Kaen University,
23 Moo 16 Mittraphap Rd., Nai-Muang, Muang District, Khon Kaen, 40002, Thailand
E-mail: urarl.urarl9@gmail.com, suwat.k@eng.kmutnb.ac.th, parinya.san@kmutt.ac.th,
thunyasethsiit.tu.ac.th, tinnakor@kku.ac.th
** Corresponding Author*

Abstract

Hepatitis B virus (HBV) infection is one of the life-threatening diseases due to causing cirrhosis and liver cancer in the infected person. Setting the policy to control the HBV epidemic is an important issue that can be achieved by using feedback controller design procedure through the compartment model. In this article, the sliding mode controller with a time-varying sliding surface was utilized to set the multiple measures control policy for controlling the HBV epidemic. The stability of the control HBV epidemic system was examined. The simulation of the control system was conducted to confirm the feasibility of applying the time-varying sliding mode controller for setting the HBV control policy. The simulation results showed that the designed control policy could drive the target subpopulation to the desired levels. The convergence rate of the control HBV system could be improved. Thus, the time-varying sliding mode controller is a feasible approach to set the measures for controlling the HBV epidemic.

Keywords: Hepatitis B, Epidemic system, Compartment model, Time-varying sliding mode control.

1. Introduction

Hepatitis B is a viral infection causing high morbidity and mortality rate from both chronic and acute liver infection[1-3]. This virus is hepatotropic DNA virus in the family of Hepadnaviridae[1, 3, 4]. The spread of hepatitis B is caused by both vertical (perinatal) and horizontal transmission routes through infected body fluids. After being exposed to the HBV, the disease can develop either acute or chronic hepatitis B[2,7,8]. Acute hepatitis B patients typically have the symptoms such as fever, fatigue, dark urine, and jaundice. Chronic infection refers to the long-term inflammation leading to liver damage. This damage causes cirrhosis and liver cancer which are deadly diseases[3, 6]. According to WHO[2], in 2019 almost three million people were chronic hepatitis B, and more than 800,000 people were dead. A feasible treatment for acute HBV patients is required to provide the comfort and to balance nutrient and fluids of the patients[2,6]. For the chronic HBV patients, the purpose of treatment is to reduce the viral load to prevent cirrhosis and liver cancer. This can be achieved by using the immunomodulatory and antiviral drugs[2,3,8]. The vaccine is an important preventive measure of the hepatitis B in susceptible individuals since it can provide high effective protection and safe[2,5]. For epidemic control, other measures such as treatment for HBV patients, screening and diagnosis of high-risk population can also be included in the control policy isolation[3,9,10].

Many mathematical models representing the dynamic of HBV epidemic have been developed based on different compartment models[10-17]. With these compartment models, control policies including measures such as vaccination, and treatment and isolation controls can be drawn by using dynamic optimization approach according to Pontryagin's maximum principle[10-17]. Another potential approach is to use feedback control approach to set the policy[18-30]. Feedback control could deal with the uncertainties and disturbances which occur in the system. Moreover, it is an efficient approach to set the policy in an analytical form[31]. Applying feedback control to define the measures of the control policy showed the feasibility and the efficiency in the previous studies of this approach[18-30]. Sliding mode control (SMC) is a robust feedback control which has been employed to control various nonlinear dynamical systems[32-34]. Moreover, the SMC method has been applied to synthesize the control policy for the biological systems such as ecosystems and epidemic models[18,24,27-30, 35-36]. For the epidemic systems, the sliding mode control was applied to set the control policy of the epidemic systems in the form of a compartment model[18,24,27-30]. The SMC with an integral sliding surface can eliminate the steady state errors of the control systems as applied to control various engineering systems[37,40,41]. Specifically, this SMC has been used for controlling epidemic systems[18, 27]. Ibeas et al. [18,27] used the integral sliding mode control to set the vaccination control policy with robustness for handling the uncertainties of the model parameters. The improvement in terms of the convergence rate of the sliding mode control systems can be achieved by using the time-varying sliding mode control (TVSMC) where the constant sliding surface is replaced by the time-varying sliding surface. As shown in Refs. 37-40, 42-44, the time varying sliding surface is defined such that the initial condition is located on or crossed by the sliding surface. The development and applications of the TVSMC method can be found in the literature[37-40, 42-44].

Even though the feasibility of applying feedback control for controlling the HBV system was presented in Ref. 21, the robustness of the control policy was not presented. As integral sliding mode control can provide the robustness and the concept of time-varying sliding surface can improve the convergence rate of a control system, this study focuses on studying of applying time-varying sliding mode control (TVSMC) with the integral sliding surface to set the HBV control policy with multiple measures based on the compartment model. To the best of authors' knowledge, setting the HBV control policy with multiple measures using the time-varying sliding mode controller design procedure has not been presented.

The rest of this paper consists of four following sections. The mathematical model of the HBV system is presented in Section 2. Section 3 provides the details of the time varying sliding mode controller design for setting the control policy. The simulation of the control system including a simulation example and simulation results is presented in Section 4. The conclusion is stated in Section 5.

2. Model of HBV Epidemic System

The mathematical model presented by Ullah et al. [17] was used for setting the control policy based on the time-varying sliding mode controller design procedure. This model consists of six subpopulations which are susceptible (S), exposed (E), acute infected (A), carrier (C), hospitalized (H), and recovered (R) individuals. The mathematical model from Ref. 17 is presented in (1):

$$\left.\begin{aligned}
\dot{S}(t) &= b - \psi bC - dS - (A + \mu C)\beta S(1-u_1) \\
\dot{E}(t) &= (A + \mu C)\beta S(1-u_1) + \psi bC - dE - \delta E \\
\dot{A}(t) &= \delta E - h_1 A - dA - d_A A - \gamma A - Au_2 \\
\dot{C}(t) &= \gamma A - dC - d_C C - h_2 C - Cu_2 \\
\dot{H}(t) &= h_2 C + h_1 A - (d+\xi)H + Au_2 + Cu_2 - Hu_3 \\
\dot{R}(t) &= \xi H - dR + Hu_3
\end{aligned}\right\}. \quad (1)$$

The parameters of this HBV system are presented as follows [17]: the rate b and the rate d represent the birth and natural death rates of the population in the HBV system, respectively. The rate h_1 and the rate h_2 refer to the rates of acute and carrier individuals who were hospitalized, respectively. The coefficient β is the transmission coefficient. The parameter δ defines the flow rate of changing from the exposed subpopulation to the carrier subpopulation. The mortality rate caused by acute infection is denoted by d_A. The carrier individual death rate is defined by d_C. The rate γ defines the rate of changing from the acute subpopulation to the carrier subpopulation. The recovery rate is defined by ξ. The rate of the unimmunized children who were born to the carrier mothers is denoted by ψ. The parameter μ is the carriers' infectiousness caused by acute infection. The first control measure, $u_1(t)$, is the isolation measure for preventing transmission between infected and uninfected individuals. The second measure, $u_2(t)$, refers the human intervention which consists of public education and awareness of sending the infected persons to the hospital. The third measure, $u_3(t)$, represents the treatment for the hospitalized individuals. All of these measures are constrained as $0 \le u_1, u_2, u_3 \le 1$ [17]. The total population of the HBV epidemic system is defined by $N(t) = S(t) + E(t) + A(t) + C(t) + H(t) + R(t)$.

After defining subpopulations of the HBV system in (1) as $S = x_1$, $E = x_2$ $A = x_3$, $C = x_4$, $H = x_5$, and $R = x_6$, the input-output model of the system in (1) can then be expressed as (2):

$$\dot{x} = f(x) + g(x)u, \quad (2)$$

where $x = [x_1\ x_2\ x_3\ x_4\ x_5\ x_6]^T$, $u = [u_1\ u_2\ u_3]^T$,

$$f(x) = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \end{bmatrix} = \begin{bmatrix} b - \psi bx_4 - dx_1 - (x_3 + \mu x_4)\beta x_1 \\ (x_3 + \mu x_4)\beta x_1 + \psi bx_4 - dx_2 - \delta x_2 \\ \delta x_2 - h_1 x_3 - dx_3 - d_A x_3 - \gamma x_3 \\ \gamma x_3 - dx_4 - d_C x_4 - h_2 x_4 \\ h_2 x_4 + h_1 x_3 - (d+\xi)x_5 \\ \xi x_5 - dx_6 \end{bmatrix}, \text{ and}$$

$$g(x) = \begin{bmatrix} g_{11} & 0 & 0 \\ g_{21} & 0 & 0 \\ 0 & g_{32} & 0 \\ 0 & g_{42} & 0 \\ 0 & g_{52} & g_{53} \\ 0 & 0 & g_{63} \end{bmatrix} = \begin{bmatrix} (x_3 + \mu x_4)\beta x_1 & 0 & 0 \\ -(x_3 + \mu x_4)\beta x_1 & 0 & 0 \\ 0 & -x_3 & 0 \\ 0 & -x_4 & 0 \\ 0 & (x_3 + x_4) & -x_5 \\ 0 & 0 & x_5 \end{bmatrix}.$$

3. Controller Design

3.1. *Time-varying Sliding Mode Controller Design*

Based on the control objective in Ref. 17, the control objective of TVSMC controller design procedure is defined to manipulate the exposed, acute, hospitalized subpopulations to the desired values. The error corresponding to the control objective is defined as (3):

$$\{e_2 = x_2 - x_{2r},\ e_3 = x_3 - x_{3r}\ \ e_5 = x_5 - x_{5r}\}. \quad (3)$$

where x_{2r}, x_{3r}, and x_{5r} are the desired values and set as $x_{2r} = x_{3r} = x_{5r} = 0$ [17].

Based on Refs. 32-34, 37, 39-41, and 45, the design procedure of synergetic control is presented as follows.

It is referred to Refs. 37, 39, and 40, the time-varying integral sliding surface are defined as

$$\left.\begin{aligned}
s_2 &= e_2 + k_{I2}\int_0^t e_2(\tau)d\tau + \phi_2 \\
s_3 &= e_3 + k_{I3}\int_0^t e_3(\tau)d\tau + \phi_3 \\
s_5 &= e_5 + k_{I5}\int_0^t e_5(\tau)d\tau + \phi_5
\end{aligned}\right\}, \quad (4)$$

where $\phi_2 = m_{c2}e^{-t/n_{c2}}$, $\phi_3 = m_{c3}e^{-t/n_{c3}}$, and $\phi_5 = m_{c5}e^{-t/n_{c5}}$. The controller parameters $k_{I2}, k_{I3}, k_{I5}, n_{c2}, n_{c3} n_{c5}$ are real positive numbers. The coefficients, m_{c2}, m_{c3}, and m_{c5} are defined based on the sliding surfaces at the initial time, $s_2(0) = s_3(0) = s_5(0) = 0$. This yields that $m_{c2} = -e_2(0)$, $m_{c3} = -e_3(0)$, and $m_{c5} = -e_5(0)$ [37, 39, 40].

The classical constant plus proportional reaching law[45] or exponential reaching law[34] was utilized in this study as (5):

$$\left.\begin{aligned} \dot{s}_2 &= -k_{sw2}sign(s_2) - k_{p2}s_2 \triangleq \theta_2 \\ \dot{s}_3 &= -k_{sw3}sign(s_3) - k_{p3}s_3 \triangleq \theta_3 \\ \dot{s}_5 &= -k_{sw5}sign(s_5) - k_{p5}s_5 \triangleq \theta_5 \end{aligned}\right\}. \tag{5}$$

Substituting sliding surfaces (4) into the set of reaching laws (5) yields

$$\left.\begin{aligned} \dot{e}_2 + k_{I2}e_2 + \dot{\phi}_2 &= \theta_2 \\ \dot{e}_3 + k_{I3}e_3 + \dot{\phi}_3 &= \theta_3 \\ \dot{e}_5 + k_{I5}e_5 + \dot{\phi}_5 &= \theta_5 \end{aligned}\right\}.. \tag{6}$$

From the HBV dynamic system, (6) can be obtained as

$$\begin{aligned} &f_2(x) - (x_3 + \mu x_4)\beta x_1 u_1 - \dot{x}_{2r} + k_{I2}e_2 + \dot{\phi}_2 = \theta_2 \\ &f_3(x) - x_3 u_2 - \dot{x}_{3r} + k_{I3}e_3 + \dot{\phi}_3 = \theta_3 \\ &f_5(x) + (x_3 + x_4)u_2 - x_5 u_3 - \dot{x}_{5r} + k_{I3}e_3 + \dot{\phi}_3 = \theta_3. \end{aligned} \tag{7}$$

The control measures can be obtained as (8):

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} -(x_3 + \mu x_4)\beta x_1 & 0 & 0 \\ 0 & -x_3 & 0 \\ 0 & (x_3 + x_4) & -x_5 \end{bmatrix}^{-1} \begin{bmatrix} -\varphi_2 + \theta_2 \\ -\varphi_3 + \theta_3 \\ -\varphi_5 + \theta_5 \end{bmatrix}, \tag{8}$$

where $\varphi_2 = f_2(x) - \dot{x}_{2r} + k_{I2}e_2 + \dot{\phi}_2$,
$\varphi_3 = f_3(x) - \dot{x}_{3r} + k_{I3}e_3 + \dot{\phi}_3$,
and $\varphi_5 = f_5(x) - \dot{x}_{5r} + k_{I5}e_5 + \dot{\phi}_5$.

3.2. *Proof of Stability*

According to Liu and Wang[34], the stability of the control HBV system can be investigated based on the Lyapunov stability. The Lyapunov function of the control HBV system is defined as (9):

$$V = 0.5s_2^2 + 0.5s_3^2 + 0.5s_5^2. \tag{9}$$

The derivative of the Lyapunov function is obtained as (10):

$$\begin{aligned} \dot{V} &= s_2\dot{s}_2 + s_3\dot{s}_3 + s_5\dot{s}_5 \\ \dot{V} &= s_2[\dot{e}_2 + k_{I2}e_2 + \dot{\phi}_2] + s_3[\dot{e}_3 + k_{I3}e_3 + \dot{\phi}_3] \\ &\quad + s_5[\dot{e}_5 + k_{I5}e_5 + \dot{\phi}_5] \end{aligned} \tag{10}$$

Assuming that there are external bounded disturbances in the $\dot{x}_2$, $\dot{x}_3$, and $\dot{x}_5$ of (2), $\dot{V}$ in (10) is obtained as (11):

$$\begin{aligned} \dot{V} &= s_2[f_2(x) - (x_3 + \mu x_4)\beta x_1 u_1 + d_2(t) - \dot{x}_{2r} + k_{I2}e_2 + \dot{\phi}_2] \\ &\quad + s_3[f_3(x) - x_3 u_2 + d_3(t) - \dot{x}_{3r} + k_{I3}e_3 + \dot{\phi}_3] \\ &\quad + s_5[f_5(x) + (x_3 + x_4)u_2 - x_5 u_3 + d_5(t) - \dot{x}_{5r} + k_{I5}e_5 + \dot{\phi}_5] \\ \dot{V} &= \begin{bmatrix} s_2 & s_3 & s_5 \end{bmatrix} \left(\begin{bmatrix} \varphi_2 \\ \varphi_3 \\ \varphi_5 \end{bmatrix} + \begin{bmatrix} d_2(t) \\ d_3(t) \\ d_5(t) \end{bmatrix} + \begin{bmatrix} -(x_3 + \mu x_4)\beta x_1 & 0 & 0 \\ 0 & -x_3 & 0 \\ 0 & (x_3 + x_4) & -x_5 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \right) \end{aligned} \tag{11}$$

where $|d_2(t)| \le D_2$, $|d_3(t)| \le D_3$, and $|d_5(t)| \le D_5$ for $D_2, D_3, D_5 \ge 0$

Substituting (8) into (11) yields

$$\begin{aligned} \dot{V} &= \begin{bmatrix} s_2 & s_3 & s_5 \end{bmatrix} \left(\begin{bmatrix} \varphi_2 \\ \varphi_3 \\ \varphi_5 \end{bmatrix} + \begin{bmatrix} d_2(t) \\ d_3(t) \\ d_5(t) \end{bmatrix} + \begin{bmatrix} -\varphi_2 + \theta_2 \\ -\varphi_3 + \theta_3 \\ -\varphi_5 + \theta_5 \end{bmatrix} \right) \\ &= \begin{bmatrix} s_2 & s_3 & s_5 \end{bmatrix} \begin{bmatrix} -k_{sw2}sign(s_2) - k_{p2}s_2 + d_2(t) \\ -k_{sw3}sign(s_3) - k_{p3}s_3 + d_3(t) \\ -k_{sw5}sign(s_5) - k_{p5}s_5 + d_5(t) \end{bmatrix}. \end{aligned}$$

$$\begin{aligned} \dot{V} &= -k_{p2}s_2^2 - k_{p3}s_3^2 - k_{p5}s_5^2 + s_2 d_2(t) + s_3 d_3(t) + s_5 d_5(t) \\ &\quad - k_{sw2}|s_2| - k_{sw3}|s_3| - k_{sw5}|s_5|. \\ \dot{V} &\le -k_{p2}s_2^2 - k_{p3}s_3^2 - k_{p5}s_5^2 + |s_2|D_2 + |s_3|D_3 + |s_5|D_5 \\ &\quad - k_{sw2}|s_2| - k_{sw3}|s_3| - k_{sw5}|s_5| \end{aligned} \tag{12}$$

If k_{sw2}, k_{sw3}, and k_{sw5} are chosen as $k_{sw2} = D_2 + \kappa_2$, $k_{sw3} = D_3 + \kappa_3$, and $k_{sw5} = D_5 + \kappa_5$, where κ_2, κ_3, $\kappa_5 > 0$, it can be obtained that

$$\begin{aligned} \dot{V} &\le -k_{p2}s_2^2 - k_{p3}s_3^2 - k_{p5}s_5^2 \\ &\quad - \kappa_2|s_2| - \kappa_3|s_3| - \kappa_5|s_5| < 0. \end{aligned} \tag{13}$$

The inequality (13) implies that control measures $u_1(t)$, $u_2(t)$, and $u_3(t)$, can stabilize the control HBV system under the bounded disturbances. Based on (4), at $s_2 = s_3 = s_5 = 0$, the error e_2, e_3, and e_5, approach to zero as time increases. Thus, the target subpopulations of the control system can track the reference signals[37,38].

4. Simulation

The control measures synthesized in (8) were applied to the simulation example of the HBV system in Section 4.1. The simulation results are presented in Section 4.2.

4.1. *Simulation Example*

The HBV epidemic system in (2) with system parameters from Ref. 17 was used as a simulation example to show the feasibility of applying TVSMC controller design procedure to set the HBV epidemic control policy. The parameters of the system and the initial conditions presented in Ref. 17 are as follows: $b = 0.5$, $d = 0.008$, $h_1 = 0.2$, $\beta = 0.5$ ($0 \le \beta \le 1$), $\delta = 6$, $\gamma = 0.5$ ($0 \le \gamma \le 1$), $\mu = 0.5$ ($0 \le \mu \le 1$), $d_C = 0.005$, $d_A = 0.005$, $\delta = 6$, $\xi = 0.1$, and $\psi = 0.2$. The initial condition was assumed based on information provided in Ref. 17 and denoted as $x(0) = [180, 40, 18, 10, 20, 0]^T$. The incremental time step is 0.01 year. The initial time and final time were t=0 day and t=10 years respectively[17]. The Runge-Kutta method was used for numerical integration.

The robustness of the control policy was demonstrated by adding the bounded disturbance into the rate of change of the hospitalized subpopulation as (14):

$$d(t) = [0 \; d_3(t) \; 0]^T, \tag{14}$$

where $d_3(t)$ is defined as (15):

$$d_3(t) = \begin{cases} 0, t < 2 \\ d_m \sin(\omega_d t) + \varphi_d, \; 2 \le t \le 4, \\ 0, t > 4 \end{cases} \tag{15}$$

where $d_m = 0.5$, $\omega_d = 20\pi$, and $\varphi_d = 2.5$.

The controller parameters for this simulation example are set as follows: $k_{I2} = 0.01$, $k_{I3} = 0.01$, $k_{I5} = 0.01$. $k_{p2} = 2$, $k_{p3} = 2$, $k_{p5} = 2$, $k_{sw2} = 4$, $k_{sw3} = 4$, $k_{sw5} = 4$ $n_{c2} = n_{c3} = n_{c5} = 0.2$.

In order to show the capability of the TVSMC policy in terms of convergence, the simulation results of the designed control policy were compared with those of the SMC policy using the sliding surface in (4) with $\phi_2 = \phi_3 = \phi_5 = 0$.

4.2. *Simulation Results*

The time responses of interested subpopulations which are the exposed, acute, hospitalized under TVSMC and SMC policies are shown in Fig. 1. The control measures of both TVSMC and SMC polices are presented in Fig. 2. The exposed, acute, and hospitalized subpopulations of the TVSMC policy converge to the zero faster than those the SMC policy as shown in Fig. 1. The isolation measures of both TVSMC and SMC start from certain levels slightly below their maximum levels. Then, they increase rapidly to their maximum levels and stay on these levels for rest of time as shown Fig 2. In Fig. 2, the human intervention and the treatment measure of both policies are at the maximum level from the beginning until the end of time. It is clear that the policy synthesized by TVSMC design procedure can suppress the HBV epidemic. Also, improvement of the convergence rate of the control system can be achieved.

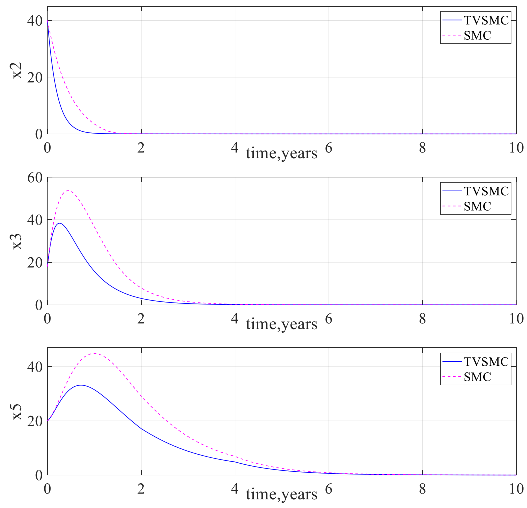


Fig. 1. Time response of interested subpopulations including exposed (x_2), acute (x_3), and hospitalized (x_5) subpopulations.

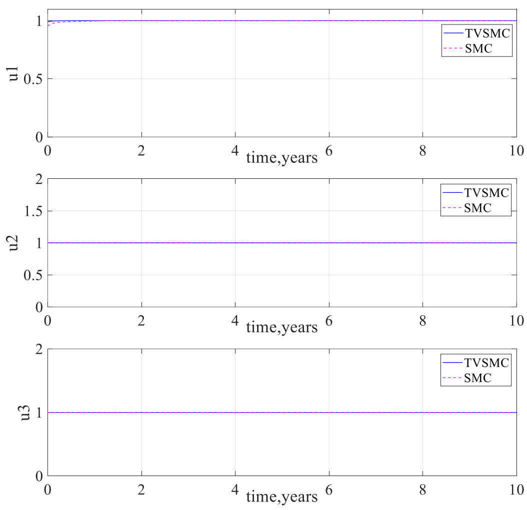


Fig. 2. Control measures including isolation (u_1), human intervention (u_2), and treatment for hospitalized individuals (u_3).

5. Conclusion

In this study, the time-varying sliding mode control was applied to set the HBV control policy with multiple measures. The HBV system can be stabilized by the synthesized control policy under the bounded disturbances. Thus, the epidemic of HBV was eradicated. The simulation results confirmed that the subpopulations according to the control objective were driven to the desired levels. According to the simulation results, the improvement of the convergence rate of the HBV control system could be achieved by using the time-varying sliding mode control policy. Therefore, it is appropriate to utilize the time-varying sliding mode controller design procedure for setting the HBV control policy.

References

1. J. Hu, "Hepatitis B Virus virology and replication," in Hepatitis B Virus in Human Diseases, Y. Liaw and F. Zoulim eds., Humana Press, series “Molecular and translational medicine,” Springer International Publishing AG, Switzerland, 2015.
2. World Health Organization, “Hepatitis B [Online],” 2021. Available: https://www.who.int/news-room/fact-sheets/
3. M. F. Yuen, D. S. Chen, G. Dusheiko, H. L.A. Janssen, D. T. Y. Lau, S. A. Locarnini, M.G. Peters, and C.-L. Lai, “Hepatitis B virus infection,” Nature Reviews Disease Primers, Vol. 4, No. 18035, 2018.
4. C. Trépo, H. L. Y. Chan, and A. Lok, “Hepatitis B virus infection,” Lancet, Vol. 384, pp. 2053-2063, 2014.
5. M. H. Nguyen, G. Wong, E. Gane, J. H. Kao, and G. Dusheiko, “Hepatitis B Virus: advances in prevention, diagnosis, and therapy,” Clinical Microbiology Reviews, Vol. 33, No. 2, p. e00046-19, 2020.
6. Centers for Disease Control and Prevention, “Viral Hepatitis [Online],” 2020. Available: https://www.cdc.gov/hepatitis/hbv/index.htm
7. C.H. Siederdissen, M. Cornberg, and M. P. Manns, “Clinical virology: diagnosis and virologic monitoring” in Hepatitis B virus in human diseases, Y. Liaw and F. Zoulim eds., Humana Press, series “Molecular and Translational Medicine,” Springer International Publishing AG, Switzerland, 2015.
8. T. Wilkins, R. Sams, and M. Carpenter, “Hepatitis B: screening, prevention, diagnosis, and treatment,” American Family Physician, Vol. 99, No. 5, pp. 314-323, 2019.
9. N. Coppola, L. Alessio, M. Pisaturo, M. Macera, C. Sagnelli, R. Zampino, and E. Sagnelli, “Hepatitis B virus infection in immigrant populations,” World journal of hepatology, Vol. 7, No. 30, pp. 2955–2961, 2015.
10. T. Khan, G. Zaman, and M. I. Chohan, “The transmission dynamic and optimal control of acute and chronic Hepatitis B,” Journal of Biological Dynamics, Vol. 11, No. 1, pp. 172-189, 2017.
11. J. Zhang and S. Zhang, “Application and optimal control for an HBV model with vaccination and treatment,” Discrete Dynamics in Nature and Society, Vol. 2018, p. 2076983, 2018.
12. C. O. Ijalana and T. T. Yusuf, “Optimal control strategy for Hepatitis B virus epidemic in areas of high endemicity,” International Journal of Scientific and Innovative Mathematical Research (IJSIMR), Vol. 5, pp. 28-39, 2017.
13. E. N. Wiah, O. D. Makinde, and I. A. Adetunde, “Optimal control of Hepatitis B virus disease in a population with infected immigrants,” Engineering Mathematics Letters, Vol. 2015, 2015.
14. F. S. Akinboro, T. O. Oluyo, O. O. Kehinde, and S. Alao, “Effect of control on the mathematical model of Hepatitis B Virus with infective migrant,” Archives of Current Research International, Vol. 20, No. 6, pp. 58-70, 2020.
15. D. Mahardika, R. H. Tjahjana, and Sunarsih, “Dynamic modelling of Hepatitis B and use of optimal control to reduce the infected population and minimizing the cost of vaccination and treatment,” AIP Conference Proceedings, Vol. 2331, No. 1, p. 020013, 2021.
16. M. A. Khan, S. A. A. Shah, S. Ullah, K. O. Okosun, and M. Farooq, “Optimal control analysis of the effect of treatment, isolation and vaccination on Hepatitis B virus,” Journal of Biological Systems, Vol. 28, No. 2, pp. 351-376, 2020.
17. S. Ullah, M. A. Khan, and J. F. Gómez-Aguilar, “Mathematical formulation of hepatitis B virus with optimal control analysis,” Optimal Control Applications and Methods, Vol. 40, No. 3, pp. 529-544, 2019.
18. A. Ibeas, M. de la Sen and S. Alonso-Quesada, "Sliding mode robust control of SEIR epidemic models," 2013 21st Iranian Conference on Electrical Engineering (ICEE), pp. 1-6, 2013.
19. R. Nistal, M. de la Sen, S. Alonso-Quesada and A. Ibeas, "A supervised multi-control for monitoring the antiviral treatment strategy for an SEIADR epidemic model," 2018 5th International Conference on Control, Decision and Information Technologies (CoDIT), pp. 524-530, 2018.
20. M. Sharifi, and H. Moradi, “Nonlinear robust adaptive sliding mode control of influenza epidemic in the presence of uncertainty,” Journal of Process Control, Vol. 56, pp. 48-57, 2017.
21. A. Boonyaprapasorn, T. Choopojcharoen, P. Sa-Ngiamsunthorn, and K. Thung-od, “Synergetic controller for Hepatitis B epidemic system,” Proceedings of the 3rd International Conference on Robotics, Control and Automation, Chengdu, China, pp. 31–36, 2018.
22. M.D. Sen, A. Ibeas, and S. Alonso-Quesada, “Feedback linearization-based vaccination control strategies for true-mass action type SEIR epidemic models,” Nonlinear Analysis-Modelling and Control, Vol. 16, pp. 283-314, 2011.

23. M. De la Sen, A. Ibeas, and S. Alonso-Quesada, “On vaccination controls for the SEIR epidemic model,” Communications in Nonlinear Science and Numerical Simulation, Vol. 17, No. 6, pp. 2637-2658, 2012.
24. H. Jiao and Q. Shen, "Dynamics analysis and vaccination-based sliding mode control of a more generalized SEIR epidemic model," IEEE Access, Vol. 8, pp. 174507-174515, 2020.
25. A. Boonyaprapasorn, T. Choopojcharoen, E. Pengwang, T. Maneewarn, S. Natsupakpong, P. Sa-Ngiamsunthorn, A. Suechoei, and W. Wechsathol, "Control of ebola epidemic system based on terminal synergetic controller design," 2019 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS), pp. 204-209.
26. A. Boonyaprapasorn, T. Choopojcharoen, E.Pengwang, S. Natsupakpong, T. Maneewarn, P. Sa-Ngiamsunthorn, and K. Thung-od, “The design of a synergetic controller for tuberculosis epidemic system,” International Journal Of Mechanical Engineering & Robotics Research, Vol. 10, No. 9, 2020.
27. A. Ibeas, M. De Sen, and S. Alonso-quesada, “Robust sliding control of SEIR epidemic models,” Mathematical Problems in Engineering, Vol. 2014, pp. 1–26, 2014.
28. D. M. L. Santos, V. H. P. Rodrigues, and T. R. Oliveira, “Epidemiological control of COVID-19 through the theory of variable structure and sliding mode systems,” Journal of Control, Automation and Electrical Systems, 2021.
29. S. Nuñez, F. A. Inthamoussou, F. Valenciaga, H. De Battista, and F. Garelli, “Potentials of constrained sliding mode control as an intervention guide to manage COVID19 spread,” Biomedical Signal Processing and Control, Vol. 67, p. 102557, 2021.
30. A. H. A. Mehra, I. Zamani, Z. Abbasi, and A. Ibeas, “Observer-based adaptive PI sliding mode control of developed uncertain SEIAR influenza epidemic model considering dynamic population,” Journal of Theoretical Biology, Vol. 482, p. 109984, 2019.
31. H. Puebla, P. K. Roy, A. Velasco-Perez, and M. M. Gonzalez-Brambila, “Biological pest control using a model-based robust feedback,” IET systems biology, Vol. 12, No. 6, pp. 233–240, 2018.
32. V. Utkin, J. Guldner, and J. Shi, Sliding mode control in electro-mechanical systems, 2nd Ed., CRC Press, Boca Raton, 2009.
33. H. Khalil, Nonlinear systems, 3rd Ed, Pearson Education, London, UK, 2002.
34. J. Liu, and X. Wang, "Advanced sliding mode control for mechanical systems," in Advanced sliding mode control for mechanical systems: Design, analysis and matlab simulation, Springer-Verlag Berlin Heidelberg, Germany, pp. 81-96, 2011.
35. H. Puebla, A. Morales-Diaz, and A. Velasco Pérez, “Sliding mode control for biological pest control problems,” Congreso Nacional de Control Automático, AMCA 2015, Cuernavaca, Morelos, México, 2015.
36. H. Puebla, H. F. Mejia, E. Hernandez-Martinez, and M. Ortiz-Vargas, "Sliding mode control of prey-predator interactions," 2008 Chinese Control and Decision Conference, 2008, pp. 4230-4233.
37. X. Zheng, X. Jian, D. Wenzheng, L. Liang, and G. Yang, “Time-varying integral adaptive sliding mode control for the large erecting system,” Mathematical Problems in Engineering, Vol. 2014, 2014.
38. S. Chen and H. Huang, "Time-varying sliding mode surface control for high order affine non-linear systems," 2017 32nd Youth Academic Annual Conference of Chinese Association of Automation (YAC), 2017, pp. 559-565.
39. T. Wang, N. Tan, X. Zhang, G. Li, S. Su, J. Zhou, J. Qiu, Z. Wu, Y. Zhai, R. D. Labati, V. Piuri, and F. Scotti, "A time-varying sliding mode control method for distributed-mass double pendulum bridge crane with variable parameters," IEEE Access, Vol. 9, pp. 75981-75992, 2021.
40. C. Zhang, L. Zhang, B. Peng, and H. Zhao, “PMLSM nonlinear integral sliding mode control based on gain time-varying reaching law,” Recent Advances in Electrical & Electronic Engineering (Formerly Recent Patents on Electrical & Electronic Engineering), Vol. 13, No. 5, pp. 743-750, 2020.
41. M. Maaruf, S. E. Ferik and M. S. Mahmoud, "Integral sliding mode control with power exponential reaching law for DFIG," 2020 17th International Multi-Conference on Systems, Signals & Devices (SSD), 2020, pp. 1122-1127.
42. S.-B. Choi, D.-W. Park, and S. Jayasuriya, “A time-varying sliding surface for fast and robust tracking control of second-order uncertain systems,” Automatica, Vol. 30, No. 5, pp. 899-904, 1994.
43. M. U. Salamci and G. S. Tombul, "Sliding mode control design with time varying sliding surfaces for a class of nonlinear systems," 2006 IEEE Conference on Computer Aided Control System Design, 2006 IEEE International Conference on Control Applications, 2006 IEEE International Symposium on Intelligent Control, 2006, pp. 996-1001.
44. A. Bartoszewicz, and A. Nowacka-Leverton, "Time-varying sliding modes for the second order systems," in Time-varying sliding modes for second and third order systems, A. Bartoszewicz and A. Nowacka-Leverton, eds., pp. 17-65, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.
45. W. Gao and J. C. Hung, "Variable structure control of nonlinear systems: a new approach," IEEE Transactions on Industrial Electronics, Vol. 40, No. 1, pp. 45-55, 1993.

Authors Introduction

Arsit Boonyaprapasorn

He received a B.S. in Mechanical Engineering from King Mongkut's University of Technology (KMUTT), Thailand, in 1998. He continued his postgraduate studies at Case Western Reserve University, Cleveland, Ohio, U.S.A., obtaining an M.S. in Systems and Control Engineering in 2003 and a Ph.D. in Mechanical Engineering in 2009. He is currently an assistant professor at the department of mechanical engineering, academic division, Chulachomklao Royal Military Academy, Nakhon Nayok, Thailand. His research interests revolve around nonlinear control and robotics, with specific foci on nonlinear feedback control and its application in biological systems, lab-on-chips, and climbing robots.

Suwat Kuntanapreeda

He received the B.Eng. degree in mechanical engineering from King Mongkut's University of Technology North Bangkok (KMUTNB), Thailand, in 1990, and the M.S. and Ph.D. degrees in mechanical engineering from Utah State University, USA, in 1992 and 1995, respectively. He is currently Professor in mechanical and aerospace engineering at KMUTNB. His research interests include the application of control theory to mechatronic, chaotic, and aerospace systems.

Parinya Sa Ngaimsunthorn

He is an assistant professor at the Department of Mathematics, Faculty of Science, King Mongkut's University of Technology Thonburi, Thailand. He obtained B.Sc. and Ph.D. degrees in Mathematics from the University of Sydney, Australia. His research interests are Nonlinear analysis, Partial differential equations, and Dynamical systems.

Thunyaseth Sethaput

He received his B.S. degree in mechanical engineering from Sirindhorn International Institute of Technology, Thammasat University, Thailand and Ph.D. in systems and control engineering, from Case Western Reserve University, Ohio, USA. He is an assistant professor and chairperson of mechanical engineering program at Sirindhorn International Institute of Technology, Thailand. His research interests are control system, systems biology, biomedical mechanics, simulation modeling, finite element analysis, and computational fluid dynamics.

Tinnakorn Kumsaen

He received his B.S degree in chemical engineering from Khon Kaen University, Thailand, M.S. degree in system engineering Case Western Reserve University (CWRU), Ph.D. in chemical engineering CWRU. A faculty member at chemical engineering Khon Kaen University, Research of interests are control system, process design simulations & integrations, renewable energy.

Blockchain Technology for Halal Supply Chain Management*

Kadeer Zulihuma†
Faculty of Business & Management, UCSI University, Cheras, Kuala Lumpur, Malaysia‡

Abdul Samad Bin Shibghatullah
Institute of Computer Science & Digital Innovation, UCSI University, Cheras, Kuala Lumpur, Malaysia§

E-mail: 1001333819@ucsiuniversity.edu.my, abdulsamad@ucsiuniversity.edu.my

Abstract

Blockchain technology (BT) is a distributed and decentralized database that store transaction information in a network. Due to providing better visibility and transparency, this technology has gained a considerable attention in the recent years. This research is carried out with the purpose of exploring the potential of blockchain technology to increase supply chain integrity in halal food industry. Therefore, a literature on BT and its adoption in the halal supply chain is given and a model is developed to identify the influence of BT on halal supply chain. Three features of smart contract including traceability, decentralized, and anonymity are added to the model as moderators to explore their influence on integrity of halal supply chain..

Keywords: Blockchain Technology; Halal Supply Chain; Supply Chain Integrity.

1. Introduction

Nowadays, due to the global safety crisis, the concerns of consumers about food safety, quality, origin, and authenticity is increased (Zailani et al, 2018), and lack of uniformity has caused issues in halal supply chain (Talib, Hamid, & Chin, 2016). The traditional supply chain is complex, and all of the information is outbreak and cannot be observed clearly. It is because everything is recorded on the paper, so there is a need to provide end to end visibility to deliver information quickly (Zailani et al, 2018). Another concern is ingredients and cleanliness of the food, and possibility of cross-contamination of halal food with any non-halal ingredients in the process of storage and distribution, which does not conform to the Islamic principle (Zailani et al, 2018). Currently, trust in the halal supply chain is solely according to presence of the halal logo or halal certification on the packages (Ali et al, 2017) as it allows the consumers to acknowledge that products are made based on the halal guideline and principles of Sharia law. Despite the existence of halal certificate, consumers are questioning the legality of the displayed certificates and halal logo (Azmi, et al, 2018). It is due to the fact that there have been several reports of displaying fake certificates or halal logo in Malaysia, which negatively impacted the reputation of Halal industry, and it cast doubt and tarnished trust among halal consumers. Hence, it is vital to improve trust between all actors of halal food network in order to enhance the integrity status of the halal food supply chain (Zulfakar et al, 2014; Zailani,et al, 2018). Putting an effort to build a comprehensive and strict halal quality assurance system or a using a dedicated transportation to deliver the halal food products, can go on waste if other actors that directly deal with them in the supply chain do not trust the firm in providing halal products (Zulfakar et al, 2014).

Moreover, the inability to trace the supply chain is common for industries such as agri-food and creates challenges in the management of the Halal supply chain. Developing a traceability system, which is trusted and reliable, in the Halal food supply chain will potentially enhance the transparency and therefore improves the integrity of halal supply chain (Zulfakar et al, 2014). Some studies offered existing technology such as Radio Frequency Identification Device (RFID) (Accorsi et al, 2016; Dabbene et al, 2016; Farooq et al, 2016) to develop and improve halal traceability, however very few studies addressed blockchain–SCM integration, and there is a need to further study adopting blockchain technology in supply chain (Azmi et al, 2018; Tsang et al, 2019; Queiroz et al, 2019).

With the purpose of addressing this research gap, this research attempts to elaborate on the integrity of the halal supply chain using blockchain technology, to understand the integrity of the halal supply chain and related factors, as well as the importance of the blockchain on integrity of halal supply chain.

2. Previous Work

Halal supply chain consists of a network of supply in which the products are given special attention to assure halal integrity. Considering the increasing demand for halal products throughout the world, both from Muslim and non-Muslim individuals, it is crucial to assure the quality of the product through an integrated supply chain (Rasi, et al, 2017). Integrity means honesty and demonstrates constant observation and strong ethical principles and values. It is essential to maintain the integrity of the halal supply chain, all aspects of halal food integrity must be cautious about the protection, and each actor involved in the supply chain must perform all necessary steps to escape any cross-contamination that cause product to become non-halal (Zulfakar, Anuar, & Talib, 2014). Moreover, food safety is part of halal integrity, and process of traceability that grant track relevant information from product source to point of sale. In this way, the entity can track and trace the food, such as ingredients of the food and expiry date, manufacturing date, packaging at all stages of the whole process of production, and record the process along the way, including when, where, and how (Kadir et al., 2016). To achieve an improved traceability, it is required to transfer the conventional supply chain and make it digitalized, which can be done by using blockchain technology in the halal supply chain management (SCM) system. It is recommended to use blockchain in halal supply chain due to following reasons: Firstly, using blockchain technology helps to solve traceability issues, because in blockchain technology all information is digitalized and every entity can have instant access to records about the food and products. Using smart contract between buyers and suppliers improve trust since data is immutable, as well as speed since third parties are eliminated. Secondly, blockchain technology avoid delays and therefore lengthen shelf live (Ibm, 2019). To elaborate blockchain technology on supply chain that related to halal, food integrity, supply chain integration, Islamic practices and firms' performance are reviewed. Literature of the relevant theories and discussion of supply chain management and halal studies form the foundation of conceptual framework.

In Islam, Muslims are banned from eating non-halal foods in daily living. The foods that are halal, namely as halal food, must be consumed by Muslims. To supply chain food industry, it has real example, food global safety alert. In 2018 December, in the US, E. coli infections linked to romaine lettuce, so United states was alert to people against to eat romaine lettuce because it has symptoms of Shiga toxin-producing, so everyone is infected with E. coli (STEC), but diarrhea and vomiting often happen, which can cause fever on some people. Most people got well in 5 to 7 days. Some infections are mild, while others are severe and even life-threating (Cdc.gov.2019). This issue causes millions of people who have eaten the popular lettuce, but investigators do not know precisely where, when or how the contamination happened. This is not only once a time that happened. So many real examples still exist. Therefore, in this example talking about the problems of supply chain flow, which is in the traditional way of processes, has decreased the efficiency of the market. If the concept of blockchain is applied to this processes, the technology can be traced through the food supply chain, if this technology applied to romaine lettuce from the beginning, such as source details, lot numbers, manufacturer, and processing data, expiration dates, storage temperatures, and shipping and shipping details from source to destination are digitally linked to actual food (Rooyen, 2017). The records during each transaction are verified by companies' in-network (farms, transport companies, packaging companies, warehouses, and stores) and a consensus is formed (Rooyen, 2017). After verifying each block (record of each transaction), it is added to a chain of transactions, which is immutable throughout the process. Ultimately, all-veggie items received in the store is validate and truthful, and this digital record can reveal food safety issues between farm and retailers (Rooyen, 2017). As well as, in individual stores that digital capture can help to market stores to well manage shelf-life products (Rooyen, 2017; Zailani et al, 2018). As refer to the Supply chain and trade finance, the Blockchain request to make all stakeholders in the supply chain to send, receive and track an electronic bill of lading, with upload and send related trade documents such as orders, invoices, and transporting orders (Zailani et al, 2018). The claim is linked to a delivered systematization network that can safety record and prove the ownership and authority of documents. Therefore, in the process of sending and sharing of verification information, the verified gross mass (VGM) information of freight containers is shared using blockchain technology. Knowledge of the VGM of the container is important to ensure the proper storage of the vessel to resist maritime and port accidents. VGM data is stored on a public blockchain, which is granted perpetual record insights from port officials, shipping firms, consigners and hosts of cargo. Also, this record will replace tedious logs, spreadsheets, third party's data and private databases. (Rooyen, 2017). In supply chain management (SCM), international clients may need to authenticate the original documents through the notary function of the blockchain and optimize the best creation, justification, and protection. When changing of ownership during a shipment called merchant, retrospective to the source of all elements in the checklist material. Monitor and automatic control of the use of third-party logistic services, transportation, and freight forwarders through irrevocable and unchanged smart contract execution. When the returned products are repaired or renovated, process management will be performed to ensure that repairs follow strict protocols, and refurbished, the products again refer the formats and standards that can be efficiently resolved by a Blockchain's notarization. Those products that are moved during the return logistics process may also be subject to warranty terms that require changing ownership. Also, failed items can be traced back to the source of defects identified to perform a root-cause analysis (Panigrahi et al, 2018).

3. The Proposed Conceptual Framework

Based on the research results by Queiroz et al (2019), very few studies discussed about the influence of blockchain on supply chain management over the past decade. Therefore, it indicates the lack of researches focusing on blockchain and integrity in supply chain (Queiroz et al, 2019). Basically, Halal Supply Chain consists of four main activities which are explained in the following.

3.1 Halal procurement on using Block chain

To use Blockchain technology in the halal supply chain, halal procurement is going to record information about the item or products, certification of origin, and details of suppliers, resources and materials, such as, halal species, and any genetic crows' contaminations from Haram species by conventional breeding. Therefore, blockchain provides complete traceability and monitoring system, which are deal with between supplier and buyer under smart contract. Besides, halal suppliers should monitor the agricultural production system, such as appropriate use of land, water, and chemicals, and provide the halal food and fodder (Rasi et al, 2017).

3.2 Halal manufacturing on using Blockchain

Halal manufacturing is an entity that is also called the halal producer. In the halal manufacturing, the raw material and also added ingredients must be certified halal. Besides, it includes a well-organized packaging process, and transforming process (Mohammed et al, 2016). To ensure all ingredients are halal, and to improve integrity, the companies in supply chain use blockchain technology to share information of raw materials and organization of the products for the benefits of the customers. Therefore, during this process, Blockchain will record this information, and share it with all of the entities from the supply chain (Rasi et al, 2017).

3.3 Halal distribution on using Blockchain

Halal distribution is an entity that provides halal packaging for the halal products or goods from the halal manufacturer which are finished products. Using Blockchain will help this entity confirm the safety of the product, expiry date, manufacturing date and ingredients or components of the Halal products. The distributor department deliver the products to the retailer that sells the Halal products or goods to the end-user customers (Rasi et al, 2017).

3.4 Halal logistic on using Blockchain

Halal logistic is considered as functions of logistics such as, transportation and packaging of products, warehouses to keep the halal products in storage, and documents of the bills, orders of financial transactions. Halal transportation has also benefited halal logistics activities (Tan et al., 2012). Logistic service provides information and communication technology, which is necessary to maintain the integrity of halal, and to enhancing the performance and efficiency of logistics and supply chain networks. Besides, the application of information technology components for halal transportation includes location tracking of the products or goods identification and data communication. (Ab-Talib, & Abdul-Hamid, 2014). Moreover, information technology controls logistic operation transparent and activities, which will improve the service and product level. Therefore, by using of blockchain technology all the information about transactions between supply chain entities will be recorded, which will also increase the efficiency of system.

Thiruchelvam et al (2018) studied the adoption of Blockchain Technology in the Coffee Supply Chain Trade. In this study the researchers investigated the effect of market access, premium pricing, traceability, transparency, and sustainability on blockhain efficiency and perceived usefulness. The results indicated that blockchain technology in the coffee supply chain helps coffee producers/farmers to gain better market access, sustainability, and traceability and therefore increase the fair trade and transparency in the supply chain. Francisco and Swanson (2018) proposed a framework for adopting technology of Blockchain for Supply Chain Transparency. The researcher concluded that using blockchain in supply chain contribute to companies' competitive advantages. Kamble et al (2018) revealed that supply chain practitioners believe that blockchain technology is beneficial for improving the supply chain effectiveness. Therefore, this study offers the following hypotheses:

Hypothesis 1. Adopting blockchain in halal supply chain positively impacts the integrity of the halal supply chain in the food industry.

3.5 Traceability

Traceability is defined as the ability to track origin and history of a product (Sutawijaya & Awangsari, 2019). In Halal food industry, with the use of traceability the practitioners can trace the halal status of a product throughout blockchain network. In this way, all the information activities that the Halal food products have went through such as origin of the product, the time and date of transferring product, etc is stored. The primary characteristics of traceability systems are (Banerjee et al, 2015): (a) register the origin of units of ingredients, (b) storing detailed information about time and location units are transformed, and (c) a comprehensive system that shares and transfers relevant traceability information with the product to the subsequent stage.

The feature of traceability helps to monitor the halal control points to retrieve detailed information, and in case of suspension of any cross contamination with non-halal ingredients, appropriate measures should be taken to control and avoid any unwanted events (Zulfakar, Anuar, & Talib, 2014). According to the study by Rohmah el al (2019), cooperation between supply chain sectors is vital to achieve ethical halal traceability. Consumers must be able to access information about product halalness, food quality and food safety. The findings by Thiruchelvam et al (2018) points out the importance of traceability and indicated that it helps to increase transparency. Therefore, this study offers the following hypotheses:

Hypothesis 2: Traceability positively moderates the relationship between blockchain-based halal supply chain and integrity of halal supply chain

3.6 Decentralize

Yli-Huumo et al (2016) stated "the goal of Blockchain technology is to create a decentralized environment where no third party is in control of the transactions and data". Eliminating a central control increase speed, and remove delays as many-to-one traffic flows are avoided (Dorri et al., 2016). Similarly, the study by Khyzer et al (2018) revealed that decentralization has a positive mediating effect on supply chain practices. Therefore, this study offers the following hypotheses:
Hypothesis 3: Decentralization positively moderates the relationship between blockchain-based halal supply chain and integrity of halal supply chain

3.7 Anonymous

The blockchain network is characterized by the anonymity aspect. Anonymity in this sense implies that the respective identities of the users in the network are anonymous, and the transaction on the ledger are treated as eliminating from an anonymous entity. The aspect of anonymity is often propagated by the existence of no-trusting parties' interaction in a decentralized environment. Depending on the nature of the contract anonymity may be necessary while in others, the anonymity aspect may not be essential (Nzuva, 2019). Therefore, this study offers the following hypotheses:
Hypothesis 4: Anonymity positively moderates the relationship between blockchain-based halal supply chain and integrity of halal supply chain

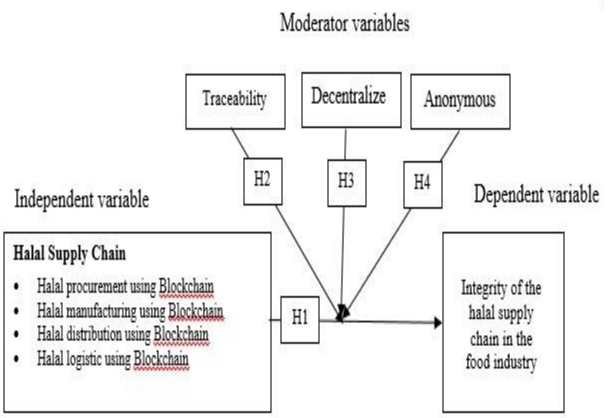


Fig.1: The Proposed Conceptual Framework

This proposed conceptual framework in fig.1 is the results of the comprehensive literature review, which includes blockchain in halal supply chain as independent variable (IV) (halal procurement, halal manufacturing, halal distribution, and halal logistic), and moderator variables of traceability, decentralization and anonymity as three moderators with the purpose of investigating the effect of variables on improving the integrity of the halal supply chain (DV).

4. Conclusion

This paper explained about blockchain, halal supply chain, and smart contract through a comprehensive literature review. The importance of blockchain in halal supply chain is discussed and the gap in the literature review is addressed. Considering the literature review, three feature of smart contract including traceability, decentralization and anonymity are studied to form the conceptual framework.

Acknowledgements

The authors wish to thank ICSDI, and CERVIE UCSI University for supporting this project.

References

1. Accorsi, R., Ferrari, E., Gamberi, M., Manzini, R., & Regattieri, A. (2016). A closed-loop traceability system to improve logistics decisions in food supply chains: A case study on dairy products. In Advances in Food Traceability Techniques and Technologies (pp. 337-351). Woodhead Publishing
2. Ali, M. H., Zhan, Y., Alam, S. S., Tse, Y. K., & Tan, K. H. (2017). Food supply chain integrity: the need to go beyond certification. Industrial Management & Data Systems.
3. Attuquayefio, S., & Addo, H. (2014). Using the UTAUT model to analyze students' ICT adoption. International Journal of Education and Development using ICT, 10(3).
4. Azmi, F., Abdullah, A., Bakri, M., Musa, H., & Jayakrishnan, M. (2018). The adoption of halal food supply chain towards the performance of food manufacturing in Malaysia. Management Science Letters, 8(7), 755-766.
5. Dabbene, F., Gay, P., & Tortia, C. (2016). Radio-frequency identification usage in food traceability. In Advances in Food Traceability Techniques and Technologies (pp. 67-89). Woodhead Publishing.
6. Farooq, U., Tao, W., Alfian, G., Kang, Y. S., & Rhee, J. (2016). ePedigree traceability system for the agricultural food supply chain to ensure consumer health. Sustainability, 8(9), 839.
7. Queiroz, M., Telles, R. and Bonilla, S. (2019). Blockchain and supply chain management integration: a systematic

review of the literature. Supply Chain Management: An International Journal.

8. Rooyen, J. V (2017). Blockchains for supply chains - part II - Resolve Solution Partners. Resolve Solution Partners.
9. Talib, H. H. A., & Mohd Ali, K. A. (2009). An overview of Malaysian food industry: The opportunity and quality aspects. Pakistan Journal of Nutrition, 8(5), 507-517.
10. Talib, M. S., Hamid, A. B. A., & Chin, T. A. (2016). Can halal certification influence logistics performance?. Journal of Islamic Marketing.
11. Tsang, Y. P., Choy, K. L., Wu, C. H., Ho, G. T. S., & Lam, H. Y. (2019). Blockchain-driven IoT for food traceability with an integrated consensus mechanism.
12. Zailani, S., Jafarzadeh, S., Iranmanesh, M., Nikbin, D., & Selim, N. I. I. (2018). Halal logistics service quality: conceptual model and empirical evidence. British Food Journal.
13. Zulfakar, M. H., Anuar, M. M., & Talib, M. S. A. (2014). Conceptual framework on halal food supply chain integrity enhancement. Procedia-Social and Behavioral Sciences, 121, 58-67.He, Y., He, Z., & Zhang, D. (2009). A Study on Prediction of Customer Churn in Fixed Communication Network Based on Data Mining. 2009 Sixth International Conference on Fuzzy Systems and Knowledge Discovery.

Authors Introduction

Ms Kadeer Zulihuma

She received her Bachelor Degree in Interior Architecture from UCSI University in 2018 and continue his Master in Business Administration at the same university and completed in 2020. Her research interest is in Halal blockchain technology.

Dr. Abdul Samad Bin Shibghatullah

He received the bachelor accounting degree from Universiti Kebangsaan Malaysia, Bangi, Malaysia in 1999, the M.Sc. degree in computer science from the Universiti Teknologi Malaysia, Skudai, Malaysia in 2002, and the Ph.D. degree in computer science from the Brunel University of Uxbridge, United Kingdom. He is currently Associate Professor at Institute of Computer Science & Digital Innovation, UCSI University, Kuala Lumpur, Malaysia. His current research interests include optimization, modelling and scheduling.

Smart Tourism Guide Application Using Location-Based Services-Go.Travel

Wong Yit Meng
Institute of Computer Science & Digital Innovation, UCSI University,
UCSI Heights, 1 Jalan Puncak Menara Gading, Kuala Lumpur, WP Kuala Lumpur 56000, Malaysia

Abdul Samad bin Shibghatullah
Institute of Computer Science & Digital Innovation, UCSI University,
UCSI Heights, 1 Jalan Puncak Menara Gading, Kuala Lumpur, WP Kuala Lumpur 56000, Malaysia

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: 1001850602@ucsiuniversity.edu.my, abdulsamad@ucsiuniversity.edu.my, kasthurisuba@ucsiuniversity.edu.my
www.ucsiuniversity.edu.my
,

Abstract

This research paper presents the creation of an assistive application that allows travellers to use tourism-related services in Malaysia to improve their travel experience. The objective of this project is to develop and enhance a Smart Tourism Guide Application, Go.Travel based on the limitations discovered. Tourists will be able to travel more efficiently and have a better travel experience in Malaysia with the use of the proposed application.

Keywords: Smart Tourism Guide Application, Smart tourism technology, tourism-related services, tourism industry.

1. Introduction

1.1. *Background*

According to World Tourism Organization, tourism is "a cultural, economic and social appearance involving any mobility of persons for business or own reasons to nations or locations beyond their general circumstance." A significant aspect of a smart city is smart tourism. In societies globally, tourism is among the key elements of economic development. Attracting new visitors from various regions of the globe has become a crucial demand of tourism. Smart tourism can be defined as the use of ICT to implement innovative tools and techniques for enhancing tourism, comparable to smart cities. [1] Technological capabilities including mobile networking, Information and Communications Technology, VR technology, cloud services and artificial intelligence depend on smart tourism. It integrates tourism's natural, information, public, and economic resources with these tools to offer intelligent incentives for tourism. [2] The goal of smart tourism is to enhance the effectiveness of resources development, optimize productivity and, by utilization of innovative techniques and processes, to achieve sustainability. The concepts of smart tourism are focused on improving the experience of tourism, improving the efficacy of resources development, optimizing the attractiveness of destinations, with such a focus on continuous aspects. Through utilizing and analysing information gathered across infrastructural facilities and social relationships, Smart Tourism has concentrated on using emerging technology to develop information into productive innovative business models. [3] Information should also be collected and delivered to promote the effective distribution of tourism resources as well as to combine macro- and micro level tourism resources to ensure that the advantages are delivered equally. In locations with advance technologies, including smart cities, they are found to be efficient and effective. The digital Internet offers immense opportunities for travellers to browse for fascinating knowledge and to schedule their trips. Current advancements in communication and information technology enable travellers during their travels to obtain useful information on the Internet. Around 50% majority of the current tourism recommendation applications are implemented for mobile platforms. Tourism has become one of the most appropriate industries for mobile technology and mobile applications, as seems to be the trend for other communication and information technologies. There are several applications and services that enable user to optimize this search, to provide information efficiently and effectively on fascinating attractions and reviews from users.

The advantage of smart tourism not only rises in tax revenue of the country, but it also brings a lot of advantages to individuals. [4] The most common challenge that tourists faced is confused about what to do and where to go while they travel to a city. It is more dangerous if tourists drive to the cities because they are unfamiliar with the environment, it can cause panic on tourists. These factors can be taken into consideration by smart tourism to figure out the best way to guide tourists during their tour. Using smart tourism technology, tourists can understand and enjoy the destination they are visiting. Besides, the travel guidance helps to guide tourists through all the fascinating destination in that area. Smart tourism technology allow tourist to browse for the best accommodation, transportation, restaurant, and destination to plan and arrange their fulfilling trip. Smart tourism technology provides self-guided tours that is more efficient and effective compared to the traditional multimedia tours guide. In 2018, the huge technology company, Microsoft has announced their collaboration with Guide Dogs to create the app called Soundscape. They aimed to put people with sight loss at the heart of the development of the new Soundscape app, enabling them to influence its design and test the product. Soundscape application allow individuals with sight loss to travel around the cities more efficient and easily. The application enhances individuals' awareness of the world as they walk, assisting them to go anywhere they want to go [5]. Hang out with friends, rush to the classes and travel to a new city are all typical activities that can be more satisfying for almost 300 million people with sight loss globally and cause less anxiety. Apart from offering audio guide of a destination, smart tourism application also can display the architecture, buildings and environment that appeared in the present, therefore tourists can understand and know how the buildings and the environment is developed over the years. [6] Furthermore, a simple design for providing feedback can be offered by individually designed applications. Tourists can provide feedback after visiting a destination or restaurant for other tourists to review for the similar destination or restaurant. It is useful because it allows the tourists to understand and know about the quality and services of certain destination, accommodation, or restaurant. Self-guided tours enable tourists to travel by following their own pace, and significant benefits also provided for tourists outside the technology sector. In addition, smart tourism application can trace where tourists visited, allowing tourists to arrange their own journey. [7] Smart tourism technology also allow tourists to see the map of the destination to search for the nearby accommodation, restaurant, and destination. This help tourists to prevent wasting time for travelling to a far place. Smart tourism technology allows tourists to make friends around the world through the travel forum. Tourists can share their own experience on travelling to a fascinating destination in the travel forum. Other tourists who saw it can also give comment on their post. Lastly, smart tourism technology allows tourists to read online latest news through their mobile device. They can know what happened recently through the online news.

1.2. *Aim*

This research project is aimed to develop a user friendly, convenient, and efficient smart tourism guide application, Go.Travel that runs on mobile platform to help all the tourists in Malaysia to maximize their travel experience.

1.3. *Objectives*

i. To review on existing Smart Tourism Guide Application and compare it to discover the advantages and limitations.
ii. To enhance the proposed Smart Tourism Guide Application to improve user's experience based on the limitations discovered.

iii. To design and develop a Smart Tourism Guide Application with all the functional requirements and non-functional requirements that are stated.
iv. To evaluate and test the proposed Smart Tourism Guide Application with other users to make sure that all the features are working well.

2. Systematic Review

Reviews on existing travel guide mobile application are essential to understand the application's functionality and how well the design is. It is a vital part of having a new idea and avoiding the application's weakness.

TripAdvisor is an online travel company that allow users to access its application through web and mobile devices. It provides a lot of functionality, and the design of its application is modern. This application also has location-based services and map which allow users to browse for nearby accommodation, restaurants, and destination. Using TripAdvisor, users can search for Hotels, Things to Do, Restaurants, Holiday Rentals, and Flights, it will recommend the best item for users. There is travel notice in TropAdvisor for users to read the online articles. Besides, users can share their own travel experience on the travel forums. Users also can communicate with other travellers who has the same interest with themselves. Users also can create and save a trip that is interested to travel in the future. To create trip and communicate with other travellers, users are required to create a profile for themselves. The benefits of TripAdvisor are it has variety of functions that allow users to plan and organize their trip more efficient and effective. Besides, users can share their experience and other travellers can comment their own opinion or idea of it. Lastly, there are a lot of reviews about the destination, restaurants, and accommodations. The limitations of TripAdvisor are it contain a lot of bogus reviews which make users hard to evaluate for a destination, restaurant, or accommodation. Besides, the home page of it has many sections which cause it to become complicated, users need to scroll through it for a long time to see all the information.

PocketGuide is the first and world's leading smartphone app, focused on GPS, audio, touring and storytelling. PocketGuide is reviewed because it provides audio guide tour that can help blind people in the world. It also provided map for users to see nearby places. If users interested in certain city, they could read their history and information in the city info. Moreover, users can create or join a travel group to travel together with their friends. Lastly, there are tour tickets for users to search and purchase for it. The benefits of PocketGuide is users can experience clear audio travel guides for more than 100 major cities throughout the world. Besides, users can use their device's integrated GPS to see their current location on a map. Moreover, users can listen as your audio tour guide provides insights to the history and culture of cities. All the map and review content is stored locally on your device so there is no need for an Internet connection. The limitations of PocketGuide are user need to purchase the audio guide.

Foursquare is a city guide application which assist users to explore and search details about attractions and restaurants. Foursquare is reviewed because users can connect with their friends. Users can connect Foursquare to users' personal contact, Facebook and Twitter account, it assists users to discover friends in their contact, Facebook and Twitter who are also using Foursquare application. This application will update users what their friends are doing after they followed their friends in Foursquare. Besides, Foursquare provide the check in function which is very useful. After users visiting a place, they can press check in. Foursquare will update their status. Therefore, their friends will know where they go and when they reached. Benefits of Foursquare are it good recommendation function. Besides, users can leave their reviews after visiting a place. Furthermore, users can earn badges and mayorships. For example, if users visit certain place frequently or as an active user, they will be rewarded due to their loyalty. Lastly, users can see history of places that is visited before and can bookmark the place that users is interested and want to visit it next time. The only limitation of this application is users are required to create an account before accessing to this application.

Tourplus is a tourism application founded in Malaysia that provide local tourism guides to users. The reason to review Tourplus is because this application provides a day tour package to users. It allows users to explore hidden gems of a place. After users booked a day tour, they can connect with a driver-guide within 2 hours. The person will guide users personally. Besides, Tourplus also provide multi-day tour for users. There are many 3 days 1 night's package of beautiful cities and destination. The benefits of Tourplus are users can book flight tickets in this application, it will recommend users the most suitable flight tickets. Besides, users can book a day tour or multi-day tour package with an experience driver-guide which allow users to enjoy the best moment of their trip. There are also review leave by other users for the tour package. Users can look on it and compare to find the most appropriate tour guides their trips. The limitations of Tourplus are it only allow users to search for nearby restaurant and some stores. Users are unable to find some interesting attractions or destination to visit. This application is lack of recommendation that allow users to decide their choice.

Yelp is a famous online directory for exploring attractions, restaurants, and entertainment. Yelp is reviewed because users can connect the application with Facebook and address book. It will help to update their friend's status. Yelp is reasonably user-friendly. To access the feedback information and data of a company on Yelp, users do not need a username. It increases the likelihood that users will check up your profile without any sign-up trouble through a relatively easy procedure. The benefit of Yelp is it annually draws 150 million tourist arrivals each month. [8] It means a great deal of attention can be provided to your company by the platform. Besides, Yelp provide a lot of trustworthy reviews. It will filter the reviews and suggest all verified reviews that is trustable. Limitations of Yelp is Yelp normally requires companies to agree to a 1-year deal on their platform. While this is far more of an assumption than a guideline, it can be a complicated challenge to choose to eliminate a paid existence well before 12 months are over. Besides, for certain small enterprises, upgraded listings are expensive for them. The expense of getting an upgraded listing is reported to be at minimum $350, although authorized rates are not released. When they involved in upgraded listings and advertising, they are encouraged to call Yelp personally.

Klook is an online tourism booking site with a lot of main attractions worldwide. Users are allowed to customize and arrange their tour schedule. The reason to review Klook is because it is user friendly application, all the detailed information of each destination with ratings is in one area, users can access to it easily. It contains a lot of categories which let users to search for certain category of information. The benefits of Klook is it provide instant confirmation to users. [9] After users' book for tickets for their trips, they will receive their voucher once they completed their booking procedure. Besides, the booking procedure is very fast, all the transaction is fast and smooth. Moreover, Klook has a good customer service, there are few customers service team that are always prepared to assist users. The limitation of Klook is users are unable to search nearby destination, restaurants, or accommodation.

Table 1. Comparison table between Existing Tourism Guide Application

	TripAdvisor	PocketGuide	Foursquare	Tourplus	Yelp	Klook
Location-Based Services	✓	✓	✓	✓	✓	✓
Web application	✓	✓	✓	✓	✓	✓
Mobile application	✓	✓	✓	✓	✓	✓
User Profile	✓	✗	✓	✓	✓	✓
Nearby Suggestions	✓	✓	✓	✓	✓	✗
Map	✓	✓	✗	✗	✗	✗
Destination Recommendations	✓	✓	✓	✓	✓	✓
Accommodation Recommendations	✓	✗	✗	✗	✓	✓
Restaurant Recommendations	✓	✓	✓	✓	✓	✓
Flights Recommendations	✓	✗	✗	✓	✓	✓
Transportation Recommendations	✗	✓	✗	✓	✗	✓
Travel Forum	✓	✗	✗	✗	✗	✗
Reviews	✓	✗	✓	✓	✓	✓
News and Notice	✓	✗	✗	✗	✗	✗
City Info	✗	✓	✗	✓	✗	✗
Save places	✓	✗ (Need download the tour)	✓	✗	✓	✗

According to table 1 above, the major limitations of existing smart tourism guide application is the travel forum and travel news features. There is only TripAdvisor application that allow users access to the travel forum to share their travel experience. However, TripAdvisor is an application not mainly for Malaysia, there are different other countries category of travel forum. Therefore, less posts is shared in travel forum of Malaysia. Besides, it is confused and complex for users when they wish to read some stories and travel experience shared by other tourists in Malaysia. This limitation will be improved in the proposed smart tourism guide application. The travel forum will be provided in the proposed application. Besides, the design and interface is easy to navigate and well categorized so that users can share their travel stories and experience or read other tourists travel stories with more satisfaction and efficiency. Another limitation which is travel news of existing application. There is only TripAdvisor provide travel news and notice features to make announcement of some important notice or some travel guidance information. However, the news and information provided by TripAdvisor is not sufficient for tourists and the news are seldom updated. In the proposed application, this limitation is improved. Different type of news and information in Malaysia will be provided sufficiently. Besides, the latest news and information will always be provided for users to get updated of the latest things happening around. Moreover, Tourplus is a smart tourism guide application created in Malaysia. However, there are some limitations in this application. One of the limitations is there are no map provided in this application for users to look on it and check for nearby places, map allow users to check a wide range of places more accurately and it is scalable. Therefore, this limitation is solved in proposed application. A map will be provided for users to press on it and check for nearby places. Users are free to zoom in and zoom out for the map to see clearer about the places around the area. In addition, Tourplus does not have the feature of save places, it is important for users to save the places when they are interested on it. Users are required to search again the places to see the information about these places after they quit the application. This limitation is addressed in the proposed application. In the proposed application, users are allowed to save the places if they are interested. Therefore, users can check on the save collection next time if they want to see the information of places that is saved previously.

3. Methodology

3.1. *Research Method*

Online survey was chosen as the research method used to collect data for this project. Conducting online surveys as quantitative data collection will collect broad information. The online survey will be done by questioning people through Google form. The google form survey questionnaire contains a set of total 18 structured questions and will be generated and sent to 120 respondents. The target participants are focusing on university students. The respondents can complete the google form survey questionnaire over the Internet through filling out the form. The duration for collecting data from the 120 respondents is 7 days. After 7 days, all the responses will be gathered, and the data collected will be analysed.

3.2. *System Methodology*

Iterative development is chosen as the system methodology for this project as this project is developing a Smart Tourism Guide Application, Go.Travel, which will be slightly complex because it included many features in the application. The reason for choosing iterative development as the system methodology for this project is because it is much more appropriate for this project and reliable compared to other system development methods. A higher quality of Smart Tourism Guide application can be developed and proposed using iterative development method as it spends more time in the design and prototyping phases to ensure that the application is completely tested to meet users' expectations and satisfaction. The proposed application requires more time in the design phase to be enhanced and improve all the weaknesses. Using iterative development method will have more time in the design phase to solve all the defects and spend less time on documentation.

4. Results and Findings

After all the survey form is completely done and collected from the 120 respondents, the data will be analyzed based on the responses from the respondents.

According to the data collected from the 120 respondents, 65 male respondents and 55 female respondents have participated in the survey. Most of the respondents are in age group of 18-24 years old and their highest education are pre-university or foundation as the targeted participants are focusing on university students. 75 of the 120 respondents have used before Smart Tourism Guide Application, it determines that they will be familiar with smart tourism technology. However, 45 of the 120 respondents may not be familiar with smart tourism technology as they did not used before any smart tourism guide application. Most of the respondents strongly agree that it is important to understand the information of the destinations before they travel there. This result in the news and information feature in smart tourism guide application is very important for all the respondents to get updated with all the latest news and information regarding the destinations. Besides, most of the respondents agree that bad impression of the places will affect their intention to revisit the destinations. Therefore, it is important to maintain the good impression of the tourists so they will have more intention to revisit the destinations. Go.Travel is to help all the tourists to have best experience during their trip so they will have good impression of the places. This result in increasing rate of intention to revisit the destinations. Furthermore, most of the respondents think that the feature of recommendation of nearby place is very important as they strongly agree that the recommendation of nearby places could solve their trouble on deciding where to go. Therefore, this feature will be added into Go.Travel to help in recommending all the nearby places of certain area. Moreover, most of the respondents strongly agree that online communities can allow them to understand more about the knowledges and cultures of other destination. Therefore, the feature of travel forum will be added into Go.Travel to allow the respondents to share their travel experience or story to other travellers. In addition, most of the respondents strongly agree that reviews and ratings of a destination is important for them when planning for a trip. Therefore, the feature of adding review and rating will be added into Go.Travel to allow the respondents to add their review and rating for certain place so other travelers can look for it while planning for a trip. Moreover, most of the respondents would like to use Smart Tourism Guide Application in the future. It results in they would be interested to use Go.Travel. Therefore, Go.Travel will be developed to meet all user's expectation so the users will be satisfied with Go.Travel and continue to use it in the future. Lastly, nearby recommendation of places as well as news and information is the top 1 and top 2 features that will attract respondents to use it. Therefore, this 2 features will be focus more in the development of proposed Smart Tourism Guide Application, Go.Travel. It will be designed to become more attractive and user-friendly so that it will achieve user's expectation and satisfaction to provide the best travel experience for users.

5. User Interface Design

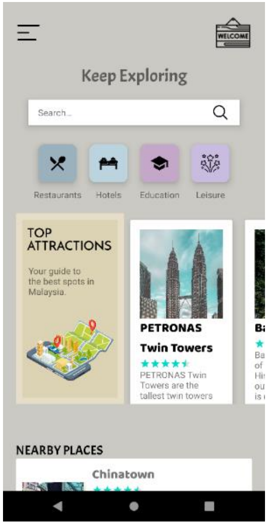


Fig. 1. Homepage

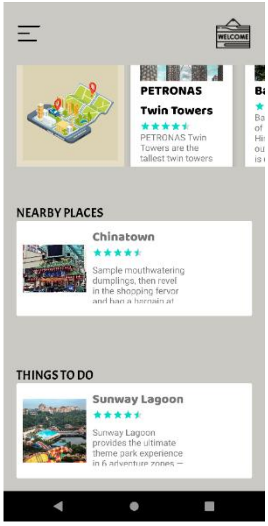
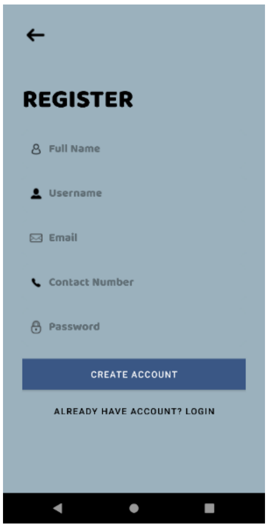


Fig. 2. Homepage

Figure 1 and 2 above shows the homepage of Go.Travel In the homepage, there are side menu bar icon that allow user to perform more tasks in Go.Travel. Welcome board icon that allows user to login or register for an account. Search bar that allows user search certain place. Some categories of places that allow user to look for category of certain place. There are more categories provided in all categories page. Top attractions, nearby places, things to do are recommended in the homepage. User can swipe to view more top attractions, nearby places or things to do that are suggested by Go.Travel. All the places are nearby Kuala Lumpur area as Kuala Lumpur is set as the location for the mobile emulator. For first time users, they can click on the welcome board icon to register for an account.



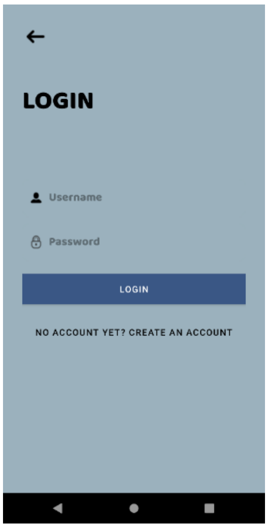
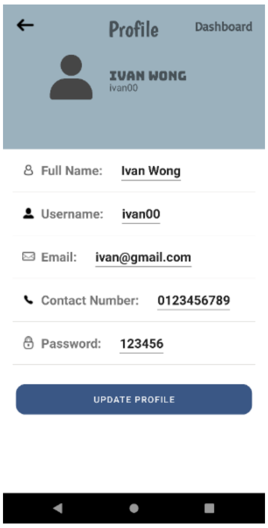


Fig. 3. Register Fig. 4. Login

Figure 3 and 4 above shows the registration page and login page of Go.Travel. Users can fill in the required information in the registration form to register for an account. The given account information will be authenticated to ensure that it is a valid information. Users are required to enter username and password to login to their account. Only registered users can login to their own account.



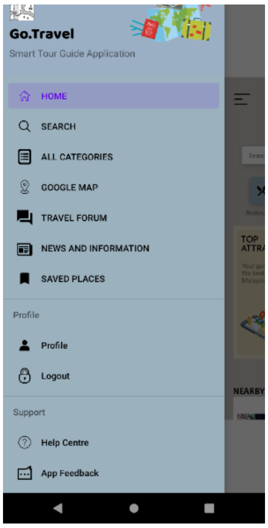


Fig. 5. Profile Fig. 6. Side Menu

Figure 5 above shows the profile page of each user. All the user's personal information will shown in this page. Users are free to edit their profile information. Figure 6 above shows the side menu of Go.Travel. It shows all the features can be performed in Go.Travel.

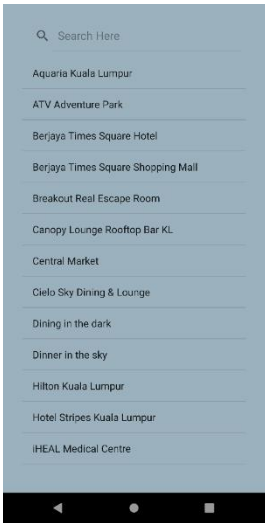


Fig. 7. Search Fig. 8. All Categories

Figure 7 above shows the search feature in Go.Travel. Users can search for all the certain places to view the detail information for each place. Figure 8 above shows all categories page. All the places in Go.Travel is divided into total of 6 categories which is restaurants, hotels, education, leisure, shopping and medical.

Fig. 9. Nearby Hotels Fig. 10. Detail Information

Figure 9 above shows all the places in hotels category. Users can search for the category of places to view all the places for each category. Users can click on each place to view the detail information page for each place. Figure 10 above shows the example of detail information page for each place.

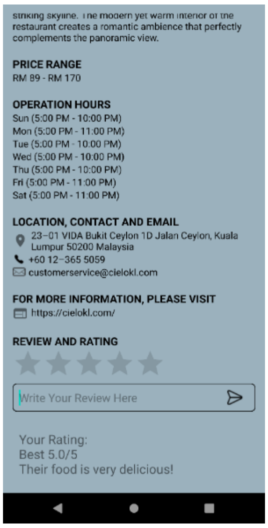


Fig. 11. Rate and review

Fig. 12. Google Maps

Figure 11 above shows rate and review feature in Go.Travel. There is rate and review section in the bottom of detail information page for each place. Users can write their own rating and review by click on the amount of star users want to rate and write their description below then click submit. The rating and review of user will shown in the bottom of rate and review section. Figure 12 above shows the Google Maps feature in Go.Travel. Users can open Google Maps in the application to view nearby places based on user's current location.

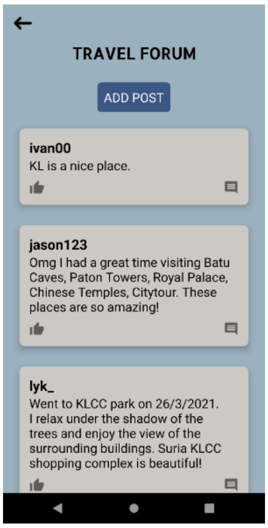


Fig. 13. Travel Forum

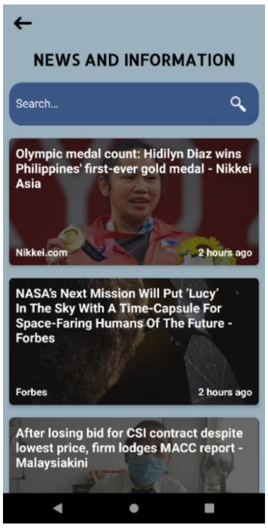


Fig. 14. News

Figure 13 above shows travel forum page in Go.Travel. Users can share their travel experience or story in the travel forum by clicking on add post button, then write their username and the description and press submit. The new post will be added in the travel forum. Figure 14 above shows news and information feature in Go.Travel. Users can read all the latest articles provided in the news and information page.

Fig. 15. Save Place

Fig. 16. Saved Places

Figure 15 and 16 above shows the save place feature in Go.Travel. Users can proceed to detail information page for each interested place. They can click on the save button beside the image as shown in figure 15 then the place will be saved into the saved places collection as shown in figure 16.

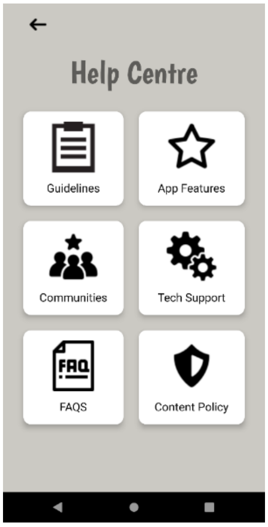


Fig. 17. Help Centre

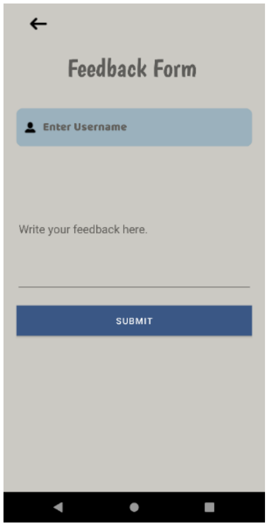


Fig. 18. Feedback Form

Figure 17 above shows the help centre in Go.Travel. Help Centre provides several type of documentation for users to read to have more understanding regarding all the features, guidelines, and rules in Go.Travel. Figure 18 above shows the feedback form provided for users to fill in and submit their own feedback. They can send feedback or opinions to the developer for future improvements.

6. Conclusion, Limitation and Recommendation

For conclude, the research project's objectives are clearly outlined in the methodology procedure, which is used to carry out the entire process for solving the issue found in existing application. The purpose of this project is to develop a user friendly, convenient, and efficient smart tourism guide application that runs on mobile platform to help all the tourists to maximize their travel experience. The research and systematic review on existing application had been carried out to determine the existing issues. A Smart Tourism Guide Application, Go.Travel was designed and implemented to meet all the objectives and improves all the existing problems. The proposed application was implemented using Java technology and is compatible with Android-based mobile devices. Although it not perfect, but it had successfully met all the objectives and improved the user's experience based on the feedback from the participants that participated in user acceptance test. The proposed application may continue to be developed and enhanced in the future.

The proposed application is not in a perfect condition due to several limitations for this project. The main limitation for this project is the developer constraints such as lack of programming knowledge to implement the proposed application. Besides, limited time is another constraint for this project. A better output of proposed application can be produced if there is more time provided. Developer has limited time to do research and learn knowledges and skills needed to develop the proposed application. It causes the developer to face several problems while implementing all the features in the proposed application. According to the feedback from the user acceptance test, there are several limitations mentioned by the participants. Some of the pages in the proposed application is slightly laggy. Besides, less category of places is provided in the proposed application. Furthermore, Google Maps does not have search function for users to search certain place or nearby suggestion of places. Moreover, like and comment feature in travel forum is not working. Lastly, the search feature in news and information is not working.

All the recommendation and feedback from users will be taken into consideration for the future enhancements and improvements to the proposed application. The proposed application is an android-based application, it can be improved to support other platforms such as IOS, windows and so on. It will result in increasing numbers of different kind of users to use the proposed application. Besides, the application's adaptability with the device's screen resolution and size may be improved to make it more dynamic. The proposed application will become more attractive and able to capture potential user attention by being built to be compatible with multiple screen sizes and resolutions. Additionally, forgot password feature can be added in Go. Travel to allow users change their forgotten password in login page. Furthermore, more categories of places can be added into the proposed application to provide more choices for tourists. Moreover, search feature can be added in Google Maps to allow tourists to search for certain place or nearby suggestion of places. In addition, like and comment feature can be added for each post in travel forum to allow user to interact with other user. Lastly, the proposed application can be improved to allow tourists post image with description in the travel forum so that other tourists can view the shared destinations by image and not imagination.

References

1. S. P. Singh, and P. Singh, “Design and implementation of a location-based multimedia mobile tourist guide system,” International Journal of Information and Communication Technology, 2015, 7(1), 40.
2. A. Jauhari, F. A. Mufarroha and M. Rofi, “ The Development of Smart Travel Guide Application in Madura Tourism,” Proceedings of the 3rd International Conference on Social Sciences (ICSS 2020), 2020, 473.
3. X. Shi, T. Sun, Y. Shen, K. Li, and W. Qu, “Tour-Guide: Providing Location-Based Tourist Information on Mobile Phones,” in 2010 10th IEEE International Conference on Computer and Information Technology, 2010.
4. H. C. Kim and Y. S. Kim, “Smart Tourism Information System using Location-based Technology,” International Journal of Software Engineering and Its Applications, 2016, vol. 10, no. 11, pp. 11-24.
5. D. Hobbs, “Guide Dogs creates app with Microsoft to help people with sight loss navigate cities,” March 2018.

6. N. B. Nugraha and E. Alimudin, "Mobile Application Development for Tourist Guide in Pekanbaru City," The 2nd International Conference on Computer Science and Engineering Technology, 2020, 14-30.
7. Y. Li, C. Hu, C. Huang, and L. Duan, "The concept of smart tourism in the context of tourism information services. Tourism Management," 2017, 58, 293–300.
8. B. Gaille, "9 Pros and Cons of Yelp," December 2016.
9. Hostelgeeks, "Klook in Review 2021 – Is it the best way to book unique experiences?" November 2020.

Authors Introduction

Mr. Wong Yit Meng

He is currently an undergraduate student who is pursuing study in BSc (Hons) in Computer Science from Institute of Computer Science and Innovation (ICSDI) in UCSI University, Malaysia.

Assoc. Prof. Dr. Abdul Samad bin Shibghatullah

He received the bachelor accounting degree from Universiti Kebangsaan Malaysia, Bangi, Malaysia in 1999, the M.Sc. degree in computer science from the Universiti Teknologi Malaysia, Skudai, Malaysia in 2002, and the Ph.D. degree in computer science from the Brunel University of Uxbridge, United Kingdom. He is currently Associate Professor at Institute of Computer Science & Digital Innovation, UCSI University, Kuala Lumpur, Malaysia. His current research interests include optimization, modelling and scheduling.

Asst. Prof. Ts. 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 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.

Gesturenomy: Touchless Restaurant Menu Using Hand Gesture Recognition

Ian Christian Susanto
Institute of Computer Science & Digital Innovation, UCSI University, UCSI Heights, 1 Jalan Puncak Menara Gading, Kuala Lumpur, WP Kuala Lumpur 56000, Malaysia

Kasthuri Subaramaniam
Institute of Computer Science & Digital Innovation, UCSI University, UCSI Heights, 1 Jalan Puncak Menara Gading, Kuala Lumpur, WP Kuala Lumpur 56000, Malaysia

Abdul Samad bin Shibghatullah
Institute of Computer Science & Digital Innovation, UCSI University, UCSI Heights, 1 Jalan Puncak Menara Gading, Kuala Lumpur, WP Kuala Lumpur 56000, Malaysia
E-mail: 1001954676@ucsiuniversity.edu.my, kasthurisuba@ucsiuniversity.edu.my, abdulsamad@ucsiuniversity.edu.my
www.ucsiuniversity.edu.my

Abstract

Hand gesture recognition is one of the more modern methods of human-computer interaction. However, study of its application in real world setting is sparse, especially its use in digital restaurant menus. We discuss various types of computer vision-based hand gesture recognition and decide on hand pose estimation as the method of recognition used. An analysis into the requirements show that respondents are concerned with the hygiene of touchscreen digital menus. We provide a description of the system being developed. The system passes the user acceptance tests given.

Keywords: digital restaurant menu, hand gesture recognition, hand pose estimation, human computer interaction

1. Introduction

A post-WIMP interface is any interface that does not use menus, forms, or toolbars but relies on more convenient interactions, of which hand gesture recognition is an example.[1] Hand gesture recognition has a few advantages, namely sterility as the user never touches the device, accessibility by people with physical handicaps, and efficiency from using three dimensions.[2] There are two main types of hand gesture recognition, with wearable technology or with computer vision.[3] While the first is simpler to implement, it may be intrusive and requires additional equipment. This project will use the latter.

Transmission of SARS-CoV-2 through inanimate surfaces is doubted[4], however other diseases may still be transmitted through fomites. Morens et al.[5] reported fomites or contaminated hands as the transmission method in a nursing home outbreak of influenza A. People should err on the side of caution. There are various ways of preventing disease transmission, such as through touchless technology. Transmission through this method is unlikely as no surface contact is performed.

This paper proposes the use of hand gesture recognition in digital restaurant menus. We discuss the analysis, design, implementation, and evaluation of the application.

We now state the research objectives:

- To investigate the current hand gesture recognition techniques.
- To design an application that utilizes the most suitable technique.

- To develop a prototype of a hand gesture control application.
- To evaluate the effectiveness of the prototype application.

The paper will be structured as follows. Section 1, this section, elaborates on the problem statement, the research objectives, and the report structure. Section 2 studies similar systems that have been developed in the past, ending with a comparison on each system. Section 3 explains the research methodology of the project, consisting of the software development life cycle model chosen and the method of requirements gathering. Section 4 shows the result of data gathering and the design of the system.

2. Literature Review

2.1. *Hand Gesture Recognition*

Oudah et al.[3] splits vision-based hand gesture recognition into seven categories:

Color-based recognition. The system determines hand gesture based on the color of each pixel. An example is by using color gloves whose colors are tracked with nearest-neighbor approach[6]. However, this still requires the subject to wear an additional glove, which is not suitable for some purposes. Another way is to use skin color. Perimal et al.[7] converts the image into YCbCr color space, determines which pixels are colored, removes parts that are not the tip of the finger, and counting the fingers.

- *Appearance-based recognition.* Features are extracted from the shape of the hand, giving the advantage of having similar behavior on varying skin colors. One way to do this is through edge detection. Kulkarni and Lokhande[8] converts the image into grayscale before performing edge detection and using neural networks to classify it into gestures.
- *Motion-based recognition.* Gesture is determined by seeing how the image changes from one image to another. Molina et al.[9] uses time-of-flight cameras to capture the depth data of an image. The closest point of the image is chosen as the point of interest. The system determines the trajectory of the hand and compares it with the synthetic patterns. The closest pattern is chosen as the gesture.
- *Skeleton-based recognition.* Geometric features of the hand such as joint orientation, space between joints, skeletal joint location, and angle between joints are extracted from the image. Konstantinidis et al.[10] extracts body and hand skeletal data from the image and compares the joint positions relative to the neck joint. Occluded joints use data from the previous frames.
- *Depth-based recognition.* Using depth information to extract gestures has a few advantages, namely that the system no longer needs to consider lighting, shade, and color. This category of recognition system usually uses depth-sensing devices such as Microsoft Kinect to get depth information. Ma and Peng[11] combines depth information and color information to obtain the region of interest and locates the fingertips.
- *3D model-based recognition.* The algorithm uses a 3D model to be compared to the 2D input. Ge et al.[12] uses constructs a 3D model of the hand using Graph CNN, which gives a more detailed output.
- *Deep learning-based recognition.* The usage of deep learning is useful as gesture recognition is a complex task, however gathering datasets to train the neural network requires much effort. Wu[13] used a double channel convolutional neural network, where one channel gets the denoised version of the image, and the other channel gets image after being processed through edge detection. The result has 98.02% recognition rate.

The method used in this project is through pose estimation, which has an advantage of the ease of adding or modifying new gestures to recognize. Such systems typically follow two steps: first, the neural network (NN) is used to find the location of the hand, which is used to crop and resize the image to contain only the hand. Another NN will perform the actual pose estimation. Zimmerman and Brox[14] treats the image as a segmentation problem, whose output is used to determine the position of the hand. The cropped RGB image is fed into an encoder-decoder model (PoseNet) that outputs 21 two-dimensional score maps. These score maps are used as input into PosePrior which regresses the 3D coordinates of each 21 joints and a 3x3 transformation matrix between the canonical coordinates and the viewpoint coordinates. Malik et al.[15] feeds a depth map into a convolutional neural network (CNN) to get a feature map, which is fed into a regression network to get the hand joint position, hand mesh vertices, and structural constraint. Spurr et al.[16] proposes the use of biomechanical constraints to increase the accuracy of already existing pose estimation system, allowing it to be weakly supervised. The biomechanical constraint function can be added as a loss function to an already existing model. Zhang et al.[17] presents a real time hand pose estimation solution with RGB input. The hand detection system only detects the palm, as it has a simpler shape compared to the entire hand. In addition to being cropped and resized, the image is also rotated so that the hand is always in the same direction. The hand landmark detection outputs 21 landmark, like Ref. 14. It also outputs the probability of a hand being present and its handedness (left or right). The system achieves real-time performance, even on mobile devices.

2.2. *Digital Restaurant Menus*

The most basic form of restaurant menu is paper based. However, this has the drawbacks where the customer needs to wait to order food. In addition, it also wastes paper, as a different customer needs a different piece of paper to track the order[18]. Some digital restaurant menu systems are built to counteract these drawbacks.

Şahin[19] classifies digital menus into touchscreen menu systems which allow for self-service and non-touchscreen menu systems which require interaction with the staff to place an order. However, this classification may be misleading as this paper describes a touchless self-service system, hence we will refer to these as *interactive* and *non-interactive* respectively.

There are two types of non-interactive menus. A static digital menu board only uses static images that may change periodically to show different items, while a dynamic digital menu board has part of or the entire screen moving (i.e. motion graphics or videos). Interactive menu boards commonly use tablets, tabletop touchscreen devices, mobile apps, and kiosks. Bhargave et al.[18] built an Android-based digital ordering system where the restaurant provides a tablet at each table to browse the menu. This reduces paper waste and allows for self-service.

Rusdi et al.[20] uses a mobile phone as a digital menu, which gives the convenience of having the entire menu on the customer's phone, reducing the need for tablets as Ref. 18 does. However, the customer might not have an internet connection or sufficient battery power. Downloading an application may not be attractive to some users.

2.3. *Comparison of Hand Gesture Recognition Systems*

We compare various existing systems that use hand gesture recognition. Desai and Desai[21] used hand gesture recognition to control a home automation system. The recognition is performed by using Microsoft Kinect as input. Four gestures are chosen to represent switching on/off lights, fan, phone charger, and TV set. A fifth gesture is used to shut down all appliances that were on. The accuracy of the system is 88%.

Suriya and Vijayachamundeeswari[22] used hand gesture recognition to control mouse on a desktop computer. Skin color detection is used, and the fingers are located by using convex hull algorithm. A linear discriminate analysis is performed to classify the gestures. The gestures being tracked are Arabic numerals alphabet.

Gestix[23] is a hand gesture-based interface used by doctors to navigate electronic medical record (EMR) databases without coming into contact with the computer. The system uses color and motion data to segment the image to determine the shape of the hand. A state machine is used to determine which gestures are being made. Surgeons' focus was shown to be improved and the interface proved to be intuitive.

Al et al.[24] use hand gesture recognition to control a Universal Robot (UR3) arm. The system uses proximity sensors to detect the position of the hand. The data is transferred from the Arduino to a MATLAB instance to be classified by an artificial neural network (ANN) into *up*, *down*, *left*, or *right*.

Zhang et al.[25] uses hand gesture recognition to control the movement of a player in a VR game. The input is gathered using the Leap Motion controller mounted in front of an Oculus Rift headset. The system was found to be intuitive, easy to learn and use, and causes low fatigue. The immersion of the player is improved and motion sickness is reduced.

Out of the system described above, the one that will most resemble this project is Gestix, due to it only using color and motion data to control a graphical user interface. The system described by Ref. 21 is also similar in spirit, in that each gestures are mapped to one action. In addition, the gestures are used to navigate a menu, in comparison to HRI or DHGI, which interacts with a robot arm or a VR player respectively.

3. Research Methodology

3.1. *System Development Life Cycle*

The project is developed on the *prototyping* model. It is an iterative development model which is used when details of the system (i.e. input, process output) are not clearly defined yet. The first step is *requirements* *gathering and analysis*, where the developer gathers expectations from users. The process of requirements gathering in this project will be elaborated in the next subsection.

The second step is *quick design*, where a simple version of the system is designed to get a brief idea of the software. In this project, this is done by listing some required features. The first prototype, for example, could only view the menu, scroll it left to right, and swap between recommended menu and the entire menu.

The fourth step is *user evaluation*, where the prototype is shown to the client or end user. The user gives feedback such as user experience or adding new features. The fifth step is *refining prototype*, which uses the feedback to refine the requirements and design of the prototype. After building a new prototype, user evaluation is performed again. This process is repeated until the result is satisfactory.

The sixth step is *product implementation and maintenance*, where a final implementation is built with care into building a high quality and maintainable system. This project stops at product implementation, as maintenance is out of the scope of the project.

3.2. *Questionnaire*

The process of requirements gathering is done through a questionnaire. The questionnaire is shared through UCSI CourseNetworking page to UCSI University undergraduate and masters students and gathered a total of 29 responses. Questions that are asked include the respondent's opinions on digital restaurant menus, what features they think should be in a menu, their concerns of the hygiene of conventional menus, and their opinions on what advantages and disadvantages using hand gesture recognition has. Most of the questions are in long form to gain a deeper understanding than possible with multiple choices.

4. Analysis and Design

4.1. *Analysis*

We analyze the findings from the questionnaire as described in the previous section. The first question asks whether restaurants should use digital restaurant menus, ranging from 1 (strongly disagree) to 5 (strongly agree). The result can be seen in Figure 1. Fifteen respondents chose 4 (agree), and the average score is 4.17. We

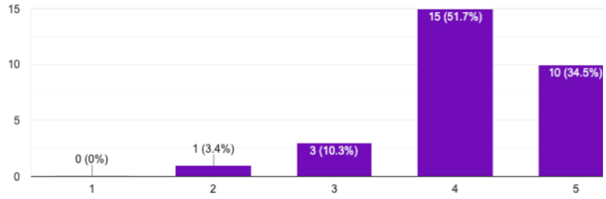


Fig. 1. Chart of responses to whether restaurants should use digital restaurant menus.

conclude that most people agree restaurant should use digital restaurant menus.

Two questions were asked relating to features that should be in a digital restaurant menu. The first question lists features found in other digital menus, while the second asks for any other features needed in a restaurant menu. Table 1 shows the average score for each features, where 1 is very undesirable and 5 is very desirable. Other features that have been requested (in decreasing frequency) include: dish price, special offers, special sets, digital payment methods, cart system, waiter notification system, and description of the freshness of the ingredients.

Table 1. Average score of each features

Feature	Avg. Score	Feature	Avg. Score
Image of the dish	4.52	Non-allergy warnings (e.g. spicy food)	4.59
Description of the dish	4.34	Recommended	4.38
Nutritional content	3.76	List of ingredients	4.17
Dish variations	4.14	Estimated preparation time	3.93
Allergy warnings	4.59		

Next, the respondents are asked about their concerns of the hygiene of touchscreen digital restaurant menus. Seventeen respondents expressed concerns about the hygiene of touchscreen digital menus. Reasons for the concern include the fact that the device is used by many people, who may leave virus or bacteria on it. Some expressed their distrust in the restaurant's ability to manage the hygiene of their touchscreen surfaces. Respondents who do not have hygiene concerns have trust in the restaurant's hygiene policies.

The next question asks whether the respondents believe that gesture recognition could solve this problem. There are 19 responses that agree with this, most citing the lack of surface contact as the reason. Three respondents are neutral. Four responses expressed distrust in the technology, and that there needs to be more testing to evaluate its feasibility.

The last question asks whether the respondents have any worries about issues that may arise from the use of this technology. Fifteen respondents have concerns with the recognition systems, as there is a risk of low accuracy and slow performance. Ease-of-learning was a concern in 7 of the responses. Maintenance issues was brought up by 2 responses. Two responses are concerned with the possibilities of making mistakes. Five responses have no concerns about any issues.

4.2. *Design*

In this section, we look at the system's design created to fulfill the requirements. Figure 2 shows the use case diagram of the system. It consists of two main parts; the customer-facing parts and the staff-facing part.

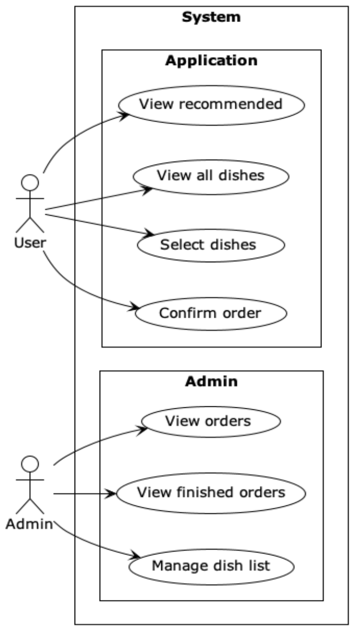


Fig. 2. Use case diagram of the system.

Figure 3 shows the activity diagram of the user's experience interacting with the system.

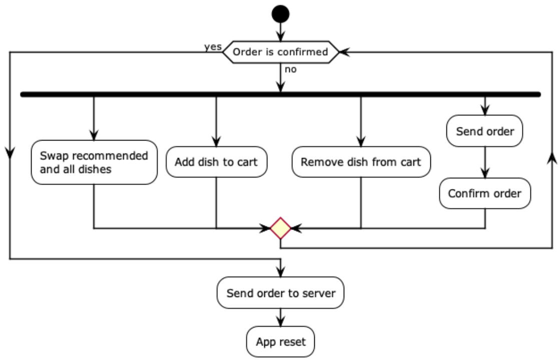


Fig. 3. Activity diagram of the user's interaction with the system.

5. Implementation and Evaluation

In this section, we look at the list of features that the system has and the list of tests that are performed to ensure the system is satisfactory.

5.1. *Implementation*

The customer-facing application is designed as follows: most of the screen is taken up by the menu itself. Below it, the gesture that is detected is shown as feedback to the user. There is also the cart. This cart shows which items the user has picked, how many of each, and the total price.

The user may swap between viewing the recommended screen and viewing all offered dishes by holding a peace sign. To prevent accidentally swapping, the user must hold this gesture for approximately one second. The user can scroll through the menu by making a fist gesture and moving it left to right.

To select an item, the user performs a "point up" gesture. A pointer that moves with the user's hand is shown when it is visible. By aiming this pointer, the user can add items into the cart. To add multiple, the user can redo the gesture again or hold the gesture. The "point down" gesture behaves similarly, but removes items from the cart instead.

Removing an item from the cart is intuitively done by performing a "point down" gesture. Removing multiple items can be done similarly to adding multiple items.

When the user is ready selecting, they can perform the "thumbs up" gesture. When this is done, the user will be asked to confirm whether they are sure of this action. To decline, a "thumbs down" gesture can be done. This will send the user back to the original screen. Otherwise, the user can do the "thumbs up" gesture again. This will send the user's order to the staff. A thank you screen is shown, along with the total price and the order number.

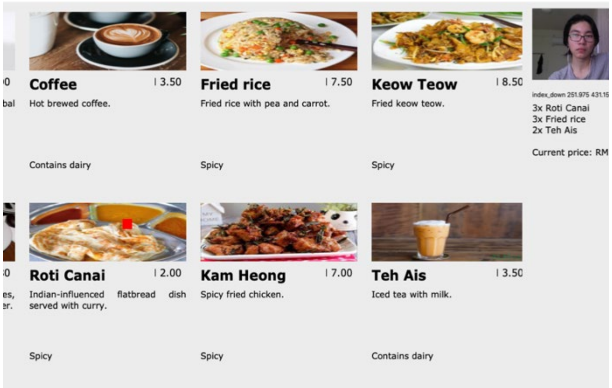


Fig. 4. Screenshot of the application with the cart being filled.

For convenience, if the cart is already filled when a user had just started using the system, they may do a "thumbs down" gesture on the menu screen to empty it.

The system is built using C++ using Qt as its GUI framework. The server uses Django backend with SQLite for its database. The two parts communicate using the REST API. Hand gesture recognition is performed by calling Python functions from the C++ application.

5.2. *Evaluation*

We describe the user acceptance tests that are conducted to deem this system satisfactory. Table 2 shows the tests that are designed to verify that each piece of the system is functioning correctly.

Table 2. The user acceptance test along with each result

No.	Description	Result
1	Start the application. Application connects to server correctly. App shows a list of recommended dishes.	Pass
2	Perform a peace sign. The menu should refresh and show more dishes that weren't in the recommended section.	Pass
3	Perform a peace sign again. This should show less items.	Pass
4	Perform a grabbing and dragging motion. This should scroll the menu in the correct direction, i.e. to the user's left if the user drags it to the left, and vice versa.	Pass
5	Move the user's hands around. The red square on the screen should follow the user's motion accurately.	Pass
6	Point up at an item. After a moment the cart should update to show that item.	Pass
7	Point up at another item, this time holding it longer. The cart should increment that item until the user stops.	Pass
8	Point down at the item specified in Test 6. This should remove that item from the cart.	Pass
9	Perform Test 7 on a different item, then point down at that item, holding it longer. This should decrease that item one by one. If the user still holds down when the item is no longer in the cart, nothing should happen.	Pass
10	Scroll the menu as described in Test 4. Point up at an item. This should update the cart with the correct item.	Pass
11	Perform a thumb up. The screen should show a dialog box confirming the user's decision.	Pass
12	Perform a thumb down. The screen should go back to the menu.	Pass
13	Perform Test 11 and perform another thumb up gesture. The screen should now show a dialog box with the total price and the order number.	Pass
14	Perform Test 6 to 10, then perform a thumb down gesture. The cart should be emptied.	Pass

6. Conclusion

We have discussed the need for a digital restaurant menu with hand gesture recognition interface, and proposed such system. An application is designed and implemented. The user testing evaluation shows that the application functions as intended.

There are many limitations present in the current system. Starting from the analysis step, the sample size and demographics of the questionnaire means that this result may not be an accurate representative of the public. Highly rated features as dish variations were not present. In addition, a more in-depth user testing would be beneficial, such as to test the intuitiveness and ease-of-use of the system.

This project contributes to the knowledge base by providing one case study of the use in hand gesture recognition in digital restaurant menus, as research in this application of gesture recognition is currently minimal. The author hopes that this project can open the doors for more research into this topic.

In the future, research into how this system affects businesses can be looked into, including its effects on the maintenance costs, efficiency of running the restaurant, and public reception. Analysis could be done more accurately by using a larger and more diverse sample size. The effectiveness of each hand gesture can be looked into more deeply, as people of different backgrounds could have different understandings on which gestures would feel more intuitive. A way to customize which gestures correspond to an action would be beneficial for this purpose. More of the features requested in the questionnaire could be implemented.

References

1. A. Van Dam, "Post-WIMP user interfaces", Communications of the ACM vol. 40(2), pp.63-67, 1997.
2. J.P. Wachs, M. Kölsch, H. Stern, and Y. Edan, "Vision-based hand-gesture applications", Communications of the ACM, vol. 54(2), pp.60-71, 2011.
3. M. Oudah, A. Al-Naji, and J. Chahl, "Hand Gesture Recognition Based on Computer Vision: A Review of Techniques", Journal of Imaging, vol. 6(8), p.73, 2020.
4. E. Goldman, "Exaggerated risk of transmission of COVID-19 by fomites", The Lancet Infectious Diseases, vol. 20(8), pp.892-893, 2020.
5. D.M. Morens, V.M. Rash, "Lessons from a Nursing Home Outbreak of Influenza A", Infection Control & Hospital Epidemiology, vol. 16(5), pp.275-280, 1995.
6. R.Y. Wang, J. Popović, "Real-time hand-tracking with a color glove", ACM Transactions on Graphics (TOG), vol. 28(3), pp.1-8, 2009.
7. M. Perimal, S.N. Basah, M.J.A. Safar, H. Yazid, "Hand-Gesture Recognition-Algorithm based on Finger Counting", Journal of Telecommunication, Electronic and Computer Engineering (JTEC), vol.10(1-13), pp.19-24, 2018.
8. V.S. Kulkarni, S.D. Lokhande, "Appearance Based Recognition of American Sign Language Using Gesture Segmentation", International Journal on Computer Science and Engineering, vol.2(03), pp.560-565, 2010.
9. J. Molina, J.A. Pajuelo, J.M. Martínez, "Real-time Motion-based Hand Gesture Recognition from Time-of-Flight Video", Journal of Signal Processing Systems, vol.86(1), pp.17-25, 2017.
10. D. Konstantinidis, K. Dimitropoulos, P. Daras, "Sign language recognition based on hand and body skeletal data", in 2018-3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-Con), pp.1-4, 2018.
11. X. Ma, J. Peng, "Kinect Sensor-Based Long-Distance Hand Gesture Recognition and Fingertip Detection with Depth Information", Journal of Sensors, 2018.
12. L. Ge, Z. Ren, Y. Li, Z. Xue, Y. Wang, J. Cai, J. Yuan, "3D Hand Shape and Pose Estimation from a Single RGB Image", in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp.10833-10842, 2019.
13. X.Y. Wu, "A hand gesture recognition algorithm based on DC-CNN", Multimedia Tools and Applications, pp.9193-9205, 2019.
14. C. Zimmermann, T. Broz, "Learning to Estimate 3D Hand Pose from Single RGB Images", in Proceedings of the IEEE International Conference on Computer Vision, pp.4903-4911, 2017.
15. J. Malik, A. Elhayek, F. Nunnari, D. Stricker, "Simple and effective deep hand shape and pose regression from a single depth image", Computers & Graphics, vol.85, pp.85-91, 2019.
16. A. Spurr, U. Iqbal, P. Molchanov, O. Hilliges, J. Kautz, "Weakly Supervised 3D Hand Pose Estimation via Biomechanical Constraints", in Computer Vision-ECCV 2020: 16th European Conference, Glasgow, UK, August 23-28, 2020, Proceedings, Part XVII 16, pp.211-228, 2020.
17. F. Zhang, V. Bazarevsky, A. Vakunov, A. Tkachenka, G. Sung, C.L. Chang, M. Grundmann, "Mediapipe Hands: On-device Real-time Hand Tracking", 2020.
18. A. Bhargave, N. Jadhav, A. Joshi, P. Oke, S.R. Lahane, "Digital Ordering System for Restaurant Using Android", International Journal of Scientific and Research Publications, vol.3(4), 2013.
19. E. Şahin, "An Evaluation of Digital Menu Types and Their Advantages", Journal of Tourism and Gastronomy Studies, vol.8(4), pp.2374-2386, 2020.
20. J.F. Rusdi, N.A. Abu, N. Agustina, M. kchouri, S. Dewi, "Software Development Stages of Mobile Computing Implementation in Restaurant Food Ordering", SciTech Framework, vol.1(1), pp.24-33, 2019.
21. S. Desai, A. Desai, "Human Computer Interaction Through Hand Gestures for Home Automation using Microsoft Kinect", in Proceedings of International Conference on Communication and Networks, pp.19-29, 2017.
22. R. Suriya, V. Vijayachamundeeswari, "A survey on hand gesture recognition for simple mouse control", in International Conference in Information Communication and Embedded Systems (ICICES2014), pp.1-5, 2014.
23. J. Wachs, H. Stern, Y. Edan, M. Gillam, C. Feied, M. Smith, J. Handler, "Gestix: A Doctor-Computer Sterile Gesture Interface for Dynamic Environments", in Soft Computing in Industrial Applications, vol.39, pp.30-39, 2008.
24. G.A. Al, P. Estrela, U. Martinez-Hernandez, "Towards an intuitive human-robot interaction based on hand gesture recognition and proximity sensors", in 2020 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI), pp.330-335, 2020.
25. F. Zhang, S. Chu, R. Pan, N. Ji, L. Xi, "Double hand-gesture interaction for walk-through in VR environment", in 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS), pp.539-544, 2017.

Authors Introduction

Ian Christian Susanto

He is currently an undergraduate student, pursuing his study in BSc (Hons) in Computing from Institute of Computer Science and Digital Innovation (ICSDI) at UCSI University, Malaysia.

Asst. Prof. Ts. 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 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.

Assoc. Prof. Dr. Abdul Samad bin Shibghatullah

He received the bachelor accounting degree from Universiti Kebangsaan Malaysia, Bangi, Malaysia in 1999, the M.Sc. degree in computer science from the Universiti Teknologi Malaysia, Skudai, Malaysia in 2002, and the Ph.D. degree in computer science from the Brunel University of Uxbridge, United Kingdom. He is currently Associate Professor at Institute of Computer Science & Digital Innovation, UCSI University, Kuala Lumpur, Malaysia. His current research interests include optimization, modelling and scheduling.

A research of infectivity rate After the Consecutive Holidays

Saori Iwanaga
Japan Coast Guard Academy (JCGA),
5-1 Wakaba, Kure, 737-8512, Japan
s-iwanaga@jcga.ac.jp
www.jcga.ac.jp

Abstract

I proposed a discrete mathematical SEPIR model for seasonal influenza. In this study, by examining affection by pre-infectious students in real data, I found that super-spreading depends on the timing of penetration. I show the students that super-spread seasonal influenza according to the day the first patients are discovered. According to the data, on average it takes for days for seasonal influenza measures to be implemented within the communities.

Keywords: Epidemic model, Flu, super-spread, infectivity rate.

1. Introduction

An epidemic of Flu: seasonal influenza occurred at JCGA: Japan Coast Guard Academy in January, 2017. After three consecutive holidays from January 7th to 9th, students returned to the dormitory and started taking classes. On Friday January 13th, two students developed Flu. At that time, there were 150 undergraduate students. Finally, 37 students, 20 freshmen, 13 sophomores and 4 juniors developed the Flu [1]. Of the 60 teachers not one developed Flu. According to our medical doctor at that time, usually about five students develop Flu at the dormitory in a usual epidemic wave.

An epidemic of SARS: Severe Acute Respiratory Syndrome occurred in Singapore in 2003[2]. Five people were categorized as super spreaders of SARS who directly affected more than ten people. At JCGA, I also found the super-spreading of Flu. In this paper, I show super-spreading according to the day that first patients are discovered. Here, I define super-spreading as that of more than certain number students which are directly affected.

2. Mathematical Model

Kermack et.al proposed SIR model for epidemics[3]. The state transition diagram of an individual is shown in Fig. 1(a). “S”, “I” and “R” means susceptible state, infectious state and recovered state. An individual of “S” can transit to “I” by contact with individual of “I”. Then, an individual moves state from “I” to "R" depending on how long it has been in “I”.

Keeling et.al[4] introduced the incubation period, that is exposed “E” state into SIR model and proposed mathematical SEIR model, which many childhood infectious diseases follow. Individual of “E” is affected but not yet infectious. Only individual of “I” can affect individual of “S”. The state transition diagram of an individual is shown in Fig. 1(b).

At JCGA, students are not affected by students of “I” because patients are isolated in sick rooms. Then, I introduced pre-infectious state “P” into SEIR model and proposed the discrete mathematical SEPIR model for Flu[1]. The incubation period is divided into two periods, the exposed period and the infectious period, but neither have any symptoms. I set the former as exposed state “E”

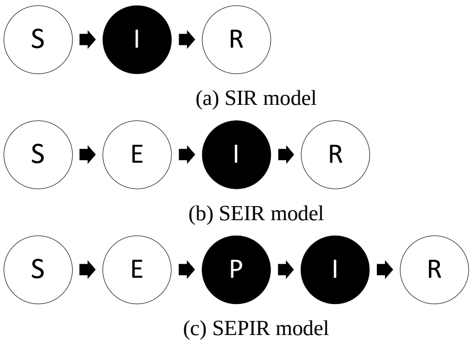


Fig. 1. The state transition diagram of an individual: circle shows state of individual. S: Susceptible state, E: Exposed state, P: Pre-infectious state, I: Infectious state and R: Recovered state. Black circle means that it can affect others.

and the latter as pre-infectious state "P". The state transition diagram of an individual is shown in Fig. 1(c).

$$\Delta S = S(t+1) - S(t) = -\alpha S(t)P(t) - \beta S(t)I(t) \quad (1)$$

$$\Delta E = E(t+1) - E(t) = \alpha S(t)P(t) + \beta S(t)I(t) - \sigma E(t) \quad (2)$$

$$\Delta P = P(t+1) - P(t) = \sigma E(t) - \tau P(t) \quad (3)$$

$$\Delta I = I(t+1) - I(t) = \tau P(t) - \gamma I(t) \quad (4)$$

$$\Delta R = R(t+1) - R(t) = \gamma I(t) \quad (5)$$

Individual of "I" or "P" can affect individual of "S". There is some contact between individual of "S" and that of "I". The probability of the contact is determined by the respective numbers of "S" and "I". Considering a mean infectivity rate β, individual of "S" moves "E" as given in Eq (1)[3]. There is also some contact between individual of "S" and that of "P". The probability of the contact is determined by the respective numbers of "S" and "P". By introducing infectivity rate α, individual of "S" moves "E" as given in Eq (1)[1]. By introducing transmission rate σ, individual of "E" moves "P" as given in Eq (2)[1]. By introducing transmission rate τ, individual of "P" moves "I" as given in Eq (3)[1]. By introducing the recovery rate γ which is the inverse of the infectious "I" period, this leads to a far more straightforward equation as shown in Eq (4)[3]. Here, S(t), E(t), P(t), I(t) and R(t) is the number of individuals of "S", "E", "P", "I" and "R". I set a base time as 08:30 and t represent days since January 9.

3. Former Results[1]

I have the JCGA data of developed students as shown by "I" in Table 1 (Case A). First column ID shows student identification number and next 18 columns show the daily state of the student. The date with underline means day off. The average of infectious "I" period is 3.86 days.

Table 1. Epidemic of Flu at JCGA in January, 2017 (Case A).

ID	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	F	C	B	Y
1	S	E	P	P	**I**	I	I	I	R	R	R	R	R	R	R	R	R	R	16	16	0	O
2	S	E	P	P	**I**	I	I	I	R	R	R	R	R	R	R	R	R	R	11	9	2	O
3	S	S	E	P	P	I	I	I	I	I	R	R	R	R	R	R	R	R	7	6	1	U
4	S	S	E	P	P	I	I	I	I	I	R	R	R	R	R	R	R	R	6	6	0	U
5	S	S	E	P	P	I	I	I	I	I	R	R	R	R	R	R	R	R	6	6	0	U
6	S	S	E	P	P	I	I	I	I	R	R	R	R	R	R	R	R	R	6	6	0	U
7	S	S	E	P	P	I	I	I	R	R	R	R	R	R	R	R	R	R	8	6	2	U
8	S	S	S	E	P	P	I	I	I	I	I	R	R	R	R	R	R	R	0	0	0	X
9	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	R	R	R	1	0	1	X
10	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	R	R	R	0	0	0	X
11	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	R	R	R	2	0	2	X
12	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	R	R	R	0	0	0	X
13	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	R	R	R	1	0	1	X
14	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	R	R	R	0	0	0	X
15	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	R	R	R	0	0	0	X
16	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	R	R	R	0	0	0	X
17	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	R	R	R	0	0	0	U
18	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	R	R	R	1	0	1	X
19	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	R	R	R	1	0	1	X
20	S	S	S	E	P	P	I	I	I	R	R	R	R	R	R	R	R	R	0	0	0	X
21	S	S	S	E	P	P	I	I	I	R	R	R	R	R	R	R	R	R	0	0	0	X
22	S	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	R	R	0	0	0	X
23	S	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	R	R	0	0	0	X
24	S	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	R	R	0	0	0	X
25	S	S	S	S	E	P	P	I	I	I	R	R	R	R	R	R	R	R	0	0	0	X
26	S	S	S	S	E	P	P	I	I	I	R	R	R	R	R	R	R	R	0	0	0	X
27	S	S	S	S	E	P	P	I	I	R	R	R	R	R	R	R	R	R	0	0	0	X
28	S	S	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	R	0	0	0	X
29	S	S	S	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	1	1	0	X
30	S	S	S	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	1	1	0	X
31	S	S	S	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	1	1	0	O
32	S	S	S	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	0	0	0	X
33	S	S	S	S	S	S	E	P	P	I	I	I	R	R	R	R	R	R	1	1	0	X
34	S	S	S	S	S	S	E	P	P	I	I	I	R	R	R	R	R	R	0	0	0	O
35	S	S	S	S	S	S	S	E	P	P	I	I	I	I	R	R	R	R	1	1	0	O
36	S	S	S	S	S	S	S	S	E	P	P	I	I	I	R	R	R	R	0	0	0	X
37	S	S	S	S	S	S	S	S	S	E	P	P	I	I	I	I	I	R	0	0	0	X
I	0	0	0	0	**2**	7	21	27	25	27	12	9	7	3	1	1	1	0				
P	0	2	7	21	25	21	13	8	8	3	2	1	0	0	0	0	0	0				
G	-	-	13	6	1	4	0	1	1	0	0	0	-	-	-	-	-	-				
CG	-	-	13	6	0	0	0	0	0	0	0	0	-	-	-	-	-	-				
BG	-	-	2	3	1	4	0	0	0	0	0	0	-	-	-	-	-	-				

I found that student 1 and 2 brought Flu to the JCGA, which is called the source of the infection. I supposed that a student is affected in their room, which is a closed space and follows the SEPIR model. By retrospective investigation of activities in closed spaces according to the schedule of students, I found that it was on campus transmission. As for infection channels, I found I just had to deal with bedrooms and classrooms as closed spaces. As for the incubation period, focusing on student of "E", I found the period of "E" is one day and the period of "P" is two days. Then, I filled in the state of the students, such as "S"," E", "P" and "R", in Table 1, as well.

4. JCGA 2017 Case

I reconfirmed to focus on the affected students by student of "P". In Table 1, F, C and B columns refer to the number of affected students by student of "P" at JCGA, in classrooms and in bedrooms. F and C of students 1 or 2 are very high. And C is more than B because students in classrooms are more than that of bedrooms. Here, F is not equal to the sum of C and B because students can affect the same student in both rooms.

I and P rows refer to the number of students of "I" and "P". I on January 13th is the most and P on January 16th or 18th is the most. G, CG and BG rows refer to the number of affected students by student of "P" at JCGA, in classrooms and in bedrooms on that day. G on January 11th or 12th is very high. CG is more than BG because students in classrooms are more than that of bedrooms. Here, G is not equal to the sum of CG and BG because students can be affected in both rooms on that day. Y means infection channel. "O" refers to outside of JCGA on weekends. "X" refers to campus transmission and 26 students of "X" derived from student 1 and 2. "U" refers to unknown.

5. JCGA 2019 Cases

In January 2019, an epidemic of Flu also occurred on JCGA training ship. Here, I define epidemic as the case that patients are more than first patients. After winter vacation until January 3rd, students returned to the dormitory and started embarkation training. At that time, there were 56 freshmen on the JCGA training ship. On Friday January 10, a student, the source of infection developed Flu. Finally, 18 students were affected. Developed students are shown by "I" in Table 2 (Case B). According to former study[1], I filled in the state of students such as "S"," E", "P" and "R", as well. F, C, B, Y, I, P, G, CG and BG are the same as in Table 1. Here, base time is 10:00. I found that F and C of student 1 or 2 is very high and G and CG on January 9th or 10th are very high. C is more than B and CG is more than BG. And G on January 12th is small though P is most.

In the two cases, it seems that the source students developed Flu at the very beginning and super-spread the virus. Here, I define super-spreading as that of more than five students directly affected at JCGA, because JCGA is a small academy. I found that the source students affected many students according to the day of the first patients being discovered, hereinafter referred to D-day. After D-day, even though P is large, G is small, few students were affected. At that time, the preventative measures against epidemic students were adapted after D-day, such as wearing medical masks, hand washing and ventilation. The measures seemed to be effective. Moreover, I found that about four days after the consecutive holidays, namely the incubation period passing, the first patients were discovered.

Table 2. Epidemic of Flu on JCGA embarkation training ship in January, 2019 (Case B).

ID	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	F	C	B	Y
1	S	S	E	P	P	**I**	I	I	I	I	I	R	R	R	R	R	R	R	6	6	4	O
2	S	S	S	E	P	P	I	I	I	I	I	R	R	R	R	R	R	R	12	12	7	U
3	S	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	R	R	7	7	4	X
4	S	S	S	S	S	E	P	P	I	I	I	I	I	R	R	R	R	R	2	2	2	X
5	S	S	S	S	S	E	P	P	I	I	I	I	I	R	R	R	R	R	2	2	2	X
6	S	S	S	S	S	E	P	P	I	I	I	R	R	R	R	R	R	R	2	2	2	X
7	S	S	S	S	S	E	P	P	I	I	I	I	I	I	R	R	R	R	2	2	2	X
8	S	S	S	S	S	E	P	P	I	I	I	R	R	R	R	R	R	R	2	2	2	X
9	S	S	S	S	S	S	E	P	P	I	I	I	I	I	R	R	R	R	2	2	2	X
10	S	S	S	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	2	2	2	X
11	S	S	S	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	2	2	2	X
12	S	S	S	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	2	2	2	X
13	S	S	S	S	S	S	E	P	P	I	I	I	I	I	R	R	R	R	2	2	2	X
14	S	S	S	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	2	2	2	X
15	S	S	S	S	S	S	E	P	P	I	I	I	I	R	R	R	R	R	2	2	2	X
16	S	S	S	S	S	S	S	S	E	P	P	I	I	I	R	R	R	R	0	0	0	X
17	S	S	S	S	S	S	S	S	E	P	P	I	I	R	R	R	R	R	0	0	0	X
18	S	S	S	S	S	S	S	S	S	S	S	S	S	S	E	P	P	I	0	0	0	U
I	0	0	0	0	0	**1**	2	3	8	15	15	12	12	4	0	0	0	1				
P	0	0	0	1	2	2	6	12	7	2	2	0	0	0	0	1	1	0				
G	-	-	-	1	5	7	0	2	0	0	0	-	-	-	-	0	0	-				
CG	-	-	-	1	5	7	0	2	0	0	0	-	-	-	-	0	0	-				
BG	-	-	-	1	3	4	0	2	0	0	0	-	-	-	-	0	0	-				

In January 2019, an epidemic of Flu also occurred at JCGA. After winter vacation, students returned to the dormitory and started taking classes. At that time, there were 109 undergraduate students. On Thursday January 9, six students developed Flu. Finally, 13 students, 11 sophomores and 2 juniors developed Flu. Developed students are shown by "I" in Table 3 (Case C). According to a former study[1], I filled in the state of students such as "S"," E", "P" and "R", as well. F, C, B, Y, I, P, G, CG and BG are the same as in Table 1. Here, the base time is 10:00.

Though six students were the source of the infection bringing Flu, super-spreading didn't occur. This case is similar to after D-day in Case A and B. Compared to Case A, I found that some factors were effective.

- No big lecture for all students: students cannot contact many students.
- No martial arts: students cannot contact physically.
- Small number of students at JCGA: students can distance from each other.

Table 3. Epidemic of Flu at JCGA in January, 2019 (Case C).

ID	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	F	C	B	Y
1	S	E	P	P	**I**	I	I	I	I	I	R	R	R	R	R	R	R	R	2	1	1	O
2	S	E	P	P	**I**	I	I	I	I	I	R	R	R	R	R	R	R	R	1	1	0	O
3	S	E	P	P	**I**	I	I	I	I	I	R	R	R	R	R	R	R	R	1	1	0	O
4	S	E	P	P	**I**	I	I	I	I	I	R	R	R	R	R	R	R	R	1	1	0	O
5	S	E	P	P	**I**	I	I	I	I	I	R	R	R	R	R	R	R	R	1	1	0	O
6	S	E	P	P	**I**	I	I	I	I	I	R	R	R	R	R	R	R	R	1	1	0	O
7	S	S	E	P	P	I	I	I	I	I	I	R	R	R	R	R	R	R	0	0	0	O
8	S	S	E	P	P	I	I	I	I	I	I	R	R	R	R	R	R	R	0	0	0	O
9	S	S	E	P	P	I	I	I	I	I	R	R	R	R	R	R	R	R	0	0	0	O
10	S	S	S	E	P	P	I	I	I	I	I	I	R	R	R	R	R	R	0	0	0	X
11	S	S	S	E	P	P	I	I	I	I	I	I	R	R	R	R	R	R	0	0	0	X
12	S	S	S	S	S	S	S	E	P	P	I	I	I	I	I	I	R	R	0	0	0	U
13	S	S	S	S	S	S	S	S	E	P	P	I	I	I	I	I	I	R	0	0	0	O
I	0	0	0	0	**6**	9	11	11	11	11	5	4	2	2	2	2	1	0				
P	0	0	6	9	5	2	0	0	1	2	1	0	0	0	0	0	0	0				
G	-	-	2	0	0	0	-	-	0	0	0	-	-	-	-	-	-	-				
CG	-	-	1	0	0	0	-	-	0	0	0	-	-	-	-	-	-	-				
BG	-	-	1	0	0	0	-	-	0	0	0	-	-	-	-	-	-	-				

6. Infectivity Rate

At JCGA, infectivity rate β by "I" is 0 because patients are isolated in a sick room. Then, I calculated infectivity rate α by "P" in classroom and bedroom, as for by D-day (α_bC and α_bB) and after D-day (α_aC and α_aB)(Table 4). I found that the average infectivity rate after D-day is smaller than that of D-day. Especially the average infectivity rate in the classroom after D-day (α_aC) is one twentieth of that of D-day (α_bC). Considering that most students are affected in classrooms, that is a significant impact on the epidemic. The measures against Flu seem to be effective after D-day. It seems that Flu has finished super-spreading by D-day.

I show the average infectivity rate in the former study in Table 4, as well[1]. The standard deviation of the infectivity rate after D-day (α_aC and α_aB) is smaller than that of the former study. They approach real data more than the former study after D-day. While, those by D-day (α_bC and α_bB) have a very wide range, it is possible to provoke a pandemic by D-day.

Table 4. Infectivity rate by "P" of JCGA 2017 Case .

Infectivity rate	Classrooms		Bedrooms	
This study	by D-day α_bC	after D-day α_aC	by D-day α_bB	after D-day α_aB
Range	0 – 0.32	0 – 0.0185	0 - 0.286	0 - 0.167
The Average	0.0423	0.00190	0.0391	0.0152
Standard Deviation	0.0886	0.00489	0.0853	0.0430
Former study[1]	α_C		α_B	
The Average	0.0206		0.0259	
Standard Deviation	0.0637		0.0660	

7. Conclusion

B y examining affection by pre-infectious students in real data, I found that super-spreading depends on the timing of penetration. I show the students that super-spread seasonal influenza according to the day the first patients are discovered. According to the data, on average it takes for days for seasonal influenza measures to be implemented within the communities.

References

1. S. Iwanaga, K. Kawaguchi, "Analysis of epidemic of Flu in closed space" proceedings of the 22nd Asia Pacific Symposium on Intelligent and Evolutionary Systems, pp. 37-44, 2018.
2. CDC, "Severe Acute Respiratory Syndrome".
3. W.O. Kermack, A.G. McKendrick, "A contribution to the mathematical theory of epidemics", proceedings of the Royal Society of London. Series A, Vol. 115, No. 772, pp. 700-721, 1927.
4. M.J. Keeling, P. Rohani, "Modeling Infectious Diseases in Humans and Animals", Princeton University Press, 2008.

Authors Introduction

Dr. Saori Iwanaga

She is Professor of Department of Maritime Safety Technology at Japan Coast Guard Academy. She received her B.E. degree in applied chemistry engineering from Utsunomiya University and her M.S. and Ph.D. degree in computer science from National Defense Academy, Japan. She is interested in complex theory, evolutionary games.

Towards the Trusted Population-Based Optimization Systems

Hiroshi Sato
Department of Computer Science, National Defense Academy
1-10-20 Hashirimizu, Yokosuka, Kanagawa 239-8686, Japan

Masao Kubo
Department of Computer Science, National Defense Academy
1-10-20 Hashirimizu, Yokosuka, Kanagawa 239-8686, Japan
E-mail: hsato@nda.ac.jp, masaok@nda.ac.jp
https://www.nda.ac.jp/~hsato

Abstract

Following the development of evolutionary computation, various population-based optimization methods have been proposed. In these systems, optimization is achieved through the interactions of many individuals/particles/agents. However, when the system is implemented in a distributed environment, reliability becomes an issue. In such an environment, it may not be possible to trust others. There are numerous cases which we cannot guarantee trust, such as malfunction of distributed parts or failure to synchronize. Therefore, we have to make trust between distributed individuals/particles/agents. The record of past actions is usually a good tool for generating trust. This paper introduces the blockchain mechanism into the population-based optimization system to make a trust management system. By using blockchain, we can implement it without a central authority. In the system, all interactions are reviewed and get feedback, and the feedback is used to calculate the trust score.

Keywords: Trusted system, Blockchain, Surrogate Assisted Evolutionary Computation.

1. Introduction

So far, various population-based optimization methods have been devised. For example, Genetic Algorithm[1], Genetic Programming[2], Evolutionary Strategies[3], and Evolutionary Programming[4] are the pioneers in this field. Following the success of evolutionary computation, a lot of other population-based algorithms have been devised such as Particle Swarm Optimization[5], Ant Colony Optimization[6], Artificial Immune System[7]. In these systems, optimization is achieved through the interactions of many solution candidates. They are called individuals, particles, and agents. We use the word "individuals" to point a solution candidate from now on.

When the system is implemented in a distributed computational environment, reliability becomes an issue. In such an environment, it may not be possible to trust others. There are numerous cases which we cannot guarantee trust in individuals, such as malfunction of distributed component, failure to synchronize the information, or injection of malicious individual. Therefore, we have to make trust between distributed individuals. In these cases, the record of past actions is usually a good tool for generating trust.

This paper introduces the blockchain mechanism into the population-based optimization system to make a trust management system. We adopt evolutionary computation as a reference model. By using blockchain, we can implement it without a central authority. In the system, all interactions are reviewed and get feedback, and the feedback is used to calculate the trust score.

2. Blockchain

A blockchain[8] is a list of records, called blocks, linked together using cryptography. Each block contains a cryptographic hash of the previous block, a timestamp, and transaction data. Fig. 1 shows the example of usage blockchain in evolutionary computation.

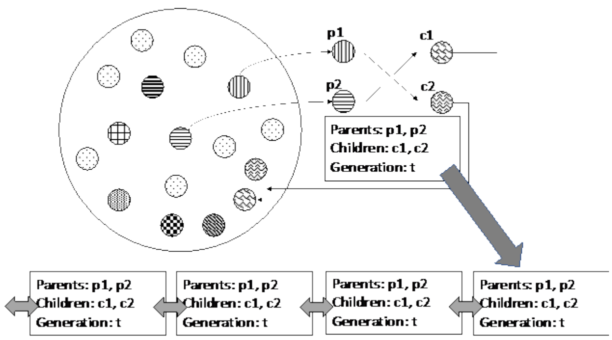


Fig. 1. The usage of blockchain in evolutionary computation.

When blockchains are used as a distributed ledger, they are usually managed by a peer-to-peer network and conform to protocols for inter-node communication and verification of new blocks. Once a block's data has been recorded, it cannot be changed retroactively without changing all subsequent blocks. For this reason, blockchain is considered secure by design and is an example of a decentralized computing system with high Byzantine fault tolerance. These make decentralized consensus a key concept in the blockchain. We use blockchain as a tool of maintaining trust.

Bitcoin[8] and Ethereum[9] are two of the most popular blockchains. While Bitcoin is a book of currency, Ethereum is a book of programs. In Ethereum, any computer program can be put on the ledger, which has attracted worldwide attention as it enables smart contracts, decentralized finance, and decentralized exchanges. Therefore, we focus on Ethereum in this study.

Ethereum is a platform for building decentralized applications and smart contracts, and the generic name of a related open-source software project being developed by the Ethereum Project. Ether is used as the internal currency required to use Ethereum. Ethereum is designed as a general-purpose computer and can run a virtual machine.

There are two consensus algorithms for Ethereum: one is for Proof of Work (POW), called “Ethash,” and the other is for Proof of Stake (POS), called “Casper.”

3. Trust in Evolutionary Computation with Blockchain

This paper concerns the reliability of each individual's information in distributed evolutionary computation. As noted in the chapter 1, the information may not be reliable in distributed environment for some reason. For example, when the computation is curried over the distributed machines, some machines may work differently from the rest by the failure or by malicious action. Moreover, the fitness information will be vague even in a single machine when the system uses a surrogate mechanism. Therefore, we have to estimate how the other individual can be trusted.

Let us assume individuals in evolutionary computation. An individual wants to know the fitness value of other individuals to produce good offspring. In usual evolutionary computation, the fitness value is assumed to be correct. However, we think distributed environment. In this case, the individual has to estimate the fitness through the record of other individuals’ actions. The individual has to decide which individual to trust. The fitness value provided by different individuals may differ. For instance, one may offer a quick answer at a lower quality while another may be slow but accurate.

While the individual will be confident of the validity of their previous interactions with the other individuals, they cannot rely on their knowledge to provide certainty in other individuals’ interactions. We can solve this problem by storing the verified feedback of the record of interaction on the blockchain. Such feedback can be accessed by any trust provider, which offers trust scores as a service. When we use blockchain, the information is available to all parties. This means that the information and trust scoring mechanisms have the following properties: Universal, Transparent, and Verifiable. As an added benefit, the integration of blockchain into the system enables the payment for resource access, including trust score estimation.

We take a quantitative approach to reason about trust, using the information which are built from the feedback of interaction between individuals. The trust calculation is done by direct experiences by aggregating individual feedback scores to form an overall individual opinion about the quality of interaction or reliability of other individuals. Sometimes, direct experience may not be possible when no interaction may have occurred between individuals. In this case, the individual would rely on third-person's information to infer information of other individuals.

4. System Architecture

Fig. 2 shows the architecture of our proposed system for evolutionary computation. This system is inspired Pal's work[10].

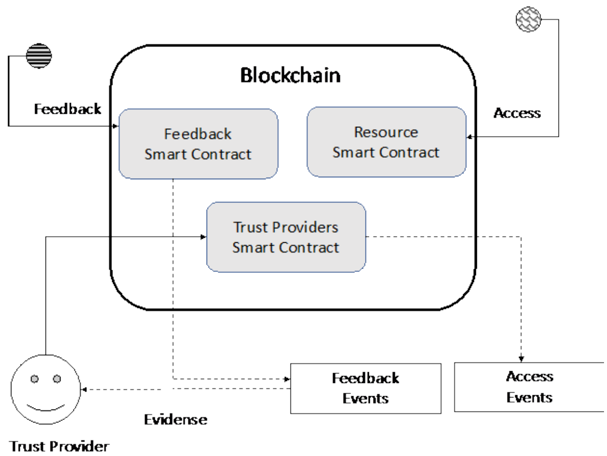


Fig. 2. The architecture of proposed blockchain platform for evolutionary computation.

There are three main components in this architecture, individuals, trust providers, and smart contracts. In the system, individuals can be both information providers and consumers. Individuals can store information, access a resource, deploy smart contracts, and communicate with one another. Trust providers maintain the trust scores. Smart contacts are collections of code and data used to execute agreements between two individuals and stored on a blockchain. We use three types of smart contracts:

- Resource smart contract: that handles access to a resource,
- Feedback smart contract: that handles the reviews submitted by the individuals,
- Trust provider's smart contract: that helps the trust providers to maintain trust scores.

The system is composed of a public blockchain that keeps track of all delegated access rights, consumer interactions, and consumer feedback that is directly linked to one consumer interaction.

The smart contract handles reviews submitted by individuals in the system. It receives a review rating and ensures that the review is linked to interaction. It ensures that a submitted review has the following parts:

- Address of the individual that submits the review,
- Details of the interaction reviewed, and
- A review rating.

The correctness of the system can be verified by all individuals interacting with the public blockchain.

Trust provider is responsible for the trust scoring functions and making the output available to the individuals for some access fee. The trust provider complements its soundness by choosing a scoring mechanism and an evidence selection to implement.

The communication between the components of the system proceed as follows:

1. When an individual wants to check the information of other individuals, he asks the trust provider.
2. The trust provider retrieves a pre-computed score or performs an on-demand trust score calculation.
3. Once the resource has been used an event will be generated on the blockchain.
4. The trust provider will be notified of the new resource access since the event is broadcast on the public chain.
5. The trust providers will then update the feedback smart contract on the blockchain to update the feedback state.

6. If the individual is willing, they can leave feedback for that resource.

5. Conclusion

This paper proposed a framework for trust systems for evolutionary computation where the record of interactions backs up evidence. This paper introduces the blockchain mechanism into the population-based optimization system to make a trust management system. By using blockchain, we can implement it without a central authority. In the system, all interactions are reviewed and get feedback, and the feedback is used to calculate the trust score. We consider several scoring methods for this type of system and averaging approach is simple yet powerful. As future works, we will implement the framework using Ethereum, and a feasibility study should be conducted.

References

1. D. E. Goldberg, "Genetic Algorithms in Search, Optimization and Machine Learning", Addison-Wesley, 1989.
2. J. R. Koza, "Genetic Programming, On the Programming of Computers by Means of Natural Selection, MIT Press, 1992.
3. H. G. Beyer and H. P. Schwefel, "Evolution Strategies: A Comprehensive Introduction", Natural Computing, 1, pp.3 - 52, 2002.
4. D. B. Fogel, "Artificial Intelligence through Simulated Evolution", in Evolutionary Computation: The Fossil Record, IEEE, pp.227 - 296, 1998.
5. M. R. Bonyadi, Z. Michalewicz, "Particle swarm optimization for single objective continuous space problems: a review", Evolutionary Computation, Vol. 25 No. 1: pp. 1 - 54, 2017.
6. M. Dorigo, M. Birattari and T. Stutzle, "Ant colony optimization", IEEE Computational Intelligence Magazine, Vol. 1, No. 4, pp. 28 - 39, 2006.
7. D. Dasgupta (Ed.), "Artificial Immune Systems and Their Applications", Springer-Verlag, Inc. Berlin, 1999.
8. S. Brotsis, K. Limniotis, G. Bendiab, N. Kolokotronis, S. Shiaeles, "On the suitability of blockchain platforms for IoT applications: Architectures, security, privacy, and performance", Computer Networks, Vol. 191, 108005, 2021.
9. Ethereum Project, https://ethereum.org/, (viewed on the 20th Dec. 2021).
10. S. Pal, A. Hill, T. Rabehaja, M. Hitchens, "A blockchain-based trust management framework with verifiable interactions", Computer Networks, Volume 200, 108506, 2021.

Authors Introduction

Dr. Hiroshi Sato

SHe is an Associate Professor of the Department of Computer Science at the National Defense Academy in Japan. He obtained a degree in Physics from Keio University in Japan and degrees of Master and Doctor of Engineering from Tokyo Institute of Technology in Japan. He was previously a Research Associate at the Department of Mathematics and Information Sciences at Osaka Prefecture University in Japan. His research interests include agent-based simulation, evolutionary computation, and artificial intelligence. Dr. Sato is a member of the Japanese Society for Artificial Intelligence (JSAI), Society of Instrument and Control Engineers (SICE), and The Institute of Electronics, Information and Communication Engineers. IEICE). He was the editor of IEICE and SICE.

Dr. Masao Kubo

He is Associate Professor of Department of Computer Science at National Defense Academy in Japan. He graduated from the precision engineering department, Hokkaido University, in 1991. He received his Ph.D. degree in Computer Science from the Hokkaido University in 1996 (multi-agent system). He had been the research assistant of the chaotic engineering Lab, Hokkaido university. He was a visiting research fellow of Intelligent Autonomous Lab, university of the west of England (2005). He is the associate professor of information system lab, Dep. of Computer Science, National Defense Academy, Japan. His research interest is Multi agent system.

Spatio-temporal prediction of crime occurrence spots

Kaede Yaji, Masao Kubo, Hiroshi Sato

Information Engineering, National Defense Academy in Japan, Hashirimizu 1-10-20
Yokosuka, Kanagawa 239-0811, Japan

E-mail: em60041@nda.ac.jp, masaok@nda.ac.jp, hsato@nda.ac.jp
www.mod.go.jp/nda/english/

Abstract

This paper proposes a method for spatiotemporal prediction of crime occurrence locations based on previous data. In recent years, Japanese government has begun to release data on crime occurrences to improve the efficiency of policing. In addition, the development of maps that can manage patrol and assist residents' crime prevention has been planned. For statistical crime prediction, while several methods are invented abroad, it has just begun to develop a specific crime prediction model for a low-crime country, Japan. One of the known methods uses LSTM to predict crime occurrences only from a temporal perspective, but it cannot predict points of crime occurrences and is insufficient to generate a map. Therefore, we propose a method that combines this LSTM based method with CNN that can adopt geographic locations. As a result of computer experiments, this method seems to be able to make predictions with a tendency to capture actual characteristics.

Keywords: Crime occurrence, Spatio-temporal prediction, LSTM, CNN

1. Introduction

This study examines the method for predicting crime locations in time and space using information on crime occurrences opened by the police to create maps that will assist police patrols and serve as crime prevention for residents. In Japan, the Basic Act on the Promotion of Government and Private Data Utilization (Act No. 103 of 2016) was enacted, and through the "Promotion of the Release the Data on Crime Occurrences Information", each prefectural police have started to release data on crime Occurrences. The purpose of this attempt is to share and utilize data on crime to improve the efficiency of police activities and prevent crimes by predicting their occurrences. In this paper, we propose a method that combines Long Short Term Memory (LSTM) and Convolutional Neural Network (CNN) to address this issue. There are studies based on the idea that the prediction needs a unique model in Japan, that is a low-crime country [2]. Among them, Method [1] predicts crimes in the same area using a neural network with LSTM layers, and it has achieved some success. Therefore, we came up with the idea that if we could combine CNN that enables spatial predictions with LSTM, we could predict crime spatiotemporally. Based on this idea, this paper proposes a method that modifies the output layer of the method created for weather predictions with ConvLSTM2D layers [4]. We verified it by comparing data from Kanagawa Prefecture and Osaka Prefecture to examine the usefulness of the proposed method. We created crime map data by dividing the information on bicycle thefts in

Kanagawa Prefecture in 2018 into monthly grid data. We used this to learn the data from January to November and predicted the data for December, and were able to create crime map data with a trend similar to the actual December data.

2. Related Studies

This section describes recent mathematical-based crime prediction in Japan and shows that there are few studies on geographic crime prediction. According to Oyama et al. (2017) [3], researches and system developments on geographical crime predictions have been conducted mainly in Europe and the United States, but there is no accumulation of researches on these methods in Japan. Therefore, it is necessary to study the prediction methods suitable for Japan. The occurrence locations and times are important in predicting crime. Existing geographical prediction methods can be classified into four categories: detections of temporal clusters, intensity estimations of crime occurrences considering spatiotemporal interaction of crimes, predictions of risks of crime occurrences from environmental factors, and prediction of the number or probability of crime occurrences by regression analysis [3]. In both cases, the prediction is conducted by setting up a model in advance and adjusting a relatively small number of parameters. Detections of temporal clusters use spatiotemporal scan statistics to find spatiotemporal clusters of crimes since the accumulation of past crimes predicts future crime occurrences. Intensity estimations of crime occurrences considering the spatiotemporal interaction of crimes focus on the near-repeat victimization and obtain the risk of crime occurrences by modeling that considers temporal proximity in addition to spatial proximity. ProMap and SEPP are typical methods. Prediction of the risk of crime occurrences from environmental factors does not rely on past information on crime occurrences but focuses on the socio-economic and physical characteristics of the region, seasonal variations, etc. The typical method is RTM using overlay analysis in GIS. Prediction of the number or probability of crime occurrences by regression analysis is a method in which regression analysis is not used to analyze past crime factors, but to predict the number of crime occurrences and crime rates in the future. ST-GAM, which also uses environmental factors as variables, is a typical example. Whereas, there are a few methods that do not assume a model and make predictions from criminal records. Zhang et al. (2019) focus only on the temporal proximity of crimes and use LSTM to predict the number of crime occurrences. LSTM is a neuronal network layer that may acquire data-specific time series models through training. In the paper, they extracted data on sneak-thieving, bicycle theft, and shoplifting from the Tokyo Metropolitan Government's open data on crime recognition (177131 crimes from 2014 to 2018). This work, used LSTM to predict the number of crime occurrences in the following month based on the number of same crime occurrences in the past six months. As a result, we have obtained some success, but there is still room for improvement in the prediction accuracy. Therefore, to improve this, we performed a verification based on the idea that the addition of convolutional neural networks (CNN) to LSTM would permit spatiotemporal predictions. CNN is a neural network developed for advanced image recognition, which can obtain a model about the positional relation specific to the data.

3. The Proposed Method

In this section, we explain the proposed method. This method combines LSTM, which handles time series data, with CNN, which can easily use geographic data, to achieve prediction based on geographic relevance, which has been a problem of the previous method [1]. Fig. 1. shows the proposed model. It is a three-layer coupling of ConvLSTM blocks, which LSTM block plus a convolutional layer, via Batch Normalization. The output function was set to Relu. The data used is crime map data, which is created by counting the number of crime occurrences in each grid cell of a certain size, divided by month, and the model predicts and outputs the crime map data for the month following the input data. We use 11 months as input and the next month as the correct answer for learning. Specifically, we take a span between Jan. 2018 to Nov. 2018 as input and set Dec. 2018 as the correct answer, then shift the data by one month until the span between Jan. 2020 and Nov. 2020 as input and set Dec. 2020 as the label of a training data, which is one epoch. We used Adam as the optimizer and MSE as the loss function.

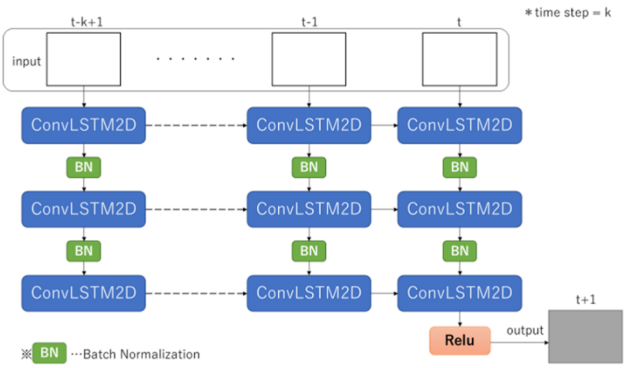


Fig. 1. Proposed Model.

An example of the learning process is shown in Fig. 2. This shows the learning process of the proposed method for a bicycle theft that occurred in Kanagawa Prefecture in 2018 (11,388 cases). The horizontal axis is the number of epochs and the vertical axis is Loss (MSE). At first, there was a major error, but in the 14th epoch, the error decreased to less than 0.3, indicating that learning had progressed.

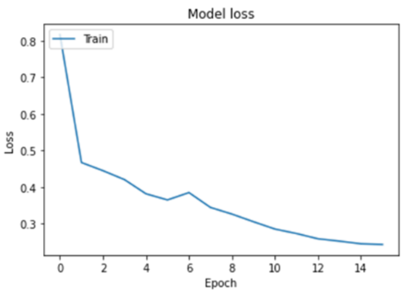


Fig. 2. Model Loss.

4. Experiments

In this section, we describe two experiments. In Experiment 1, we verify the crime prediction using the proposed method. In Experiment 2, we examine whether the proposed method can predict crime in any region of Japan using data from a different region than in Experiment 1.

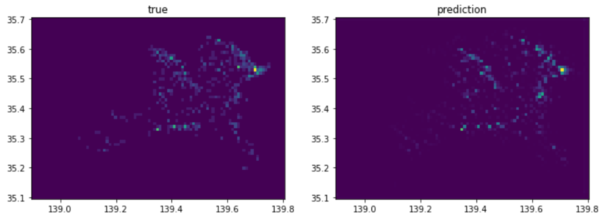


Fig. 3. Experiment 1. Kanagawa’s 2018 bicycle thefts data

4.1. *Experiment 1: Kanagawa’s 2018 bicycle thefts data*

In this experiment, we predicted the December by learning the period from January to November in the data of Kanagawa’s 2018 bicycle thefts. The left of Fig. 3. shows the actual number of this crime of this month, the right illustrates the predicted number of this crime. Comparing them, we could see that the hotspots (the areas close to yellow in the figure) generally match, indicating that the predicted crime map captures the characteristics of the actual data.

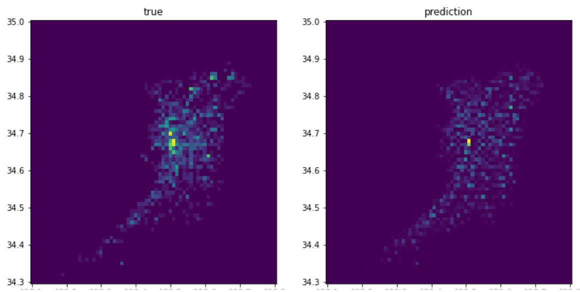


Fig. 4. Experiment 2. Osaka's 2018 bicycle thefts data

4.2. *Experiment 2: Osaka's 2018 bicycle thefts data*

In Experiment 2, we used Osaka's 2018 bicycle thefts data to make predictions in the same way as Experiment 1, and the results are shown in Fig. 4. Comparing the actual crime raster data with the predicted one, we can say that the hotspots are generally captured the same as Experiment 1, although the prediction is slightly rougher overall.

5. Conclusion

In this paper, we proposed a method to predict crime spatially and temporally by combining CNN with the temporal crime prediction method using LSTM in existing research based on the idea that a unique model is necessary to predict crime in Japan with a low-crime rate. We examined this method using three years of bicycle thefts data in Japan, and found that the predictions were generally good, indicating that the proposed method is effective in predicting crime in Japan.

References

1. Xiaoqi Zhang, et al. "Study on Forecasting Crime and its Temporal Proximity Using Open Data and LSTMs", journals of Informatics, vol16.1, pp56-70, 2019.
2. Tomoya Oyama, et al. "A Metric Comparison of Geographical Crime Prediction Techniques — An Examination for Developing the Crime Prediction Technique for Japan—", 2017.
3. Tomoya Oyama, et al. "Recent Research Trends on Geographical Crime Prediction", Theory and applications of GIS, 25(1), pp. 33-43, 2017.
4. https://qiita.com/eycjur/items/85baee1c397b74c3b6bd

Authors Introduction

Miss. Kaede Yaji

She is currently a student in National Defense Academy in Japan.

Dr. Masao Kubo

He is an Associate Professor of the Department of Computer Science at National Defense Academy in Japan. He graduated from the precision engineering department, Hokkaido University. He received his Ph.D. degree in Computer Science from the Hokkaido University.

Dr. Hiroshi Sato

He is an Associate Professor of the Department of Computer Science at the National Defense Academy in Japan. He obtained a degree in Physics from Keio University in Japan and degrees of Master and Doctor of Engineering from Tokyo Institute of Technology in Japan.

Cross-view Image Geo-Localization using Multi-Scale Generalized Pooling with Attention Mechanism

Duc Viet Bui[*]
Department of Computer Science, National Defense Academy, 1-10-20 Hashirimizu Yokosuka City, Kanagawa Prefecture, Japan[†]

Masao Kubo
Department of Computer Science, National Defense Academy, 1-10-20 Hashirimizu Yokosuka City, Kanagawa Prefecture, Japan

Hiroshi Sato
Department of Computer Science, National Defense Academy, 1-10-20 Hashirimizu Yokosuka City, Kanagawa Prefecture, Japan
E-mail: ed21009@nda.ac.jp, masaok@nda.ac.jp, hsato@nda.ac.jp

Abstract

Cross-view image matching for geo-localization is the task of finding images containing the same geographic target across different platforms. This task has drawn significant attention due to its vast applications in UAV's self-localization and navigation. Given a query image from UAV-view, a matching model can find the same geo-referenced satellite image from the database, which can be used later to precisely locate the UAV's current position. Many studies have achieved high accuracy on existing datasets, but they can be further improved by combining different feature processing methods. Inspired by previous studies, in this paper, we proposed a new strategy by using a channel-based attention mechanism with a generalized mean pooling method to enhance the feature extracting process, which improved accuracy.

Keywords: Cross-view image matching, UAV, attention mechanism, generalized mean pooling.

1. Introduction

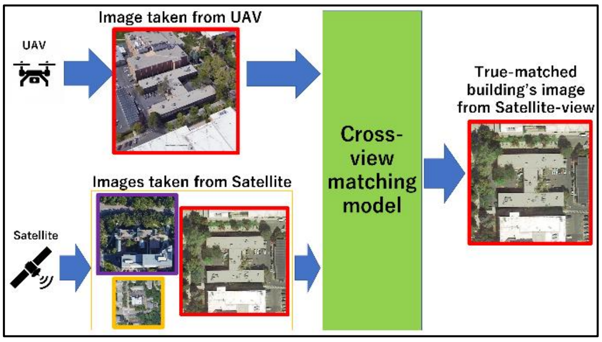
The applications of unmanned aerial vehicles (UAV) in daily life have been rapidly increasing. UAV has become an essential part of various fields such as aerial surveillance[1], agriculture[2], transportation, search and rescue missions. Along with their applications, to further reduce human's work, the need for autonomous drones has been increasing over time. However, most researches failed to achieve a fully autonomous drone system, as the most used navigation system (Global Positioning System - GPS) has many limitations in real-life missions. For example, GPS is not powerful enough when high buildings or trees block GPS signals, leading to difficulties in applying UAVs in cities and urban areas.

Many solutions for navigation in autonomous drone systems have been proposed, and among them, cross-view image matching-based methods have received lots of researchers' attention due to the vast application value in geo-localization[3,4,5,6)]. Cross-view image matching is the task of matching a satellite-view image with

.

geographic location tags and a UAV-view image without a geographic location tag or vice versa to locate a UAV's position based on information from taken images. Figure 1 shows an example of cross-view matching methods. Given a UAV-view image of a building, the matching model searches for the image of that building in the satellite-view image gallery. The output is a satellite-view image similar to the query UAV-view image. This output can be used to locate the current position of the UAV.

Fig. 1: Example of UAV-view → Satellite-view



Early-stage cross-view image matching researches[7] focused on using traditional image processing methods, which used hand-crafted features from images. In recent years, with the rapid development of machine learning and deep learning in image processing, especially the convolutional neural network, many studies attempted to apply them in cross-view image matching for geo-localization, some of which have achieved significant results[3,6)]. Also, deep learning's well-known self-attention mechanisms have been used to understand image representations further and bridge the gap between images from different views[8]. However, the self-attention mechanisms usually require lots of computational costs, which is entirely unsuitable for current UAV systems. Moreover, pooling strategies in the previous cross-view matching method were mainly average pooling and max pooling, limiting the feature learning process from extracting important global features.

Inspired by cross-view image matching related works, in this paper, we ensembled several feature processing methods with a channel-based attention mechanism and multi-scale generalized pooling strategy. Our proposed model has shown an increase in performance through experiments compared to the state-of-the-art method (SOTA).

2. Related works

The previous studies[9,10)] often consider cross-view image matching as an image retrieval problem since they aim to find similar images to a query image among an image dataset. In the cross-view matching problem, the main task is to learn image representation that varies in different views, thus bridging the gap between multi-view images. The schematic diagram of cross-view matching is described in Figure 2.

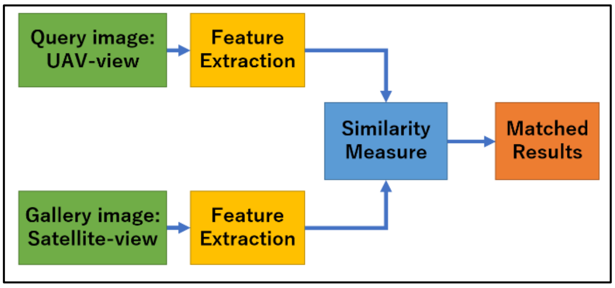


Fig. 2: The schematic diagram of cross-view image matching problem (UAV-view → Satellite-view)

At first, features from query images and a database (gallery images) are extracted using different feature processing methods. After that, features' similarities were calculated by using distance similarity measures such as cosine similarity or Euclidean distance. The results are later used to create a ranking list, from which the model will determine the true-matched image to the query image.

Traditional image processing methods such as SIFT[11] or SUFT[12] have been implemented in early research for the feature extracting phase. However, as the gap between different viewpoints is enormous, the matching results are not high as expected. As a result, more and more researchers have been paying attention to the powerful convolutional neural networks (CNN), which is well-known for its abilities to learn high-level features. Another work focused on learning discriminative features using metric learning and proposing various loss functions. This line of work has similar approaches with face verification and person re-id problems cause they adopt ranking losses such as contrastive loss or triplet loss to learn the relative distances between different inputs. For example, Lin et al.[13] adopted contrastive loss to optimize network parameters, and Hu et al.[14] designed

CVM-Net, which employed weighted-soft-margin ranking loss.

Moreover, some works try to enhance the feature learning phase by applying the attention mechanism. Attention-mechanism is a method invented to make neural networks learn the most relevant features from inputs, thus increasing the network's learning abilities. Novel self-attention model – Transformer[15] has been studied in many natural language processing problems, and now its application in vision processing (known as Vision Transformer[16]) has been adopted in numerous cross-view matching researches[8]. However, the disadvantages of Transformer architecture are high computational cost and a massive amount of data required for training.

Generalized mean pooling (GeM pooling) was first proposed in Ref. 17 as an alternative pooling method for image retrieval. Since then, it has been widely applied in many retrieval systems and achieved promising results.

Following related works on cross-view image matching, the architecture of our proposed model was based on Siamese networks (twin neural network), with two branches for each view's input, and we investigated the effectiveness of the channel-based attention with GeM pooling towards the cross-view matching problem.

3. Materials and the proposed method

3.1. Dataset and evaluation metrics

In this work, we use the University-1652 dataset published by Zheng et al.[18], as it is the only benchmark dataset with both satellite-view and UAV-view images, which helps solve cross-view matching for UAV navigation. This dataset contains 1652 geographic targets from 72 universities all over the world. Each target contains three views: satellite-view, UAV-view, and street-view. To reduce the high cost in airspace control and flying UAV, all UAV-view and street-view images were collected by a 3D engine named Google Earth, while satellite-view images were captured by Google Map. All images in the dataset have geo-tags as their class labels. The view of UAVs in Google Earth was controlled by simulated camera-view, and the height of view descends from 256 to 121.5m. Each target consisted of 1 satellite-view image, 54 UAV-view images, and a few street-view images. The dataset was split into the training and test sets with no overlapped classes. The captured images have an original size of 512x512. The distribution of data in each set is described in Table 1.

Table 1. Distribution of data in University-1652

	Images	**Class**	**Universities**
Training	50218	701	33
Query (UAV)	37855	701	39
Query (Satellite)	701	701	
Query (Ground)	2579	701	
Gallery (UAV)	51355	701	
Gallery (Satellite)	951	951	
Gallery (Ground)	2921	793	

Regarding the evaluation metrics, most of the image retrieval and cross-view image matching researches has been using Recall@K and Average Precision (AP) as the main indicator for evaluating proposed systems. Recall@K is computed by calculating the ratio of the true-matched image in the top-K results of the ranking list. On the other hand, AP is a popular metric in measuring the precision of a retrieval system. The higher Recall@K and AP, the better the model performs.

3.2. Proposed method

The overview of proposed network architecture is described in Figure 3. From subsection 3.2.1 to 3.2.3, we explain the major components of the model.

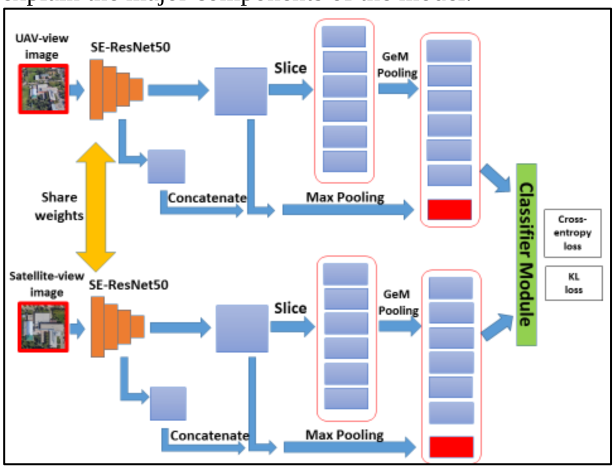


Fig. 3: Proposed network architecture

3.2.1. *Feature extraction strategy*

Because of excellent accuracy and inference time, other existing methods used backbone from ResNet50[19] model or VGG16[20] as the main feature extractor, while the usage of attention modules in these backbones is rarely seen. However, we believe that an attention mechanism can strengthen the saliency value of each view and restrain the unnecessary features from affecting the final results. Therefore, among various attention mechanisms in literature, channel-based attention - the SE block from Ref.21 is chosen for its efficiency in reinforcing the backbone while requiring very little computation cost. The SE block can also be easily implemented in ResNet50's Residual block as follow (Figure 4).

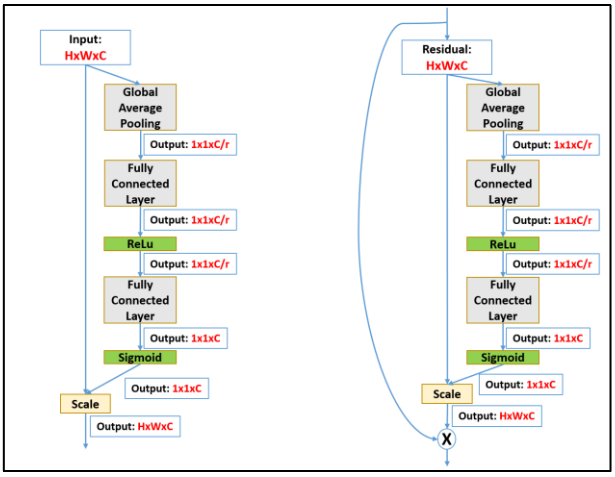


Fig. 4: SE block (left) and its implementation in Residual block (right)

Deep layers in CNN tend to learn high-level features (object, human), while shallow layers extract low-level features such as shapes, edges in the image. Geographic targets in the dataset are mainly buildings, so these low-level features play an important role in understanding the entire view. For that reason, in our proposed model, not only the last layer of the model but the global features from shallow layers (here, we chose the third and the fourth layers) are also extracted. Feature maps were concatenated, and Global max pooling method was applied after that.

3.2.2. *Multi-scale block and pooling strategy*

Previous works in Ref. 22 and Ref. 23 proposed a feature partition strategy to take advantage of contextual information. In particular, the output feature map is divided into several blocks, and then the global average pooling method is performed. Here we also applied the block strategy, but GeM pooling was put in practice instead of global average pooling. The formula of GeM pooling can be defined as follow:

$$f^{(g)} = [f_1^{(g)} .. f_k^{(g)} .. f_K^{(g)}]^T, f_k^{(g)} = \left(\frac{1}{|X_k|}\sum_{x \epsilon X_k} x^{p_k}\right)^{\frac{1}{p_k}} \quad (1)$$

with X_k represents feature map, K is the number of channel and p_k is the pooling parameter. This pooling parameter can be manually set or changed through learning process.

The way of creating multi-scale feature blocks is described in Figure 5:

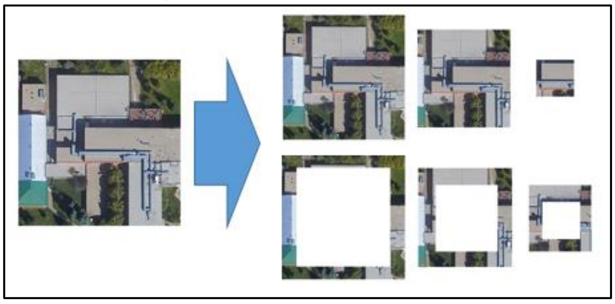


Fig. 5: Multi-scale block partition

3.2.3. *Loss functions*

In Ref. 18, instance loss was adopted to train the multi-branches networks, and results have shown the good effect of this loss function compared to other ranking losses in terms of cross-view matching accuracy. Instance loss was first proposed in Ref. 24, an alternative way to learn the distance between features. Extracted features from each branch were sent to the shared fully connected layer in order to map the features of all sources into one shared feature space. Finally, the cross-entropy loss function was applied to optimize the network.

Additionally, in Ref. 23, Kullback-Leibler divergence (KL divergence) was first applied in the training phase. In the fields of machine learning, KL divergence is a measure of how a probability distribution differs from another probability distribution. KL divergence is expected to close the gap between two different domains (UAV and satellite).

Here we use the Softmax function to obtain the normalized probability scores, and KL divergence is then computed and added to the training loss. The KL divergence formula is defined as follow:

$$L_{KL(p_2||p_1)} = \sum_{n=1}^{N} p_2^n log \frac{p_2^n}{p_1^n} \quad (2)$$

p_1, p_2 are predicted results of each branch.

4. Experiments and discussions

4.1. Experiments

4.1.1. *Implementation details*

We performed training our proposed model with the University-1652 dataset and some ablation experiments with different backbones and pooling layers to understand the effect of attention mechanism and pooling strategy towards the results.

The experiments were carried out with three different input sizes (256x256, 384x384, 512x512). Random flipping and random cropping were used to augment the training images. SE-ResNet50 was pre-trained with the ImageNet dataset, and the stride of the final down-sampling layer was fine-tuned from 2 to 1 to increase the size of the feature map output by the backbone. The optimizer was Stochastic Gradient Descend (SGD) with a momentum of 0.9 and weight decay of $5\text{x}10^{-4}$. The training lasted for 120 epochs, with an initial learning rate of $1\text{x}10^{-4}$ for backbone layers and $1\text{x}10^{-3}$ for other layers. The pooling parameter p_k in GeM pooling was initially set to 3.

The classifier layer was removed in the testing phase, so the model returns only extracted features as outputs. Euclidean distance was applied to compute the similarity between feature vectors from different views.

4.1.2. *Experiment Results*

In Table 2, we compared our proposed method with the SOTA in Ref. 23. In Table 3, we showed the results of the ablation experiments on SE block and pooling strategy.

Table. 2 Comparison of proposed method with SOTA

Method	UAV → Satellite		Satellite → UAV	
	R@1	**AP**	**R@1**	**AP**
MSBA (256x256)	82.33	84.78	90.58	81.61
MSBA (384x384)	86.61	88.55	92.15	84.45
MSBA (512x512)	86.69	88.66	92.01	84.45
Ours (256x256)	**82.87**	**85.13**	**90.87**	**82.06**
Ours (384x384)	**86.96**	**88.88**	**92.30**	**84.92**
Ours (512x512)	**87.90**	**89.71**	**92.58**	**85.49**

Table. 3 Results of ablation experiments of using different strategies

Method	UAV → Satellite		Satellite → UAV	
	R@1	**AP**	**R@1**	**AP**
ResNet50 + Average Pooling (256x256)	78.99	81.85	87.30	78.18
ResNet50 + GeM Pooling (256x256)	79.58	82.22	86.16	78.56
SE-ResNet50 + Average Pooling (256x256)	81.44	84.09	90.16	79.44
SE-ResNet50 + GeM Pooling (256x256)	82.87	85.31	90.87	82.06

4.2. *Discussions*

From the results of the experiments shown in Table 2, it can be seen that our proposed model has an increase in terms of accuracy. The improvement in 256x256 and 384x384 input size was little, but in the original size (512x512), we achieved a slight increase compared to SOTA methods (1% or more in all evaluation indicators). Table 3 gave us a more specific view of how SE block and GeM Pooling affected the final results. GeM pooling showed its good performance in cross-view matching problems when using the same ResNet50 backbone, and the attention feature created by the SE block gave an important contribution to the final learning results compared to the basic residual block (approximately 4% increase in accuracy). This gave promising capability of doing further studies on the effectiveness of SE block and GeM pooling on the cross-view problem.

5. Conclusion

In this paper, we addressed the problems in navigating UAVs without GPS and focused on solving cross-view image matching tasks for navigation. We designed a network using a channel-based attention mechanism with a multi-scale generalized mean pooling strategy and verified its effectiveness on the University-1652 dataset. Experiment results showed that our proposed model has an increase in accuracy compared to the previous SOTA result. In future works, to increase the model's robustness to features in cross-view domains, we plan to further utilize the attention mechanism in the feature extracting phase, and some other pooling strategies are under-considered. Instead of 3D images, real environment data should also be tested in the next phase of our research.

References

1. M. Kontitsis, K. P. Valavanis and N. Tsourveloudis: "A UAV vision system for airborne surveillance", IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004, New Orleans, LA, USA, pp. 77-83 Vol.1, 2004,
2. S. W. Chen, et al: “Counting apples and oranges with deep learning: a data-driven approach”, IEEE Robotics and Automation Letters, vol. 2, no. 2, 2017,
3. L. Ding, et al.: "A Practical Cross-View Image Matching Method between UAV and Satellite for UAV-Based Geo-Localization.", Remote Sensing 13.1: 47, 2021,
4. L. Liu, H. Li, and Y. Dai.: “Stochastic Attraction-Repulsion Embedding for Large Scale Image Localization”, Proceedings of the IEEE International Conference on Computer Vision, pp. 2570–2579, 2019.
5. Y. Shi, Xin Yu, Liu Liu, Tong Zhang, and Hongdong Li. : “Optimal Feature Transport for Cross-View Image Geo-Localization”. arXiv preprint arXiv:1907.05021 , 2019.
6. Y. Tian, C. Chen, and Mubarak Shah: “Cross-view image matching for geo- localization in urban environments”, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3608–3616, 2017,
7. F. Castaldo et al.: “Semantic cross-view matching”, Proceedings of the IEEE International Conference on Computer Vision Workshops, p. 9-17, 2015.
8. H. Yang et al.: “Cross-view Geo-localization with Evolving Transformer”,arXiv preprint arXiv:2107.00842, 2021.
9. N. Khurshid et al.: “Ground to Aerial Image Retrieval Through Deep Learning”, Proceeding of International Conference on Neural Information Processing, Springer, p. 210-221, 2019.
10. S. Zhu et al.: “Cross-view image geo-localization beyond one-to-one retrieval”, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, p. 3640-3649, 2021.
11. D. G. Lowe et al.: “Object recognition from local scale-invariant features”, Proceedings of the IEEE International Conference on Computer Vision, 1999.
12. H. Bay, T. Tuytelaars, and L. Van Gool.: “Surf: Speeded-up robust features”, Proceedings of European Conference on Computer Vision, 2006.
13. T. Lin et al.: "Learning deep representations for ground-to-aerial geolocalization.", Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.
14. S. Hu et al.: “Cvm-net: Cross-view matching network for image-based ground-to-aerial geo-localization”. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7258–7267, 2018.
15. V. Ashish, et al.: “Attention is all you need”, Advances in neural information processing systems. pp. 5998-6008, 2017.
16. A. Kolesnikov et al. “An image is worth 16x16 words: Transformers for image recognition at scale”, arXiv preprintarXiv:2010.11929,2020.
17. F. Radenović et al.: “Fine-tuning CNN image retrieval with no human annotation”, IEEE transactions on pattern analysis and machine intelligence, 41(7), pp 1655-1668, 2018.
18. Z. Zheng, Y. Wei, and Y. Yang. : "University-1652: A multi-view multi-source benchmark for UAV-based geo-localization.", Proceedings of the 28th ACM international conference on Multimedia, 2020.
19. H. Kaiming, et al. : "Deep residual learning for image recognition.", Proceedings of the IEEE conference on computer vision and pattern recognition, 2016.
20. K. Simonyan, Z. Andrew. : “Very Deep Convolutional Networks for Large-Scale Image Recognition”, arXiv 1409.1556, 2014.
21. J. Hu, L. Shen, and G. Sun.: "Squeeze-and-excitation networks.", Proceedings of the IEEE conference on computer vision and pattern recognition, 2018.
22. W. Tingyu, et al.: "Each part matters: Local patterns facilitate cross-view geo-localization.", IEEE Transactions on Circuits and Systems for Video Technology, 2021.
23. J. Zhuang et al.: “A Faster and More Effective Cross-View Matching Method of UAV and Satellite Images for UAV Geolocalization”, Remote Sensing, 13.19: 3979, 2021,
24. Z. Zheng, et al. "Dual-Path Convolutional Image-Text Embeddings with Instance Loss." arXiv preprint arXiv:1711.05535,2017.

Authors Introduction

Mr. Duc Viet Bui

He received his M.S degrees from Department of Computer Science, National Defense Academy of Japan in 2021. He is currently a doctoral student at Department of Computer Science in National Defense Academy of Japan. His research related to different applications of computer vision in aerial robotics. His interests include computer vision, machine learning, deep neural networks and aerial robotics.

Dr. Masao Kubo

He is Associate Professor of Department of Computer Science at National Defense Academy in Japan. He graduated from the precision engineering department, Hokkaido University, in 1991. He received his Ph.D. degree in Computer Science from the Hokkaido University in 1996 (multi-agent system). He had been the research assistant of the chaotic engineering Lab, Hokkaido university. He was a visiting research fellow of Intelligent Autonomous Lab, university of the west of England (2005). He is the associate professor of information system lab, Dep. of Computer Science, National Defense Academy, Japan. His research interest is Multi agent system.

Dr. Hiroshi Sato

He is an Associate Professor of the Department of Computer Science at the National Defense Academy in Japan. He obtained a degree in Physics from Keio University in Japan and degrees of Master and Doctor of Engineering from Tokyo Institute of Technology in Japan. He was previously a Research Associate at the Department of Mathematics and Information Sciences at Osaka Prefecture University in Japan. His research interests include agent-based simulation, evolutionary computation, and artificial intelligence. Dr. Sato is a member of the Japanese Society for Artificial Intelligence (JSAI), Society of Instrument and Control Engineers (SICE), and The Institute of Electronics, Information and Communication Engineers. IEICE). He was the editor of IEICE and SICE.

Recommendation an Emergency Patient Destinations by LightGBM

Ryota Kawaguchi, Masao Kubo, Hiroshi Sato
Information Engineering, National Defence Academy in Japan, Hashirimizu 1-10-20
Yokosuka, Kanagawa 239-0811, Japan
E-mail: em60041@nda.ac.jp, masaok@nda.ac.jp, hsato@nda.ac.jp
www.mod.go.jp/nda/english/

Daizoh Saitoh
National Defence Medical College of Japan, Namiki 3-2 Tokorozawa, Saitama
359-8513, Japan
www.mod.go.jp/ndmc/

Takao Oya
Saitamaseibu Fire Bureau, Keyakidai 1-13-11
Tokorozawa, Saitama 359-1118, Japan
www.saisei119.jp/

Abstract

We propose to adopt the LightGBM method for recommending a hospital through we conduct some experiments by using patient transport data, and we found that location information, age, degree of injury are important elements in the situation of selecting a hospital. In addition, we found that accuracy can be obtained without detailed personal information, and the task of selecting a destination hospital can be decentralized.

Keywords: emergency patient, LightGBM, Machine Learning.

1. Introduction

In this study, we propose a method to recommend an emergency patient destination to reduce the burden of selecting a destination hospital for rescue teams which run in parallel with first aid and selecting a destination hospital. Since we use only patients' basic status that may be allowed to be used as input, it can be implemented even in local governments with insufficient understanding of IT utilization in medical care. Concretely speaking, we use patient summary information, location information, and conversations with call centers as input data for machine learning. The data used for the verification here are the data of the 5 most frequently used hospitals in the case of emergency transportation in western Saitama prefecture. The destination hospital was predicted based on the above data, and the accuracy was obtained by comparing it with the actual destination. We have investigated several machine learning methods. As a result, the LightGBM, a decision tree-based machine learning method, achieved the best accuracy compared to some other methods, with a score of about 70%. In addition, as a result of analyzing the emergency transport data, it was found that age, degree of injury, and location information are important for improving the accuracy.

2. Previous Researches

According to the data released by the Fire and Disaster Management Agency of Japan, the number of emergency patients is increasing year by year due to the aging of the population[1]. Therefore, there is an urgent need to reduce the burden of transport operations for rescue teams. In addition, the transportation time from the report to the hospital is increasing. We think that it needs to streamline transportation operations to save critically ill patients who require hospital treatment at an early stage. Against this background, there are some previous studies based on emergency transport data. In the study [2], machine learning methods are used to determine whether hospitalization is necessary or not. In the study [3], it was investigated the conditions of patients who are denied hospitalization. In the study [4], machine learning predicts whether a hospital can admit a patient, but the algorithm is not disclosed because it contains personal information. Therefore, in this research, detailed personal information is not used as input data for machine learning. Not handling the detailed personal information will make it difficult to predict, but will facilitate system introduction.

3. Data Set

The dataset contains 155,369 records taken to 287 hospitals in western Saitama Prefecture. We use only the records of the 5 most used hospitals in this dataset. Since the name and address of the hospital are not recorded, they are called A to E. The number of extracted records is 63,829, which includes 10 elements of information such as notification hour, notification month, day property, sex, age, degree of injury, consciousness level, location information, conversation with a call center, and destination hospital number. The breakdown of this dataset is shown in Table 1. There are 26,186 records to Hospital A, accounting for about 41% of the total. Even the smallest E has about 11%. In the following, Summary Vector(SV) includes information of notification hour, notification month, day property, sex, age, degree of injury, consciousness level. Location Vector(LV) includes information about the location. Conversation Vector(CV) includes information of conversation with a call center. All elements of these 3 vectors were normalized to the range 0 to 1.

Table 1. Data Set Breakdown.

Hospital	Number	Percentage
A	26,186	41.0
B	12,661	19.8
C	9,644	15.1
D	7,848	12.3
E	7,490	11.7
Total	63,829	100.0

3.1. *Summary Vector(SV)*

A SV includes 7 elements. Notification hour and notification month are the information when the call center is notified. The day property is a holiday or weekday. The degree of injury has one of mild, moderate, severe, or death. Consciousness level is an integer from 0 to 300, where 0 is normal and 300 is unconscious. Convert each of these elements from 0 to 1. For example, 0 for men and 1 for females.

3.2. *Location Vector(LV)*

A LV is anonymized location location information of a 2-dimensional. The dataset contains information up to the village section of a patient's location to protect personal information. Convert this information into a vector with two-dimensional elements. First, convert the address to latitude and longitude using the Yahoo API[5]. Then normalize to the range 0 to 1.

3.3. *Conversation Vector(CV)*

A CV includes 300 elements. Each conversation in this data describes in short sentences for example "I'm bleeding in contact with a car". The sentences are converted into a 300-dimensional vector using the Doc2Vec model learned on Japanese Wikipedia[6].

4. Experiments

We conducted the following three experiments.

1) Examination of prediction algorithm: First, we examined the optimal prediction algorithm. We compared the five methods of Support Vector Machine (SVM), k-nearest neighbor method, logistic regression, neural network, and LightGBM, and clarified which method is suitable for this problem. For the input data, 309-dimensional vectors of SV (7-dimensional), LP (2-dimensional), and CV (300-dimensional) were used. we calculate the score with the following formula (1).

$$Score = \frac{Number\ of\ Correct\ Answers}{Number\ of\ All\ Data}. \qquad (1)$$

2) Sensitivity analysis of input vectors: Next, we clarify what vector is important. We experiment using the best prediction method in the above experiment. We experiment using the best prediction method in the above experiment with combination of various vectors.
3) Sensitivity analysis of Summary Vector: Third, we clarify what element is important in SV. We experiment with combinations of all elements of SV.

4.1. *Comparison by Prediction Algorithms*

The ratio of training data to test data was 9: 1, and each experiment was randomly divided. Table 2 shows averages of the accuracy of each method when 10 experiments were performed. From this result, it was found that LightGBM gives the best accuracy. We decided to use LightGBM for the subsequent experiments.

Table 2. Comparison by Prediction Algorithm.

Method	Score[%]
SVM	67.8
k-NN(k=10)	59.8
Logistic Regression	66.3
Neural Network	68.8
LightGBM	70.7

4.2. *Comparison by Input Vectors*

Next, we analyzed the important vectors among the three input vectors. Compare the accuracy by combinations of the 3 vectors. The prediction method used was LightGBM, which was the most accurate in the first experiment. The results are shown in Table 3. From this result, it was found that all vectors contribute to the improvement of accuracy, but LV is an important factor for improving accuracy.

Table 3. Comparison by Input Vectors.

Input vector	Number of input element	Score[%]
SV	7	44.4
LV	2	65.4
CV	300	41.7
SV + LV	9	69.7
SV + CV	307	45.1
LV + CV	302	67.4
SV + LV + CV	309	70.7

*SV:Summary Vector
LV:Location Vector
CV:Conversation Vector

4.3. *Comparison by Summary Vector Elements*

Third, we experimented with combination of various elements of SV by LightGBM. PV and CV are always used in this experiment. There are 2 ^ 7 = 128 patterns for all combinations. The results are shown in Table 4. Looking at Score Ranking 1 to 5 in Table 4, all 7 elements were needed because almost all 7 elements were used. In particular, the top 31 cases always included age and degree of injury, indicating that these two factors are important.

Table 4. Comparison by Summary Vector Elements.

Score Ranking	Notification Hour	Notification Month	Day property	Sex	Age	Degree of injury	Consciousness level	Score[%]
1	○	×	○	×	○	○	○	70.7
2	×	×	○	○	○	○	○	70.6
3	○	×	○	○	○	○	○	70.6
4	○	○	○	○	○	○	○	70.6
5	○	○	×	○	○	○	○	70.5
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
124	○	×	×	×	×	×	×	67.4
125	×	×	○	×	×	×	×	67.3
126	×	○	×	×	×	×	×	67.3
127	○	○	×	×	×	×	×	67.3
128	×	○	○	○	×	×	×	67.0

* ○:use element, ×:not use element

5. Discussion

In the first experiment, the five prediction methods were evaluated with all available information. As a result, it makes LightGBM the most effective for this problem. In this study, about 41% (63,829 / 155,369) of the total transport data was extracted and used as the input data. It shows that software can replace the job in about 70% of cases. In the second experiment, the sensitivities of three vectors(SV, LV, CV) were investigated. As a result, it founds all three vectors is indispensable. It was also found that LV was the most important. In the third experiment, the sensitivities of the seven elements of SV were investigated. As a result, it became clear that it is effective to use all seven elements. It was also found that gender and degree of injury are particularly important.

From the above three results, it was found that location information, age, and degree of injury should always be collected when constructing a system for recommending a destination hospital. Furthermore, the task of selecting the destination hospital can be decentralized by starting hospital selection from the time of notification.

6. Conclusion

In this study, to reduce the burden of selecting a destination hospital, we investigated a method of recommending a destination hospital from patient information without detailed personal information. As a result, LightGBM was excellent, and we were able to achieve a score of about 70% for the accuracy of selecting a hospital. We also tried various input data patterns and found that age, degree of injury and location information contributed to improving learning accuracy in particular.

Since information on the above three important elements can be transmitted from the any person to the call center, it is possible to promptly recommend a hospital and reduce the burden of selecting a hospital for an on-site rescue team. We also found that accuracy can be obtained without detailed personal information, and what kind of information is referred by the rescue team when selecting a hospital.

References

1. Fire and Disaster Management Agency of Japan, 2020.
2. Satoshi Iwai, et al. "Admission predicting system using pre-arrival information in emergency departments", AIM, SIG-AIMED-001-04, 2015.
3. Mika Kumagai, et al. Japanese Society of Public Health, 63(3) 116-124, 2018.
4. "NTT Data News Release".
5. Yahoo JAPAN!.
6. Yuki Okuda.

Authors Introduction

Mr. Ryota Kawaguchi

He received his Bachelor's degree from the Department of Information Engineering, Kagawa University, Japan in 2018. He is currently a Master course student in National Defense Academy in Japan.

Dr. Masao Kubo

He is an Associate Professor of the Department of Computer Science at National Defense Academy in Japan. He graduated from the precision engineering department, Hokkaido University. He received his Ph.D. degree in Computer Science from the Hokkaido University.

Dr. Hiroshi Sato

He is an Associate Professor of the Department of Computer Science at the National Defense Academy in Japan. He obtained a degree in Physics from Keio University in Japan and degrees of Master and Doctor of Engineering from Tokyo Institute of Technology in Japan.

A Framework for Understanding the Neural Underpinnings of Symbolic and Non-Symbolic Communication Based on Global Synchronization in Human Brain Activity

Masayuki Fujiwara, Takashi Hashimoto
School of Knowledge Science, Japan Advanced Institute of Science and Technology,
1-1 Asahidai, Nomi, Ishikawa, 923-1211 JAPAN
E-mail: m-fujiw@jaist.ac.jp, hash@jaist.ac.jp
www.jaist.ac.jp

Abstract

We propose a framework for understanding the neural underpinning of communication-based processes, using electroencephalogram (EEG) synchronization. The framework comprised four stages: (i) characterization of the target communication in a two-dimensional space defined by symbolic/embodied (non-symbolic) and voluntary/involuntary processes, (ii) a focus on the level of synchronization analysis on an ontological hierarchy, (iii) a construction of a neurocognitive model to explain neural mechanism, and (iv) empirical hypothesis testing of neural underpinning with model-based EEG connectivity neurofeedback processes. We claim that following the framework will advance our understanding of neural dynamics and mechanisms for communication. During this study, we analyzed two EEG experiments, while implementing two former stages: the formation of symbolic communication changing from voluntary to involuntary and embodied communication competing between voluntary and involuntary. Their outcome was a hypothesis that three brain regions were involved in interpreting symbols, motor intentions, and social coordination. Finally, we described the advantages and limitations of the proposed framework, following a discussion concerning its operational validation in the latter stage.

Keywords: Communication, Framework, Synchronization, Electroencephalogram, Neurocognitive modeling

1. Introduction

Human communication takes various modalities, such as using symbols (*symbolic communication*) and body movements (*embodied*, and often *non-symbolic, communication*), which enables communication with others. Therefore, the neural underpinning of symbolic and non-symbolic communication and its relationship to neural synchronization within and between the brains are currently being investigated. However, although such studies on the neural underpinning of these communication systems have been conducted independently between the symbolic and non-symbolic, differences and similarities between them are still not clearly understood. Thus, by devising a comprehensive framework for explaining these communication modalities, we can discuss a unified neural underpinning for them.

Hence, we propose a framework for understanding the neural underpinning of symbolic and non-symbolic communication systems from the viewpoint of synchronization. The framework comprised four stages:

(i) Characterization of the target communication in a space defined with two axes: symbolic/embodied (non-symbolic) and voluntary/involuntary,
(ii) A focus on the level of synchronization analysis on an ontological hierarchy of "micro-macro loop chains" from individual neural activities to social behavior, in which the upper level is organized from and constrains the lower level,

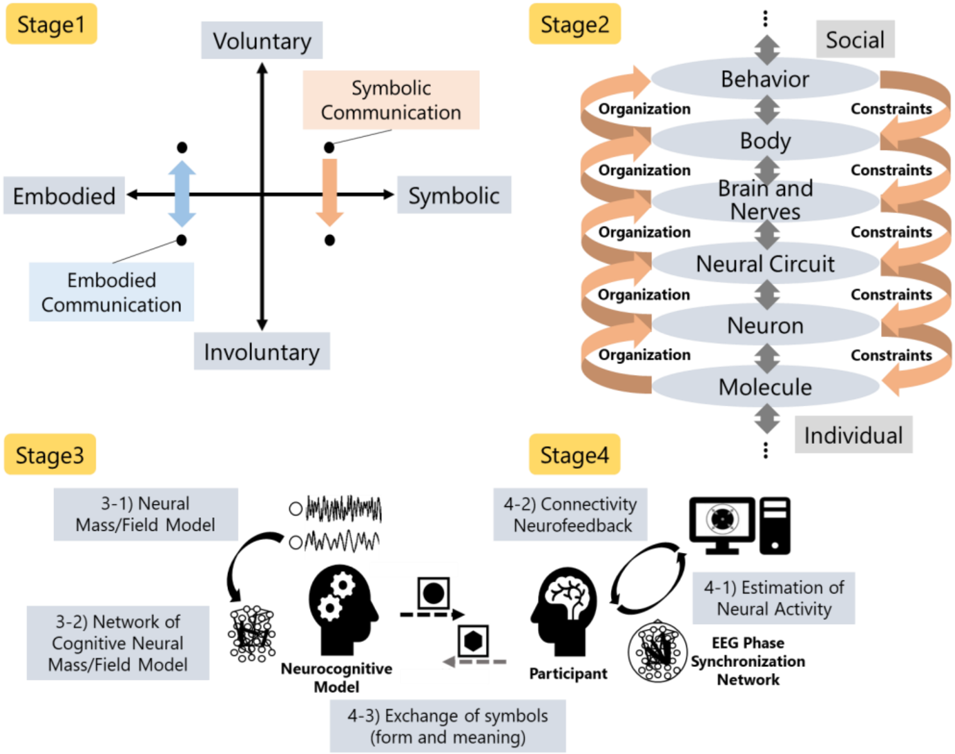


Fig. 1. A Framework for Understanding the Neural Underpinning of Symbolic and Non-Symbolic Communication. (Stage 1) A two-dimensional space to characterize communication modality. The horizontal and vertical axes were symbolic vs embodied and spontaneous vs involuntary, respectively. The two communication experiments summarized in this paper are illustrated with red and blue arrows. (Stage 2) A conceptual diagram showing a micro-macro loop chain with organizations and constraints. Upper levels focused on the social phenomena, whereas the lower focused on the individual phenomena in this ontological hierarchy. (Stage 3) Steps to construct a neurocognitive model, using the neural mass/field model and (Stage 4) to empirically validate the neural underpinning, using an EEG connectivity neurofeedback, during communication with a virtual partner based on the neurocognitive model.

to conduct empirical measurements during communication,

(iii) The construction of a neurocognitive model to explain neural mechanisms, and

(iv) Empirical hypothesis testing, with EEG connectivity neurofeedback methods, based on the model.

By going through these four stages, we expect to understand neural dynamics and communication mechanisms of communication deeply.

This paper also provides an overview of the findings from two EEG experiments on symbolic and non-symbolic communication, following this framework, and presents a working hypothesis of the neural underpinning of a communication system. Finally, we summarize the proposed comprehensive framework, describe its advantages and limitations, and discuss further operational validation methods of the neural underpinning hypothesis in the brain.

2. A Framework for the Neurological Understanding of Communication

2.1. *A two-dimensional space to characterize human communication*

At first, we positioned the target communication as a research subject in a two-dimensional space according to the target's communication modality. The space comprised two axes: the symbolic/physical (non-symbolic) axis and the voluntary/involuntary axis. In Fig. 1 (Stage 1), our two EEG experiments that show communication were placed on the space.

Note that the positioning is either a point or a changing path in the space. Targets in this paper were represented as paths having directionalities since we were interested in dynamic phenomena, such as the formation of communication systems and intentional switching during communication processes. This stage makes it possible to characterize the target communication clearly.

2.2. *The micro-macro loop chain on cognitive neuroscience*

In the second stage, we examined which ontological level should be focused on to empirically investigate the targeted communication in the first stage. Therefore, we proposed a "micro-macro loop chain" with emergent and constrained feedback loops among ontological levels (Fig. 1, Stage 2). This concept is inspired by the micro-macro loop in social science (organization theory[1] and economics[2]), in which micro information is connected to macro information which is then fed back to the micro-level.

The critical issue is the level of focus. Once the target level is determined, upper and lower levels become apparent. Then, we conducted empirical measurements and analyses to clarify the self-organization and constraints between the target level and the upper/lower levels.

2.3. *Computational modeling using a neural mass/field model*

In the third stage, we construct a computational model of phenomena at the target level (Fig. 1, Stage 3) through two steps: 3-1) building a model for the lower level and 3-2) making a network of the lower level models. This model construction approach is a sort of constructive approach[3], which is complementary to predictions and inferences from laboratory experiments and is effective to understanding complex phenomena and mechanisms. Following this approach, we construct a model or system that is based on a specific prediction or inference. Although the model is difficult to validate in actual situations, we run the model on a computer to verify it by comparing the computation results with real-world phenomena. The constructive approach is especially beneficial for neuroscientific studies of communication, focusing on the human brain.

Various models have been explored in computational neuroscience, starting with neuron models, neural mass models, such as the Wilson-Cowan model[4], next-generation neural field models[5], and neurocognitive models. By exploring what conditions are necessary for neural networks and the neural underpinning of human communication with a constructive approach, using these models, an understanding on the self-organization and constraints in the micro-macro loop chain will be achieved. While it is necessary to validate the constructed neurocognitive model by corresponding with actual phenomena, it is also possible to investigate more comprehensive models based on specific phenomena.

2.4. *A model-based EEG connectivity neurofeedback*

In the fourth stage, we 4-1) estimate neural activities during communication with a virtual partner[6,7], using a computational model constructed in the third stage, 4-2) conduct an EEG connectivity neurofeedback[8,9], and 4-3) perform an operational validation of neural mechanism of human communication. Specifically, the functional connectivity in the human brain is estimated through quantification of neural synchronization processes, following feeding to a computational model to reflect this quantification in the model's communicative behavior (e.g., decision making). Manipulating the model also allowed us to approach the empirical validation from the angle of indirect neural causalities.

3. EEG Recording Experiments

This section presents an overview of our two communication experiments, working on the proposed framework. For details on experiments and results, please refer to [10,11] for the former, and [12] for the latter.

3.1. *Symbolic communication tasks*

The symbolic communication task (red arrows in Fig. 1, Stage 1) is a coordination game[13,14], in which two participants performed a task through an exchange of symbols only. The game comprised a laboratory experiment, based on experimental semiotics[15,16], which allowed us to observe the emergence of artificial language by deliberately restricting the means of communication.

This task was designed to observe the emergent process of symbolic communication. Specifically, participants were required to move their avatars, placed in one of the four rooms, to the same room as their partner, only by exchanging predetermined and meaningless figures (Fig. 2). The task was designed such that it was impossible to successfully achieve the task without inferring implicit intentions as well as the correspondence between figures and rooms.

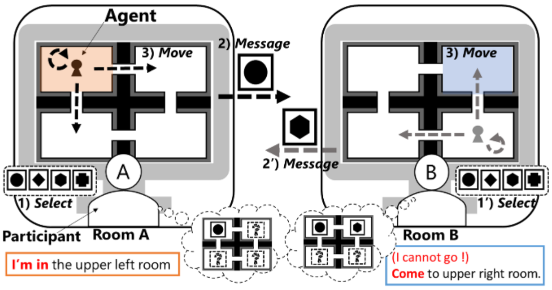


Fig. 2 An overview of the symbolic communication task.

3.2. *The Look This Way! task*

Regarding embodied non-symbolic communication, we proposed a new experimental paradigm called the "Look This Way!" task. In this task, pairs of participants played "janken" (rock-paper-scissors), followed by a finger-wagging task that was a modified version of a traditional Japanese game "Acchi-Muite-Hoi."

This task was designed to observe the representation and understanding of dynamic motor intentions when two participants switched cooperation and competition with each other (blue arrows in Fig. 1, Stage 1; Fig. 3). Specifically, during the "rock-paper-scissors," the two participants involuntary synchronized their rhythms when they shook their arms. In contrast, the subsequent "Look This Way!" task required voluntary *competitive* motions, particularly pointing the finger in a different direction from the partner. By comparing neural and physical activities with that of cooperative and scramble finger-pointing conditions, this experimental design allowed us to observe understanding and switching dynamic motor intentions that were involved in embodied communication.

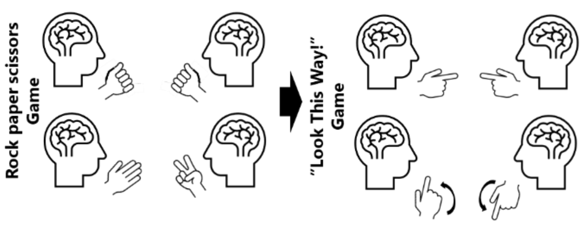


Fig. 3. An overview of the look this way! task.

3.3. *A working hypothesis of the neural underpinning of symbolic and non-symbolic communication systems*

The target level of the hierarchy in the experiments described in this paper is the neural activity during EEG recordings, with the upper level being the whole brain as functional connectivity between brain regions, and the lower level being the hierarchy at the neuronal level. The EEG recordings reflect a neural oscillation by electrical activities, in which the sum of the action and synaptic potentials of neurons appears with a certain rhythm. Neural oscillation is used to observe self-organization at the level of the functional network in the brain. On the contrary, neural oscillation is also maintained in a certain state due to the constraints of the brain network and is the constraint on neuronal activities. Epilepsy, for example, can be regarded as a condition, in which for some reasons, constraints of the brain network (or neural oscillation) are broken, and neurons become spontaneously and continuously active and synchronized.

Actually, we observed the neural activations and synchronization between the frontal and right centro-parietal regions from the former EEG experiment, and activations of the left fronto-central and right centro-parietal regions from the latter. The results suggested that three brain regions were involved in interpreting symbols, motor intentions, and social coordination processes. Therefore, we proposed a working hypothesis on the neural underpinning of symbolic and non-

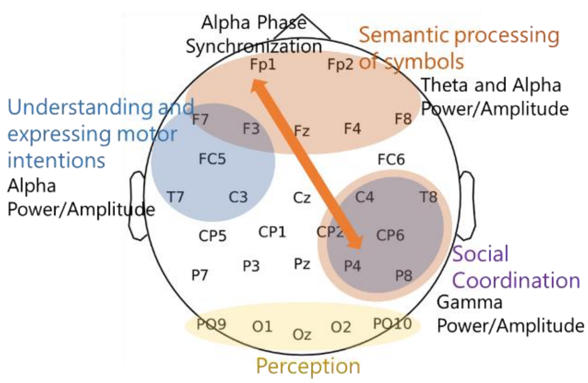


Fig. 4. A neural underpinning hypothesis for symbolic and non-symbolic communication processes

symbolic communication processes, as described in Fig. 4. We will also perform stages 3 and 4 in the proposed framework to validate this working hypothesis.

4. Discussions and Conclusion

We propose a framework, comprising four stages, to understand the neural underpinning of symbolic and non-symbolic communication. One advantage of the framework is the unified handling of symbolic/non-symbolic communication processes. Actually, we performed symbolic and non-symbolic communication experiments according to the framework. As a result, we proposed a new working hypothesis on neural underpinnings of symbolic and non-symbolic communication processes.

Meanwhile, although the validity of the framework has not been demonstrated yet, it is necessary to achieve this validation by confirming the working hypothesis based on the latter two stages of computational modeling, and the operational experiment, using a model-based EEG connectivity neurofeedback of human communication processes. While we focused on human communication processes so far, to consider the mechanisms of communicating with animals and machines, we need to also specify important factors to distinguish among humans, animals, and machines. It is therefore expected that the proposed framework will be helpful for such universal communication processes, including its understanding.

Acknowledgements

We are grateful to Jiro Okuda (Kyoto Sangyo University, Japan) and Hiroaki Wagatsuma (Kyushu Institute of Technology, Japan) for collaboration on each EEG experiment. This work was supported by JSPS KAKENHI Grant Numbers JP17H06383, JP17J06623 and JP26240037.

References

1 K. Imai and I. Kaneko, Network Organization Theory. Iwanami Shoten, 1988 (in Japanese).

2 Y. Shiozawa, Introduction to Complex Systems Economics. Chikuma Shobo, 2020. (in Japanese)

3 T. Hashimoto, "The Emergent Constructive Approach to Evolinguistics: Considering Hierarchy and Intention Sharing in Linguistic Communication," *J. Syst. Sci. Syst. Eng.*, Vol. 29, pp. 675–696, October. 2020.

4 H. R. Wilson and J. D. Cowan, "Excitatory and Inhibitory Interactions in Localized Populations of Model Neurons," *Biophys. J.*, Vol. 12, No. 1, pp. 1–24, 1972.

5 Á. Byrne, D. Avitabile, and S. Coombes, "Next-generation neural field model: The evolution of synchrony within patterns and waves," *Phys. Rev. E*, Vol. 99, No. 1, January. 2019.

6 G. Dumas, G. C. de Guzman, E. Tognoli, and J. A. S. Kelso, "The human dynamic clamp as a paradigm for social interaction," *Proc. Natl. Acad. Sci.*, Vol. 111, No. 35, pp. E3726–E3734.

7 V. Kostrubiec, G. Dumas, P. G. Zanone, and J. A. Scott Kelso, "The virtual teacher (VT) paradigm: Learning new patterns of interpersonal coordination using the human dynamic clamp," *PLoS One*, Vol. 10, No. 11, p. e0142029, 2015.

8 A. Mottaz, M. Solcà, C. Magnin, T. Corbet, A. Schnider, and A. G. Guggisberg, "Neurofeedback training of alpha-band coherence enhances motor performance," *Clin. Neurophysiol.*, Vol. 126, No. 9, pp. 1754–1760, September. 2015.

9 A. Yamashita, S. Hayasaka, M. Kawato, and H. Imamizu, "Connectivity Neurofeedback Training Can Differentially Change Functional Connectivity and Cognitive Performance," *Cereb. Cortex*, Vol. 27, No. 10, pp. 4960–4970, October. 2017.

10 F. Masayuki *et al.*, "Characteristics and connectivity of phase synchrony between success and failure group during symbolic communication task," *Proc. 27th Annu. Conf. Japanese Neural Netw. Soc.*, pp. 53–54, 2017.

11 M. Fujiwara *et al.*, "Changes in Phase Synchronization of EEG during Development of Symbolic Communication Systems," *Adv. Cogn. Neurodynamics (VI)*, pp. 327–333, 2018.

12 M. Fujiwara, M. R. V. Sanchez, S. Mishima, T. Hashimoto, and H. Wagatsuma, "An EEG Power Analysis with Hyperscanning EEG-Motion-Gaze data in the "Look This Way!" Game: A Pilot Study." (in preparation)

13 T. Konno, J. Morita, and T. Hashimoto, "Symbol communication systems integrate implicit information in coordination tasks," *Adv. Cogn. Neurodynamics (III)*, pp. 453–459, 2013.

14 G. Li, T. Konno, J. Okuda, and T. Hashimoto, "An EEG Study of Human Mirror Neuron System Activities During Abstract Symbolic Communication,"*Cogn. Neurodynamics (V)*, pp. 1–6, 2015.

15 B. Galantucci and S. Garrod, "Experimental Semiotics: A Review," *Front. Hum. Neurosci.*, Vol. 5, February, pp. 1–15, 2011.

16 B. Galantucci, "Experimental Semiotics: A New Approach for Studying Communication as a Form of Joint Action," *Top. Cogn. Sci.*, Vol. 1, No. 2, pp. 393–410, April. 2009.

Authors Introduction

Mr. Masayuki Fujiwara

He graduated from the Department of Information and Systems Engineering, Faculty of Information Engineering, Fukuoka Institute of Technology (FIT) in 2015 and received his master's degree from the School of Knowledge Science, Japan Advanced Institute of Science and Technology (JAIST) in 2017. He was a Research Fellow of the Japan Society for the Promotion of Science (JSPS) (DC1) until 2020. He has been studying the EEG research and neural modeling of symbolic and non-symbolic communication from the viewpoint of complex systems and phase synchronization.

Dr. Takashi Hashimoto

He is a Professor at School of Knowledge Science, Japan Advanced Institute of Science and Technology (JAIST). He was given his Ph.D. degree in 1996 from the Graduate School of Arts and Sciences, the University of Tokyo. He has been studying the origin and evolution of language, the dynamics of communication, and the design of social institution from the viewpoint of complex systems, with pursuing to construct a scientific field called "knowledge science" for creation, sharing, and utilization of knowledge.

Characterization of randomness tests by using tests results of weakly correlated chaotic sequences

Akihiro Yamaguchi
Department of Information and Systems Engineering, Fukuoka Institute of Technology, 3-30-1 Wajiro-higashi, Higashi-ku, Fukuoka, 811-0295, JAPAN

Asaki Saito
Department of Complex and Intelligent Systems, Future University Hakodate 116-2 Kamedanakano-cho, Hakodate, Hokkaido 041-8655, JAPAN
E-mail: aki@fit.ac.jp, saito@fun.ac.jp

Abstract

For the test suite of randomness, a characterization method of the individual tests is proposed. The proposed method is based on the clustering of the results of individual tests for the alternative hypothesis constructed by the weakly correlated chaotic sequences. These sequences are generated by the chaotic true orbit of the piecewise linear chaotic map corresponding to the Markov process. Then we apply our proposed method to the test suite of NIST SP800-22 and try to construct an optimal subset in terms of the distance among tests.

Keywords: randomness test, NIST SP800-22, chaotic true orbit, piecewise linear chaotic map

1. Introduction

High quality pseudo-random number sequences are required in various fields of engineering, and the statistical test of randomness is one of the important subjects. A typical test suite of randomness, e. g. NIST SP800-22[1], is defined as a set of several different kinds of randomness tests. One problem here is that the similarity between the individual tests included in the test suite is not obvious, and it is difficult to make an argument for the optimality of a set of randomness tests. For this problem, Doganaksoy et al. and Sulak et al. studied the independency among tests and the construction of optimal subset of tests based on the experimental results[2,3]. Iwasaki theoretically proved the equivalency among some tests[4]. These studies based on the null hypothesis. To analyze the effectivity of test subsets, the analysis of the statistical power to detect the alternative hypothesis is also necessary.

In this study, we propose a characterization method of randomness tests based on the test results of weakly correlated binary sequences that corresponds to the alternative hypothesis. Theses weakly correlated binary sequences are generated by the piecewise linear chaotic map and its chaotic true orbits[5-8] to guarantee their stochastic properties, exactly[9]. We apply our proposed method to characterize randomness tests included in the test suite of NIST SP800-22 and try to construct the optimal subsets in terms of the distance among tests based on the test results of the alternative hypothesis.

2. Randomness Tests and Definitions

Let the target randomness test suite Γ_{ts} consist of N tests

$$\Gamma_{ts} = \{T_1, T_2, \cdots, T_N\}. \tag{1}$$

For example, the randomness test suite NIST SP800-22 consists of 15 kinds of 188 random number tests.

The null hypothesis H_0 of each randomness test is that the target sequence has the ideal statistical properties of the random number. Let the target sequences of the randomness test be the m binary sequences with length n. In the randomness test $T_i \in \Gamma_{ts}$, each target sequence is tested, and the p-value is obtained from its test statistic.

The p-value corresponds to the probability that the test statistic is equal to or more extreme (farther from H_0) than the test statistic calculated from the target sequence when the null hypothesis H_0 is true. If the p-value is less than the significance level α, H_0 is rejected, otherwise H_0 is accepted, which means the target sequence is judged to be random. As results, we obtain m p-values and decisions of randomness for each test $T_i \in \Gamma_{ts}$.

2.1. *Feature Vector and Distance Matrix*

To characterize the randomness test, we consider the set of different M alternative hypotheses

$$\Lambda_{alt} = \{H_1, H_2, \cdots, H_M\}. \tag{2}$$

For each alternative hypothesis $H_a \in \Lambda_{alt}$, m target sequences that obeys H_a are generated and tested for each $T_i \in \Gamma_{ts}$. Let the obtained p-value of the k-th tested sequence be $p_{T_i}(k; H_a)$ and the mean p-value be $\bar{p}_{T_i}(H_a)$. The feature vector of the randomness test T_i is defined as

$$\bar{v}(T_i; \Lambda_{alt}) = \{\bar{p}_{T_i}(H_1), \bar{p}_{T_i}(H_2), \cdots, \bar{p}_{T_i}(H_M)\}. \tag{3}$$

Then we construct the distance matrix such as

$$D(\Gamma_{ts}; \Lambda_{alt}) = \begin{pmatrix} d_{1,1} & \cdots & d_{1,N} \\ \vdots & & \vdots \\ d_{N,1} & \cdots & d_{N,N} \end{pmatrix}, \tag{4}$$

where $d_{i,j}$ is the distance between two feature vectors,

$$d_{i,j} = \left|\bar{v}(T_i) - \bar{v}(T_j)\right| \tag{5}$$

and $|\cdot|$ denotes the Euclidean norm. Furthermore, we define the total distance of the tests set $\Gamma \subseteq \Gamma_{ts}$ such as

$$td(\Gamma) = \sum_{T_i, T_j \in \Gamma} d_{i,j}. \tag{6}$$

We propose this distance (Eq. (5)) as a measure of the similarity between two tests and the total distance (Eq. (6)) as a measure of the variety of the tests set.

2.2. *Empirical Power of Multiple Testing*

The power of the test $T_i \in \Gamma_{ts}$ with the significance level α for the alternative hypothesis $H_a \in \Lambda_{alt}$ is defined as

$$Pw(T_i; H_a, \alpha) = \frac{1}{m} \#\left(\left\{ k \mid p_{T_i}(k; H_a) < \alpha \right\}\right), \tag{7}$$

where $\#(\cdot)$ denotes the number of elements. The power of the multiple testing with the K tests

$$\Gamma_K = \{T_{i1}, T_{i2}, \cdots, T_{iK}\} \subseteq \Gamma_{ts} \tag{8}$$

with the significance level α for the alternative hypothesis $H_a \in \Lambda_{alt}$ is defined as

$$\begin{aligned} &Pw(\Gamma_K; H_a, \alpha) \\ &\quad = \frac{1}{m} \#\left(\left\{ k \,\middle|\, \exists T_i \in \Gamma_K : p_{T_i}(k; H_a) < \frac{\alpha}{K} \right\}\right). \end{aligned} \tag{9}$$

Here, the Bonferroni correction is applied to the significance level of each test as α/K to maintain the significance level of the multiple testing that are rejected if even one of the tests in Γ_K is rejected. The Bonferroni correction, however, assumed the independency between tests. Since the some of the tests included in NIST SP800-22 are not independent, the actual significance level is expected to be smaller than α.

3. Construction of Alternative Hypotheses

In this study, we construct alternative hypotheses using weakly correlated chaotic sequences[9]. As an alternative hypothesis, we consider the weakly correlated sequences that has the following stochastic properties. (i) The probability of the length l subsequences are equal to 2^{-l}. (ii) The conditional probability of 0 and 1 following the specified length l subsequences

$$s = 0s_2 \cdots s_l, \qquad s' = 1s_2 \cdots s_l \qquad (s_i \in \{0,1\}) \tag{10}$$

is given as

$$\begin{cases} P(0|s) = P(1|s') = 1/2 + e \\ P(1|s) = P(0|s') = 1/2 - e \end{cases}, \tag{11}$$

where $-1/2 < e < 1/2$. This binary sequence can be generated using the chaotic dynamical system that exactly corresponds to the Markov process. This chaotic dynamical system

$$x_{i+1} = g(x_i) \quad \left(x_i \in [0,1)\right) \tag{12}$$

is given by the piecewise linear map

$$g(x) = \begin{cases} \xi_+ \cdot (x + 2he\Delta) & (x \in I_1) \\ \xi_- \cdot (x - 2(h+1)e\Delta) & (x \in I_2)) \\ 2x & (x \in I_0 \cup I_3) \\ \xi_- \cdot (x - 2^{-1} - 2he\Delta) & (x \in I_5) \\ \xi_+ \cdot (x - 2^{-1} + 2(h+1)e\Delta) & (x \in I_6) \\ 2x - 1 & (x \in I_4 \cup I_7) \end{cases}, \tag{13}$$

where $h = 0,1,\cdots,2^{l-1}-1$ is a number that corresponds to the subsequence s as binary number, $\xi_\pm = 2/(1 \pm 2e)$, $\Delta = 2^{-l}$, $I_i = [a_i, a_{i+1})$, $a_0 = 0$, $a_1 = h\Delta$, $a_2 = (h + \xi_+^{-1})\Delta$, $a_3 = (h+1)\Delta$, $a_4 = 2^{-1}$, $a_5 = 2^{-1} + h\Delta$, $a_6 = 2^{-1} + (h + \xi_-^{-1})\Delta$, $a_7 = 2^{-1} + (h + 1)\Delta$, and $a_8 = 1$. The example of $g(x)$ is shown in Fig. 1. The shape of $g(x)$ is a partly modified form of the Bernoulli map that generate the ideal random sequences. This dynamical system corresponds to the Markov process shown in Fig. 1(b).

By using the dynamical system (Eq. (12)), we can obtain the binary sequence $s_1 s_2 \cdots s_n$ such as

$$s_i = \begin{cases} 0 & (0 \le x_i < 1/2) \\ 1 & (1/2 \le x_i < 1) \end{cases}, \tag{14}$$

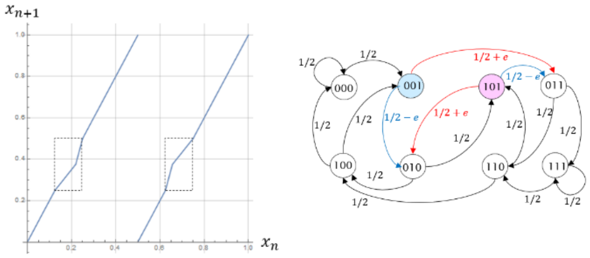
for the given initial point x_0. This binary sequence obeys the stochastic properties (i) and (ii), exactly. In this study, we calculate exact chaotic true orbit[5-8] of the dynamical system (Eq. (12)) and generate binary sequences that obey the alternative hypothesis.

4. Numerical Experiments

In this study, we try to characterize 14 randomness tests included in the NIST SP800-22 test suite. The target randomness tests Γ_{ts} are listed in Table 1. The non-overlapping template matching test and the random excursions, and its variant were excepted.

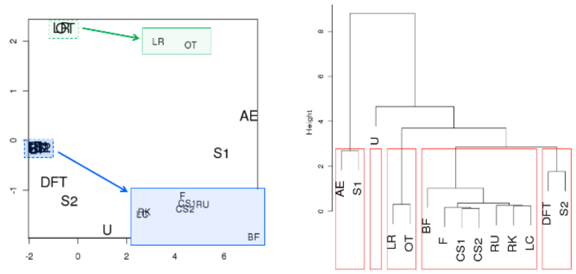
We generated $m = 10^3$ binary sequences with length $n = 10^6$ using Eq. (12) and (14), where $l = 4,\cdots 10$, $e = \pm 4^{-1}, \pm 8^{-1}, \pm 16^{-1}$, and 8 different patterns of s and s' for each l. Totally, $M = 336$ alternative hypotheses were constructed as Λ_{alt}. For the generated binary sequences, we applied the target 14 randomness tests in Γ_{ts} and calculated the p-value for each test and sequence. Then, we obtained the feature vectors (Eq. (3)) and the distance matrix (Eq. (4)). Examples of the cluster analysis based on the feature vector and the distance matrix are shown in Fig. 2. Fig. 2(a) is an example of the two-dimensional representation of the distance relation in $D(\Gamma_{ts}; \Lambda_{alt})$. This figure indicates the similarity among tests, e. g., AE and S1, LR and OT, DFT and S2, and the others except U. These similarities are also confirmed by the results of hierarchical cluster analysis shown in Fig. 2(b). These cluster analyses were performed using R.

The mean power of each test for Λ_{alt} is shown in Table 1. As a result, AE and S1 have high power to detect Λ_{alt}. On the other hand, F, BF, CS, RU, RK and LC could not detect Λ_{alt}. Then, we constructed the optimal subset



(a) (b)

Fig. 1. (a) An example of the piecewise linear chaotic map and (b) corresponding Markov process for the case that $l = 3$, $h = 1$ $(s = 001)$, and $e = 1/4$. The rectangular part of (a) is modified from Bernoulli map.



(a) (b)

Fig. 2. Examples of the cluster analysis based on the distance matrix $D(\Gamma_{ts})$. (a) A two-dimensional representation of $D(\Gamma_{ts})$ using the classical multidimensional scaling. Rectangular are magnifications of dense areas. (b) A cluster tree obtained by the hierarchical cluster analysis.

and the worst subset in terms of the maximization of the total distance (Eq. (6)) with respect to the size of subset $K = 2,\cdots,7$. Results are shown in Table 2 and Table 3, respectively. Here, $Pw(\Gamma_K)$ is the mean power of the multiple testing Γ_K (Eq. (8)) for Λ_{alt}. For the case of $K = 2$, the most distant pair AE and LC is optimal and the closest pair CS1 and CS2 is worst. The obtained optimal subsets are consistent with the results of the cluster analysis shown in Fig. 2. These results suggest that our method can construct the optimal subset with low similarity. However, our method cannot distinguish several tests that have low sensitivity to the proposed alternative hypothesis. To improve this problem, a more varied set of alternative hypotheses is necessary.

5. Conclusion

In this study, we proposed the characterization method of randomness tests and applied to 14 randomness tests in NIST SP800-22. As results, we obtained the optimal subsets in terms of the maximization of the total distance among tests. The characterization of all tests in NIST

Table 1. The mean power of randomness tests in Γ_{ts} for the constructed alternative hypotheses Λ_{alt}.

Target tests Γ_{ts} and abbreviations		$Pw(T_i)$
Frequency Test	F	0.010
Block Frequency Test	BF	0.012
Cumulative Sums Test	CS1	0.010
	CS2	0.010
Runs Test	RU	0.010
Longest Run Test	LR	0.122
Matrix Rank Test	RK	0.010
DFT Test	DFT	0.041
Overlapping Template Matching Test	OT	0.124
Universal Test	U	0.198
Approximate Entropy Test	AE	0.878
Serial Test	S1	0.628
	S2	0.090
Linear Complexity Test	LC	0.010

SP800-22 and the construction of other types of alternative hypothesis are our next future works.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Numbers 16KK0005, 17K00355. The computation was carried out using the computer resources offered under the category of General Projects by the Research Institute for Information Technology, Kyushu University.

Table 2. The optimal subset that maximizes the total distance.

K	$td(\Gamma_K)$	$Pw(\Gamma_K)$	Γ_K that maximize the total distance $td(\Gamma_K)$.
2	8.8	0.865	AE, LC
3	20.3	0.864	BF, LR, AE
4	37.9	0.860	BF, LR, AE, S1
5	59.4	0.857	BF, LR, U, AE, S1
6	84.4	0.854	BF, LR, U, AE, S1, LC
7	111.2	0.852	BF, LR, U, AE, S1, S2, LC

Table 3. The worst subset that minimizes the total distance.

K	$td(\Gamma_K)$	$Pw(\Gamma_K)$	Γ_K that minimize the total distance $td(\Gamma_K)$.
2	0.1	0.007	CS1, CS2
3	0.5	0.005	F, CS1, CS2
4	1.4	0.006	F, CS1, CS2, RK
5	2.6	0.007	F, CS1, CS2, RU, RK
6	4.2	0.008	F, CS1, CS2, RU, RK, LC
7	9.6	0.009	F, BF, CS1, CS2, RU, RK, LC

References

1. L. E. Bassham et al., NIST SP800-22 Rev. 1a: A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications, NIST CSRC, 2010.
2. A. Doganaksoy, et al., Mutual correlation of NIST statistical randomness tests and comparison of their sensitivities on transformed sequences, Turkish Journal of Electrical Engineering & Computer Sciences, 25, pp. 655-665, 2017.
3. F. Sulak, et al., On the independence of statistical randomness tests included in the NIST test suite, Turkish Journal of Electrical Engineering & Computer Sciences, 25, pp. 3673-3683, 2017.
4. A. Iwasaki, Study of the relationships among some tests in NIST SP800-22 (in Japanese), Proc. of the 2020 JSIAM Annual Meeting, pp.46-47, JSIAM, 2020.
5. A. Saito and S. Ito, Computation of true chaotic orbits using cubic irrationals, Physica D, 268, pp. 100–105, 2014.
6. A. Saito, et al., True orbit simulation of piecewise linear and linear fractional maps of arbitrary dimension using algebraic numbers, Chaos 25, 063103, 2015.
7. A. Saito and A. Yamaguchi, Pseudorandom Number Generation using Chaotic True Orbits of the Bernoulli Map, Chaos, 26, 063112, 2016.
8. A. Saito and A. Yamaguchi, Pseudorandom Number Generator based on the Bernoulli Map on Cubic Algebraic Integers, Chaos, 28, 103122, 2018.
9. A. Yamaguchi and A. Saito, Construction of the Markov process by using chaotic true orbit of piecewise linear map toward the independence analysis of randomness tests (in Japanese), Proc. of the 2017 JSIAM Annual Meeting, pp.255–256, JSIAM, 2017.

Authors Introduction

Dr. Akihiro Yamaguchi

He is a Professor of the Department of Information and Systems Engineering at Fukuoka Institute of Technology. He received his doctor degree in Science from Hokkaido University in 1997. His research interest includes neural networks and applications of chaotic systems.

Dr. Asaki Saito

He is a Professor of Department of Complex and Intelligent Systems at Future University Hakodate in Japan. He received his Ph.D. from University of Tokyo in 1999. His research interest includes true orbit computation of dynamical systems and learning dynamics.

A Mutual Control Method for a Multi-layered Non-contact Impedance Model-based Mobile Robots

Masaru Sasaki
Graduate School of Science and Engineering, Ibaraki University, 1-4-1 Nakanarusawa-cho Hitachi, Ibaraki 316-8527, Japan

Taro Shibanoki
Graduate School of Natural Science and Technology, Okayama University, 3-1-1 Tsushima-naka, Kita-ku Okayama, Okayama 700-8530, Japan

Hideyuki Tonooka
Graduate School of Science and Engineering, Ibaraki University, 1-4-1 Nakanarusawa-cho Hitachi, Ibaraki 316-8527, Japan

Toshio Tsuji
Faculty of Engineering, Hiroshima University, 1-4-1 Kagamiyama Higashi-Hiroshima, Hiroshima 739-8527, Japan

E-mail: 21nm726x@vc.ibaraki.ac.jp, shibanoki@okayama-u.ac.jp, hideyuki.tonooka.dr@vc.ibaraki.ac.jp, tsuji@bsys.hiroshima-u.ac.jp

https://www.okayama-u.ac.jp/user/mechatro/

Abstract

In this paper, a mutual control method is proposed for multi-layered non-contact impedance model-based mobile robots. In the proposed system, the motion priority is set to the robot, and the stiffness, viscosity, and inertia parameters of the non-contact impedance model are changed appropriately according to the priority value allowing the robots to avoid collisions with each other and obstacles at the same time. In the experiment, two mobile robots one controlled by electromyogram signals were prepared and operated to intersect. The other mobile robot automatically stopped and resumed its movement in response to the EMG-controlled robot with high priority, indicating that the proposed method can be used to control multiple robots.

Keywords: electromyogram (EMG), collision avoidance, noncontact impedance control, recurrent probabilistic neural network

1. Introduction

To support and enable the physically disabled to live independently, it is essential to operate various types of equipment and support mobility. Several studies have proposed biological signal-based interfaces for mobility support[1-5]. However, collision accidents caused by the use of these systems depend on the skills of the user.

Several conventional techniques have been proposed for obstacle avoidance[6-8]. These methods can avoid obstacles and other robots by planning the paths using artificial potential fields[9]. The author’s research group has also developed a natural obstacle-avoidance method for an EMG-controlled mobile robot that generates a multi-layered virtual wall based on the mechanical impedance model[10]. However, this system does not consider simultaneous control of multiple mobile robots. Therefore, in an environment where multiple mobile robots are present, the robots may recognize each other as an obstacle and an unintended movement may occur, such as repulsion caused by virtual forces.

In this paper, a multi-layered non-contact impedance (MLNCI) model is proposed that can automatically adjust the parameters according to the priority of each robot and apply them to the mobile robot control. Mobile robots within a specific region communicate with each other and adjust their parameters according to their control priority, enabling voluntary control even in an environment with multiple MLNCI model-based robots.

2. An EMG-controlled Mobile Robot Based on a Multi-layered Non-contact Impedance Model

In this section, an EMG-controlled mobile robot is described based on a multi-layered non-contact impedance model. EMG signals measured from L pairs of electrodes at a sampling frequency of f_s [Hz] are rectified and filtered using a second-order low-pass Butterworth filter (cut-off frequency: f_c [Hz]) and the values are then normalized such that the maximum value for each channel is 1 ($E_l(t)(l = 1, \dots, L)$). the feature vector used to estimate the movement direction of the robot is extracted from the normalized signals such that the sum of the channels is 1. The force information $F_{\mathrm{EMG}}(t)$, which is the average of $E_l(t)$, is assigned to the robot velocity $V_{\mathrm{EMG}} = V_{\max} F_{\mathrm{EMG}}(t)$.

The direction of movement of the robot is discriminated using a recurrent log-linearized Gaussian mixture network to enter the extracted features. It is possible to learn the motions corresponding to the reference directions in advance and output the posterior probabilities of the new input feature vectors for each reference direction arranged in a clockwise direction from $\pi/2$, and the movement direction vector $\boldsymbol{e}(t)$ at time t is extracted [11]. The moving direction $\theta(t)$ of the robot is determined from $\boldsymbol{e}(t)$, and the velocity [w_l, w_r] of both wheels of the robot is controlled.

In addition, I virtual walls are deployed around the mobile robot, which can generate virtual external forces according to the mechanical impedance. When an obstacle o enters the i_{th} virtual wall, the virtual external force $\boldsymbol{F}_o^i$ is defined by the following equation using the distance $\boldsymbol{X}_o$ from the virtual wall to obstacle o.

$$\boldsymbol{F}_o^i = \boldsymbol{M}_i \Delta \ddot{\boldsymbol{X}}_o^i + \boldsymbol{B}_i \Delta \dot{\boldsymbol{X}}_o^i + \boldsymbol{K}_i \Delta \boldsymbol{X}_o^i, \qquad (1)$$

where $\boldsymbol{M}_i, \boldsymbol{B}_i,$ and $\boldsymbol{K}_i$ represent the inertia, viscosity, and stiffness parameters, respectively, and $\boldsymbol{F}_O^i$ is zero if the obstacle is outside the i_{th} virtual wall ($|\boldsymbol{X}_o| > q_i$) or inside the $(i+1)_{\mathrm{th}}$ virtual wall ($|\boldsymbol{X}_o| \leq q_{i+1}$). The virtual force ${}^{(p)}\boldsymbol{F}_o$ received by all virtual walls $q_i(i =$

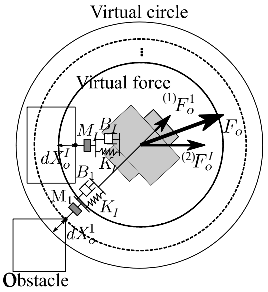


Fig. 1 Multi-layered non-contact impedance model [10].

$1,2, \dots, I)$ from all obstacles $o(o = 1,2, \dots, O)$ is expressed as the sum of ${}^{(p)}\boldsymbol{F}_o^i$. The equations of motion of the robot are defined as:

$${}^{(p)}\boldsymbol{M}\Delta\ddot{\boldsymbol{X}} + {}^{(p)}\boldsymbol{B}\Delta\dot{\boldsymbol{X}} + {}^{(p)}\boldsymbol{K}\Delta\boldsymbol{X} = {}^{(p)}\boldsymbol{F} + \boldsymbol{F}_o. \qquad (2)$$

By solving the equation, the velocity of the robot $\Delta\dot{\boldsymbol{X}} = \left[{}^{(p)}\dot{x}, {}^{(p)}\dot{y}\right]^{\mathrm{T}}$ is obtained, the velocities of both wheels of the robot based on the MLNCI model are calculated, and the velocity of the robot is adjusted ($\left[{}^{(p)}\dot{x} + V_{\mathrm{EMG}}, {}^{(p)}\dot{y}\right]^{\mathrm{T}}$)[10].

3. A Mutual Control Method for a Multi-layered Non-contact Impedance Model-based Mobile Robots

For example, when two robots with the same parameters collide, the virtual repulsive force causes them to move backward. In an environment with multiple robots, it is necessary to determine the robot that takes priority allowing the other robots to automatically stop or take evasive action. In such cases, to give priority to other robots, it is necessary to change each parameter of the virtual non-contact impedance model shown in Equation (1) to achieve appropriate control, such as automatic stopping. In this study, the priority *n* for each robot is set, and the coefficient vectors m_i^n, b_i^n, k_i^n required to adjust each parameter of the MLNCI model and adjust the virtual repulsive force according to the priority are as follows.

$$\boldsymbol{F}_o^i = m_i^n \boldsymbol{M}_i \Delta \ddot{\boldsymbol{X}}_o^i + b_i^n \boldsymbol{B}_i \Delta \dot{\boldsymbol{X}}_o^i + k_i^n \boldsymbol{K}_i \Delta \boldsymbol{X}_o^i. \qquad (3)$$

To achieve this priority-based control in a real environment, robots must be able to communicate with

Table 1 Parameters used in the experiment.

	$\boldsymbol{M}_1$[kg]	$\boldsymbol{B}_1$[Nm]	$\boldsymbol{K}_1$[N/m]
Robot 1	(0.21,0.21)	(0.8,0.8)	(1.0,1.0)
Robot 2	(2.0,2.0)	(2.0,2.0)	(2.0,2.0)

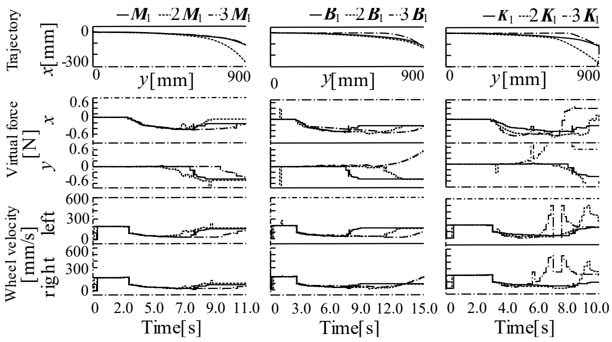


Fig. 2 Relationship between MLNCI parameters and virtual forces and velocities

each other and set their priorities. In this study, a server PC with a video camera that recognizes the robots by discriminating the markers attached to the robots is prepared, and the priorities are communicated through the server. Based on this priority, each parameter of the multi-layer non-contact impedance model is automatically adjusted, and the voluntary control of multiple robots can be realized.

4. Experiments

A validation experiment was conducted to identify the coefficient of varying for the MLNCI model. Two robots ($p = \{1, 2\}$) with one virtual wall ($I = 1$, see Table 1) were placed facing each other 900 [mm] apart. In the experiment, Robot 2 was moved forward at a speed of 200 [mm/s], and the parameters of inertia, viscosity, and stiffness were multiplied by 1, 2, and 3, respectively. RPLiDAR A1 (SLAMTEC) was used to measure the distances between robots (sampling frequency: 10 [Hz]).

The experiment results are shown in Fig. 2. The scatter plot at the top shows the trajectory of Robot 2, and the line graphs at the bottom show the results of varying the (a) inertia, (b) viscosity, and (c) stiffness of Robot 2.

The results show that the trajectory of Robot 2 until it starts to avoid Robot 1 is steeper when the stiffness parameter is varied compared to the viscosity and inertia parameters. The trajectory is also steeped in the second half when the inertia parameter is doubled. However, the velocities of both wheels change more slowly than when the stiffness parameter is varied. In addition, focusing on the change in velocity, the velocity slowly approached zero when the viscosity parameter increased, and the change in velocity was small when the inertia parameter increased. From these results, it can be observed that the robot can be stopped naturally by adjusting the inertia and viscosity parameters. In addition, the stiffness parameter can be used to rapidly move the robots away from each other. Based on the results, the next section shows that the proposed method can be used to preferentially control a mobile robot.

Table 2 Parameters used in the EMG-control experiment.

	$\boldsymbol{M}_1$[kg]	$\boldsymbol{B}_1$[Nm]	$\boldsymbol{K}_1$[N/m]
Robot 1	(0.21,0.21)	(0.8,0.8)	(1.0,1.0)
Robot 2	(6.0,6.0)	(6.0,6.0)	(2.0,2.0)
	$\boldsymbol{M}_2$[kg]	$\boldsymbol{B}_2$[Nm]	$\boldsymbol{K}_2$[N/m]
Robot 1	(2.1,2.1)	(8.0,8.0)	(10.0,10.0)
Robot 2	(6.0,6.0)	(6.0,6.0)	(4.0,4.0)

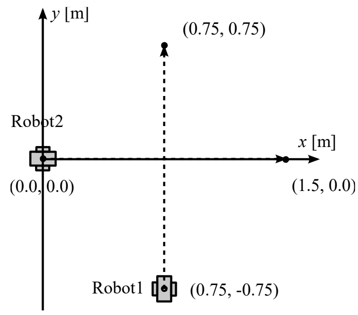


Fig. 3 Experimental setup

5. An example of EMG-based control

Based on the results presented in Section 4, mobile robot control was conducted in a real environment. Robot 1, which headed in the direction of the participant, was controlled using EMG signals measured from eight pairs of electrodes (Myo, Thalmic Labs, Inc.), and Robot 2 automatically moved toward $(x, y) = (1500, 0)$ [mm] (see Fig. 3). Two virtual walls ($I = 2$) were surrounded by each robot, and the parameters of the virtual walls were set as listed in Table 2.

Figure 4 shows the experimental scenes and trajectory of Robot 2. Figure 5 shows an example of the experiment results (from the top: virtual forces, wheel velocities, distance and angle to Robot 1, and the priority of Robot 2). The results show that from 0.5 [s], the parameters of Robot 2 were changed because of its lower priority, and it stopped to give priority to the myoelectric-controlled robot (Robot 1) between 1.5 to 11 [s]. At this time, the virtual repulsive force to Robot 2 was greater than that of Robot 1, and the velocity of each wheel was gradually decreasing. After Robot 1 passed, Robot 2 resumed its

(a) t = 4.0 [s] (b) t = 10.0 [s] (c) Trajectory

Fig. 4 Experimental scene and trajectory of Robot 1,2

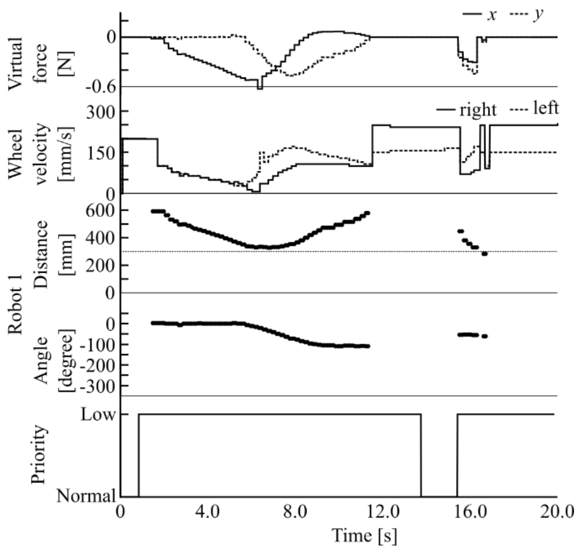


Fig. 5 An example of experimental results (Robot 2)

movement toward the position 1.5 [m] from the initial position.

These results show that the proposed method can be used for the mutual control of mobile robots based on a multi-layered non-contact impedance model.

6. Conclusion

In this paper, a mutual control method for mobile robots based on a multi-layered non-contact impedance model is proposed. The proposed method can simultaneously avoid collisions between robots and obstacles at the same time by changing the parameters of the model according to the motion priority of the robots. In the experiment, the myoelectric-controlled robot and automatic-controlled robot were manipulated such that they crossed each other. The experiment results showed that the automatically controlled robot was able to stop and resume its motion on its own in response to a myoelectric-controlled robot with higher priority.

In future research, we will consider the implementation of the model on aerial and undersea drones as well as the determination of priority using the movement direction and velocity of the robot and the surrounding environment through object recognition.

Acknowledgement

This work was supported by JSPS KAKENHI Grant Numbers: JP26330226 and 20K20212.

Reference

1. K. Tanaka, *et al*., Electro-encephalogram-based Control of an Electric Wheelchair, IEEE Trans. on Robotics 21 (4) (2005) 762-766.
2. M. Nakazawa, *et. al*., Implementation and its Evaluation of Autonomous Moving Wheelchair System Using EEG, IPSJ SIG Technical Report, (2015).Vol.2015-DPS-163 No.14
3. M. Yamamoto, *et al*., Analogue Input Interface for the Operation of Motored Wheelchair Using Respiratory Pressure for Severely Disabled Person, The Japan Society of Mechanical Engineers, (2009), Vol. 75, No. 751, pp. 703–709
4. Z. Bien, *et al*., Development of a wheelchair-based rehabilitation robotic system (kares ii) with various human-robot interaction interfaces for the disabled, In Proceedings 2003 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM 2003), (2003). Vol. 2, pp. 902-907.
5. K. Shima, *et. al*., CHRIS: Cybernetic Human-Robot Interface Systems in in Proc. 36th International Symposium on Robotics, (2005). WE1C3K.
6. D. Wang, *et al*., An obstacle avoidance strategy for the wave glider based on the improved artificial potential field and collision prediction model. Ocean Engineering, (2020). Vol. 206, p. 107356.
7. R. Singh, *et al*., Obstacle avoidance of mobile robot using fuzzy logic and hybrid obstacle avoidance algorithm. IOP Conference Series: Materials Science and Engineering, (2019). Vol. 517, p. 012009.
8. D. Wang, *et al*., UAV environmental perception and autonomous obstacle avoidance: A deep learning and depth camera combined solution. Computers and Electronics in Agriculture, (2020). Vol. 175, p. 105523.
9. X. Huang, *et al*., Mixed path planning for multi-robots in structured hospital environment. The 14th Asian Conference on Computer Aided Surgery (ACCAS 2018), eISSN 2051-3305.
10. T. Shibanoki, *et al*., An-EMG-controlled Mobile Robot Based on a Multi-layered Non-contact Impedance Model,

2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech), (2021). pp. 126-127.
11. O. Fukuda, *et al.*, An EMG-Controlled Omnidirectional Pointing Device, IEICE Transactions on Information and Systems, Pt.2, (2004). Vol. 87, No. 10, pp.1996-2003.

Authors Introduction

Mr. Masaru Sasaki

He received the B.E. degree from Ibaraki University in 2021. He is a master course student with the graduate school of science and Engineering, Ibaraki University, Hitachi, Japan.

Dr. Taro Shibanoki

He received the B.E. degree from the University of Tokushima in 2008, and the M.E. and D.Eng. degrees from Hiroshima University in 2010 and 2012, respectively. He was a Research Fellow of the Japan Society for the Promotion of Science since 2011, an Assistant Professor (Special Appointment) in Hiroshima University since 2013 and an Assistant Professor, Lecturer, and an Associate Professor in Ibaraki University since 2014, respectively. He is currently an Associate Professor in Graduate School of Natural Science and Technology, Okayama University. His current research interests focus on medical robotics, biological signal processing, and human-machine interface.

Dr. Hideyuki Tonooka

He is a professor at Ibaraki University, Japan. His main research areas are calibration, ground processing, and applications in remote sensing. He has been a member of the project science team for the Advanced Spaceborne Thermal Emission and Reflection radiometer (ASTER), part of NASA's Earth Observing System (EOS), since 1995. He is also joining the Compact Infrared Camera (CIRC) project of the Japan Aerospace Exploration Agency (JAXA).

Dr. Toshio Tsuji

He received the B.E. degree in industrial engineering and the M.E. and Dr. Eng. degrees in systems engineering from Hiroshima University, Japan, in 1982, 1985, and 1989, respectively. He has been a Full Professor at Hiroshima University since 2002. During 1992-1993, he was a Visiting Professor at the University of Genoa, Italy. His research interests range from engineering to human science with focuses on cybernetics, medical electronics, computational neural sciences, and particularly biological Kansei modeling. Dr. Tsuji has published more than 650 academic papers, and received 52 academic awards including the IEEE King-Sun Fu Memorial Best Transactions Paper Award (2003).

Relationship Between Delay Time and Sensation in Tactile Feedback for Myoelectric Prosthesis

Taro Shibanoki
Graduate School of Natural Science and Technology, Okayama University, 3-1-1 Tsushima-naka, Kita-ku Okayama, Okayama 700-8530, Japan[*]

Kosuke Jin
Graduate School of Science and Engineering, Ibaraki University, 1-4-1 Nakanarusawa-cho Hitachi, Ibaraki 316-8527, Japan
E-mail: shibanoki@okayama-u.ac.jp, 19nm718s@vc.ibaraki.ac.jp
https://www.okayama-u.ac.jp/user/mechatro/

Abstract

In this paper, we aim to develop a new tactile feedback method for myoelectric prosthetic hands and model the relationship between delay time and sensation in vibration stimulation. For myoelectric prosthetic hands, the tactile sensation can be expressed by vibrating an oscillator attached to the socket based on information obtained from tactile sensors attached to the prosthesis's fingertips. In this case, if there is a time gap between the sensory input and the stimulus, there is a possibility of causing discomfort. Therefore, in the experiment performed, a delay time, D_u [s], is set between the start of contact with an object and the start of vibration using a tactile sensor and conducted NRS evaluation. The results showed that the discomfort was generated up to D_u = 0.4 [s] and then decreased according to the delay time. The results showed that the discomfort was induced by controlling the timing of the vibration stimulus.

Keywords: electromyogram, prosthetic hand, tactile feedback, tactile sensation.

1. Introduction

Although there are various studies on myoelectric prosthetic hands to compensate for the lost functions of upper limb amputees, to create a truly human prosthetic hand, it is necessary not only to realize control that mimics human hand movement, but also to equip the prosthetic hand with a means of obtaining sensory information such as the tactile sensation that is inherent in the human arm. One of the reasons why there are so few myoelectric prosthetic hand users, despite the development of various advanced prosthetic hands, is that there is little sensory feedback to the user.

To solve this problem, research has been conducted to feed back the information obtained when the prosthetic hand comes in contact with an object to the operator. Feedback for myoelectric prosthetic hands can be divided into invasive and non-invasive methods, among which it is important to know what kind of stimuli to use and how to use them. Various types of stimuli such as vibration [1-2], temperature, pressure [3-4], electrical stimulation [5], and phantom limb stimulation [6] have been studied as possible feedback. On the other hand, there is research that uses electrical stimulation to present the pain of a localized strong stimulus, and research that combines image recognition to present the texture of an object. Pain, for example, may be useful for presenting contact with objects that should not be touched with a prosthetic hand, such as hot or sharp objects, but it is usually difficult to provide direct feedback of pain to the user.

The proposed method aims to present pain sensation, and to clarify the relationship between the timing of vibration stimuli and human sensation.

2. Method

The time taken for the sensory stimuli to be recognized as contacting an object is approximately 50 to 300 [ms].

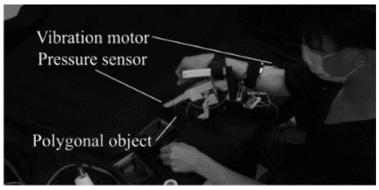


Fig. 1. Overview of the myoelectric prosthetic hand feedback system.

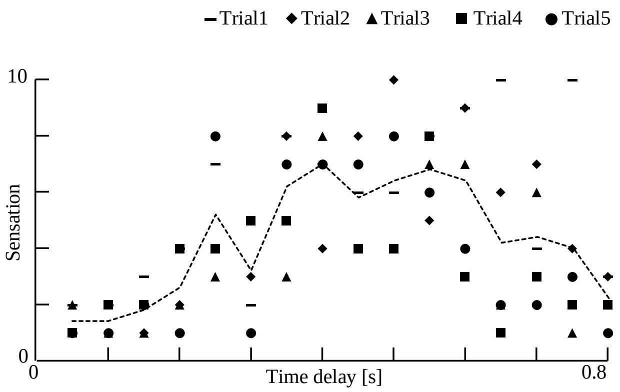


Fig. 2. Experimental results.

In a music game, the timing of the button operation and its feedback was varied, and the results were evaluated by questionnaire, and the larger the deviation from the operator's expected timing, the greater the discomfort. Therefore, when the pressure sensor value exceeds the object contact judgment threshold th_s and the threshold value th_u during the contact judgment time S, the vibration stimulus is given after the delay time D_u [s] from the end of the contact judgment time S, and it is verified whether the operator feels discomfort against the unexpected stimulus. Figure 1 shows an overview of the myoelectric prosthetic hand feedback system using an oscillator. A tactile sensor is attached to the fingertip of the myoelectric prosthetic hand, and feedback using an oscillator is possible according to the sensor value.

3. Experiments

In order to verify the optimal delay time D_u for expressing pain as discomfort using the proposed method, an experiment was conducted on a healthy university student. In the experiment, when the subject initiated feedback at an arbitrary timing, the vibration stimulus generation was delayed by a total of 10 random delay times D_u between [0.1, 1.0] [s] in 0.1 [s] increments. In this experiment, the subject was asked to evaluate how much discomfort he felt by a questionnaire with a 10-point scale. In the experiment, the subject presses a button to generate the feedback once and the vibration stimulus ends as one trial. In this experiment, all the delay times of 10 steps are presented 10 times. In addition, the vibration stimulus is set to 2.0 [s].

Figure 2 shows the results of the experiment. The vertical axis is the evaluation value of the questionnaire, and the horizontal axis is the delay time D_u. From the graph, the optimal delay time for expressing discomfort by feedback is 0.4 [s].

4. Conclusion

This paper verified a time-delay of tactile stimulation whether it is possible to express the discomfort to the operator. In the experiment, the time-delay of vibratory feedback after contact determination to the feedback was evaluated. The time at which the feedback by the vibration stimulus started was varied in 10 steps of 0.1 [s] between [0.1, 1.0] [s], and presented randomly. The time-delay that caused the most discomfort was 0.4 [s] for the participant.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Numbers: 17K12723 and 20K20212.

References

1. T. Rosenbaum-Chou, W. Daly, R. Austin, P. Chaubey and D. A. Boone, Development and real world use of a vibratory haptic feedback system for upper-limb prosthetic users, J. Prosthetics. Orth., vol. 28, no. 4, 136-144, 2016.
2. F. Clemente, M. D 'Alonzo, M. Controzzi, B. B. Edin and C. Cipriani, Non-invasive temporally discrete feedback of object contact and release improves grasp control of closed-loop myoelectric transradial prostheses, IEEE Trans. Neural Syst. Rehabil. Eng., vol. 24, no. 12, 1314-1322, 2016.
3. M. Aziziaghdam and E. Samur, Providing contact sensory feedback for upper limb robotic prosthesis, Proc. IEEE Haptics Symp. (HAPTICS), 575-579, 2014.
4. T. Morita, T. Kikuchi and C. Ishii, Development of sensory feedback device for myoelectric prosthetic hand to provide hardness of objects to users, J. Robot. Mechatron., vol. 28, no. 3, 361-370, 2016.
5. M. A. Schweisfurth, M. Markovic, S. Dosen, F. Teich, B. Graimann and D. Farina, Electrotactile EMG feedback improves the control of prosthesis grasping force, J. Neural Eng., vol. 13, pp. 056010, Oct. 2016.

6. Osborn, Luke E et al., Prosthesis with neuromorphic multilayered e-dermis perceives touch and pain, Science robotics vol. 3, 19, 2018.

Authors Introduction

Dr. Taro Shibanoki

He received the B.E. degree from the University of Tokushima in 2008, and the M.E. and D.Eng. degrees from Hiroshima University in 2010 and 2012, respectively. He was a Research Fellow of the Japan Society for the Promotion of Science since 2011, an Assistant Professor (Special Appointment) in Hiroshima University since 2013 and an Assistant Professor, Lecturer, and an Associate Professor in Ibaraki University since 2014, respectively. He is currently an Associate Professor in Graduate School of Natural Science and Technology, Okayama University. His current research interests focus on medical robotics, biological signal processing, and human-machine interface.

Mr. Kosuke Jin

He received the B.E. and M.E. degrees from Ibaraki University in 2019 and 2021, respectively. His research interests focus on myoelectric prosthetic hand.

Effects of Tactile Stimulation Near the Auricle on Body Sway During Foot Stamping

Masaya Tadokoro
*Graduate School of Science and Engineering, Ibaraki University, 1-4-1 Nakanarusawa-cho
Hitachi, Ibaraki 316-8527, Japan*

Taro Shibanoki
*Graduate School of Natural Science and Technology, Okayama University, 3-1-1 Tsushima-naka, Kita-ku
Okayama, Okayama 700-8530, Japan*

Hideyuki Tonooka
*Graduate School of Science and Engineering, Ibaraki University, 1-4-1 Nakanarusawa-cho
Hitachi, Ibaraki 316-8527, Japan*

E-mail: 21nm726x@vc.ibaraki.ac.jp, shibanoki@okayama-u.ac.jp, hideyuki.tonooka.dr@vc.ibaraki.ac.jp

https://www.okayama-u.ac.jp/user/mechatro/

Abstract

This paper describes the effects of tactile stimulation near the auricles on body sway in during foot stamping. The system measures center of pressure, acceleration of upper body, and whole-body movements by skeleton tracking with a depth camera while the participant performs foot stamping on the stabilometer, and extracts evaluation indices based on four perspectives: the amplitude, variation of body sway, rhythm of body sway, and correlation between of each limb. In the prototype experiment conducted with one healthy male, the body sway during foot stamping for ten trials in non-stimulus condition is compared with that for two trials in stimulus condition that the constant tactile stimulus is applied. As a result, the variation of the movement of lower limb and upper body sway were significantly reduced. This result implies that applying the constant tactile stimulation near both of auricles may be effective in stabilizing posture during foot stamping.

Keywords: dynamic balance function, body sway, tactile stimulation, center of pressure, foot-stamping

1. Introduction

According to the 2021 Annual Report on the Aging Society,[1] the number of elderly people requiring care is increasing every year, which is one of the most important issues considering the recent super-aging population and the number of elderly people in Japan (approximately 28.4% of the total population[2]). It is important for elderly people to extend their healthy life expectancy, maintain their quality of life (QOL), and support their independent lives. Falls and fractures are a major cause of the major causes of decreased QOL and account for 18.1 % of the factors causing the need for care.[1] Even if the elderly do not require nursing care, the fear of falling can affect their cognitive ability and gait abilities[3], causing a decline in the QOL and affecting the activities of daily living (ADL). In addition, it is said

(a). Initial stance (b). Left stamp (c). Right stamp

Fig. 1. Protocol of the experiment.

that decline in cognitive function increases the risk of falls[4], and therefore, preventing falls is an urgent and essential issue.

There are two types of equilibrium functions in humans: static balance and dynamic balance, and measures to prevent falls are required for both.

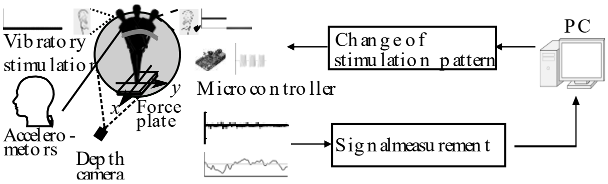


Fig. 2. The overview of the dynamic balance function analysis system based on tactile stimuli.

In the study of static balance function, body sway has been mainly analyzed by upright testing on a stabilometer, and light touch contact (LTC)[5], a phenomenon in which body sway is reduced by touching the fixed point with a fingertip at a force of less than approximately 1 [N], auditory stimulation[6], tactile stimulation[7], and galvanic vestibular stimulation (GVS) have been reported to reduce body sway[8]. In the study of dynamic balance function, rhythmic auditory stimulation (RAS) has been reported to improve gait function in gait analysis[9]. However, because of the possibility of being distracted by external stimulation owing to attention to rhythmic sounds, they are difficult to use in daily life. Our research group has shown that body sway can be reduced by applying vibratory stimulation around both sides of the pinna during tandem limb stance[10]. However, the effects of tactile stimulation on dynamic balance function have not been discussed.

This paper describes the evaluation of the effect of tactile vibratory stimulation on dynamic balance function by measuring the body sway during foot stamping on a stabilometer.

2. Method

The participant in the experiment was one healthy male who provided informed consent. The oscillator motors (KD18B1) were attached near both sides of the mastoid region of the participant and were connected to a Raspberry Pi 3 B and controlled via pulse width modulation (PWM) control with a duty cycle of 0.1 [s]. First, the participant was asked to face the front with both feet shoulder-width apart on the stabilometer (see Fig.1 (a)). Then, the participant was required to perform foot stomping from the left foot with natural arms swing and height of thigh in synchronization with a 1.0 [Hz] beep sound from the PC during 40 [s] with his eyes closed. The evaluation time was $T = 30$ [s], excluding the 5 [s] before and after the start of the measurement. A steady-state stimulus was applied to the participant with a constant duty ratio ($D = 1.0$) on both sides of near the auricles. The number of trials was 10 for both the stimulus and non-stimulus conditions, and the trials were conducted at intervals.

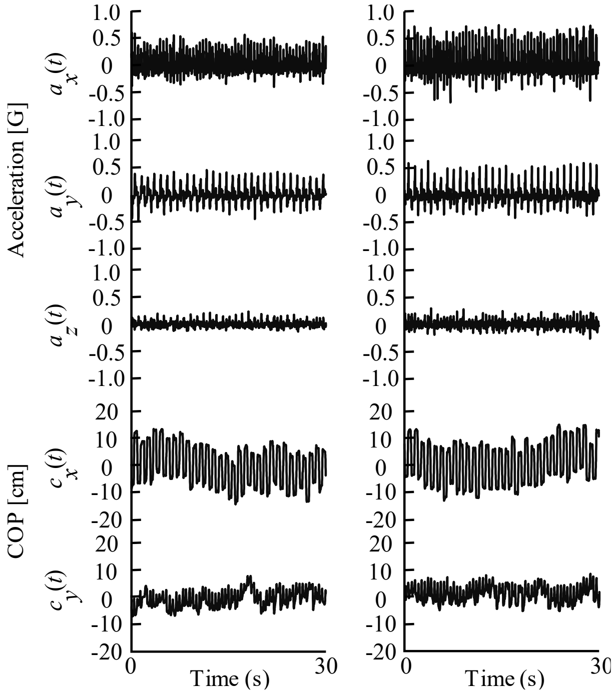


Fig. 3. An example of experimental results.

Figure 2 shows an overview of the dynamic balance function analysis system based on tactile stimuli. Two-dimensional time series data of center of pressure (COP) $c_{\{x,y\}}(t)$ were obtained from a Wii balance board (Nintendo, Co., Ltd.) at a sampling frequency of 100 [Hz] and smoothed using a second-order digital Butterworth low-pass filter (cut-off frequency: $f_c^L = 10$ [Hz]). To measure the participant's upper body sway, acceleration signals $a_{\{x,y,z\}}(t)$ were also measured using acceleration sensor (IMU-Z2, ZMP Inc.) attached to the back neck. The DC component in the signal was removed using a second-order digital Butterworth high-pass filter (cut-off frequency: $f_c^H = 0.5$ [Hz]). In addition, an RGB-D camera (Intel RealSense D435, Intel Corp.) was used to measure the body movements (sampling frequency: 30 [Hz]), and skeleton tracking was performed in real time with the skeleton detection middleware Nuitrack (3DiVi Inc.).

The filtered COP $c_{\{x,y\}}(t)$ and acceleration signal $a_{\{x,y,z\}}(t)$ were used to evaluate body sway from three perspectives: (i) amount and (ii) variation of movement and (iii) rhythm. Each index value was categorized as follows:

(i) Amount of movement

The evaluation indices based on COP:

- I_1: Movement width for $c_{\{x,y\}}(t)$ [cm]
- I_2: Total trajectory length [cm]

and that based on acceleration:

- I_3: Root mean square (RMS) that can be expressed by the combination of axes of $a_{\{x,y,z\}}(t)$ [cm]

(ii) Variation of movement

Mainly to evaluate the density of the body sway.

- I_4: RMS deviation area [cm^2]
- I_5: Rectangle area [cm^2]
- I_6: A standard deviation area [cm^2]

(iii) Rhythm

Mainly to evaluate the periodicity of the body sway.

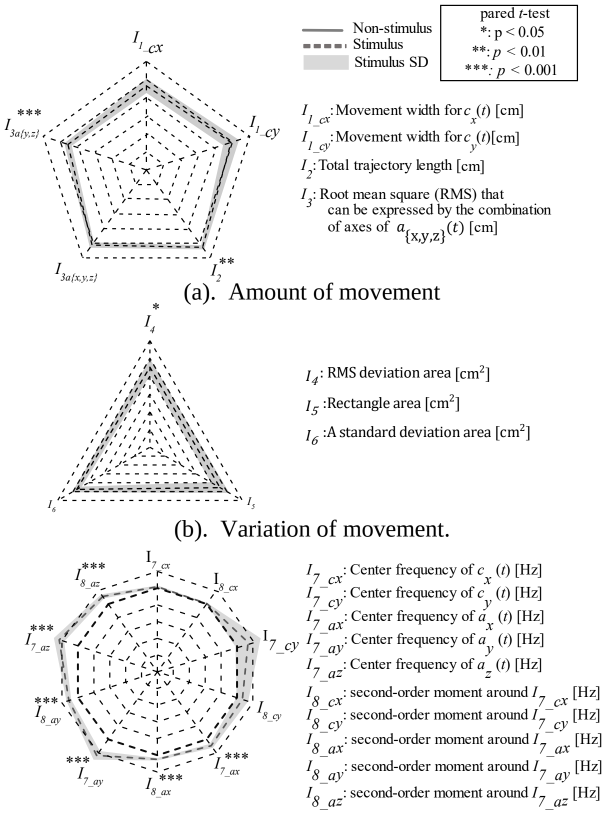
- I_7: Center frequency of $c_{x,y}(t)$ and $a_{\{x,y,z\}}(t)$ [Hz]
- I_8:Second-order moment of I_7 [Hz]

3. Results and discussion

Figure 3 shows the COP and acceleration signals measured in the experiment. The results shows that waveforms on the right fluctuated slightly.

Figure 4 shows the radar charts of each evaluation index for each category. Note that Fig. 4 shows the mean values of normalized the no-stimulus condition is 1. The results of a pared *t*-test are simultaneously shown in the figure. As shown in Fig. 4(a), the amount of movement, trajectory length I_2 significantly increase and horizontal plane composite acceleration I_3 decreased ($p < 0.01$). In addition, variation of COP (I_4) decreased ($p < 0.05$, see Fig. 4(b)). Although the trajectory length in the present measurement was considered to be affected by the change in foot width during foot-stamping and the force of stepping on the stabilometer, the velocity of COP movement was faster in stimulus condition that may have resulted in denser lower limb sway. Focusing on the index of rhythm, the center frequency and second-order moment of $a_{\{x,y,z\}}(t)$ increased significantly ($p < 0.001$).

Taguchi *et al*. reported that the amount of head sway in the horizontal plane increased in participants who were prone to body sway, such as abnormal peripheral vestibular function, as a result of attaching an accelerometer to the head and measuring the amount of sway[11]. In this study, these results implied that body sway was stabilized because of applying tactile stimulation to both auricles of the participant.



(a). Amount of movement

(b). Variation of movement.

(c). Rhythm.

Fig. 4. The experimental results.

4. Conclusion

This study investigated the effects of tactile stimuli applied to around the both sides mastoid region on dynamic equilibrium function by measuring body sway during foot stamping. In the experiment, the participant was asked to perform foot stomping in sync with a beep sound played every 1 [s] from the PC, and the changes in body sway with and without stable stimulation were compared. The results showed that the COP variation ($p < 0.05$) and the sum of horizontal RMS of acceleration ($p < 0.001$) decreased and the COP trajectory length ($p < 0.01$), and acceleration in all directions center frequency, and second-order moment increased ($p < 0.001$).

In future research, we plan to increase the number of participants and deepen the knowledge in this study and applying various stimulus patterns, and we will investigate the relationship between dynamic balance function and tactile stimulation.

Acknowledgement

This work was supported by JSPS KAKENHI Grant Numbers: 17K12723 and 20K20212.

References

1. Cabinet Office Japan, "Annual Report on the Ageing Society.
2. Cabinet Office Japan, "Annual Report on the Ageing Society [Summary] FY2020", July 2020.
3. B. J. Vellas, S. J. Wayne, L. J. Romero, R.N. Baumgartner, P.J.Garray, "Fear of falling and restriction of mobility in elderly fallers", Age and Aging, Vol. 26, pp.189-193, 1997.
4. L. Vereeck, F. Wuyts, S. Truijen, P.V de Heyning, "Clinical assessment of balance: normative data, and gender and age effects.", International Journal of Audiology, Vol. 47 No.2, pp.67-75, 2008.
5. M.A. Riley, T.A. Stoffregen, M.J. Grocki, M.T. Turvey: "Postural stabilization for the control of touching", Human Movement Science, Vol. 18, No. 6, pp. 795–817, 1999.
6. A. Majewska, W. K. Lkiewicz, D. H-Jezierska, A. Jezierska, L. Kubisz, "The influence of the acoustic stimulus on postural stability", Acta of Bioengineering and Biomechanics, Vol. 19, No. 2, pp.113-119, 2017.
7. L. Lipsitz, M. Lough, J. Niemi, T. Travision, H. Howlett, and B. Manor, "AShoe Insole delivering Subsensory Vibratory Noise Improve Balance and gait in Healthy Elderly People", Arch Phys Med Rehabili, Vol. 96, No. 3, pp. 432-439, 2015.
8. C. Fujimoto, Y. Yamamoto, T. Kamogashira, M. Kinoshita, N. Egami, Y. Uemura, F. Togo, T. Yamasoba and S. Iwasaki, "Noisy galvanic vestibular stimulation induces a sustained improvedment in body balance in elderly adults", SCIENTIFIC REPORTS, 2016
9. J.M.Hausdorff, J.Lowenthal, T.Herman, L.Gruendlinger, C.Peretz and N.Giladi, "Rhythmic auditory stimulation modulates gait variability in Parkinson's disease", European Journal of Neuroscience, Vol. 26, pp.2369-2375, 2007.
10. S.Watanabe, T.Shibanoki, K.Shimatani, *A Body Sway Mitigation Method Based on Tactile Stimulation*, Proceedings of the SICE Annual Conference 2016, pp.518-19, 2016.
11. K.Taguchi, C. Hirabayashi, and M. Kikukawa, "Clinical Significance of Head Movement while Stepping", The Oto-rhino-and laryngological clinic, Vol. 76, pp. 2418-

Authors Introduction

Mr. Masaya Tadokoro

He received his B.E. degree in 2021 from the College of Engineering, Ibaraki University, Ibaraki, Japan. He is currently a master course student in Ibaraki University. His research interests focus on balance function analysis.

Dr. Taro Shibanoki

He received the B.E. degree from the University of Tokushima in 2008, and the M.E. and D.Eng. degrees from Hiroshima University in 2010 and 2012, respectively. He was a Research Fellow of the Japan Society for the Promotion of Science since 2011, an Assistant Professor (Special Appointment) in Hiroshima University since 2013 and an Assistant Professor, Lecturer, and an Associate Professor in Ibaraki University since 2014, respectively. He is currently an Associate Professor in Graduate School of Natural Science and Technology, Okayama University.
His current research interests focus on medical robotics, biological signal processing, and human-machine interface.

Dr. Hideyuki Tonooka

He is a professor at Ibaraki University, Japan. His main research areas are calibration, ground processing, and applications in remote sensing. He has been a member of the project science team for the Advanced Spaceborne Thermal Emission and Reflection radiometer (ASTER), part of NASA's Earth Observing System (EOS), since 1995. He is also joining the Compact Infrared Camera (CIRC) project of the Japan Aerospace Exploration Agency (JAXA).

A Monitoring System of a Hamster Based on Video Image Analysis

Yugo Yamazaki
Graduate School of Science and Engineering, Ibaraki University, 1-4-1 Nakanarusawa-cho Hitachi, Ibaraki 316-8527, Japan

Taro Shibanoki
Graduate School of Natural Science and Technology, Okayama University, 3-1-1 Tsushima-naka, Kita-ku Okayama, Okayama 700-8530, Japan

Hideyuki Tonooka
Graduate School of Science and Engineering, Ibaraki University, 1-4-1 Nakanarusawa-cho Hitachi, Ibaraki 316-8527, Japan

E-mail: 21nm760a@vc.ibaraki.ac.jp, shibanoki@okayama-u.ac.jp, hideyuki.tonooka.dr@vc.ibaraki.ac.jp

https://www.okayama-u.ac.jp/user/mechatro/

Abstract

This paper proposes a monitoring system for a hamster that uses a video camera. The proposed system first processes the video image taken from the top of the cage, and then extracts features related to posture information and internal state. These features are used to discriminate between daily activities and other activities. This allows the system to alert when the hamster behaves differently from its daily routine. We analyzed the daily behavior of a hamster by using the proposed system. The results showed that the behaviors of a hamster changed when stimuli are given from outside the cage, and the system could discriminate them appropriately.

Keywords: abnormal detection, rodents, video camera, monitoring system, behavior tracking, heart rate variability

1. Introduction

As COVID-19 has been rampant, the number of pets has increased [1][2]. Therefore, it is helpful for pet owners if they can continuously monitor their pets. Conventional behavioral analyses of living things include visual observation of behavioral changes associated with changes in the living environment [3][4]. However, it is difficult to quantify the results, as the evaluation varies from researcher to researcher.

During the development of new drugs, research has been conducted to create behavioral models by measuring the movements of laboratory mice with a camera [5][6]. However, this method is not suitable for detecting behaviors that are not performed in daily life because it aims to quantify specific behaviors. Even though, there is a system to measure heartbeat and behavior by using a sensor module that includes a heart rate sensor, it is not suitable for measuring daily activities because the stress caused by such devices attached to a body changes its state [7].

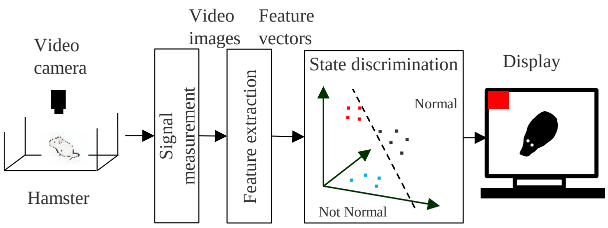
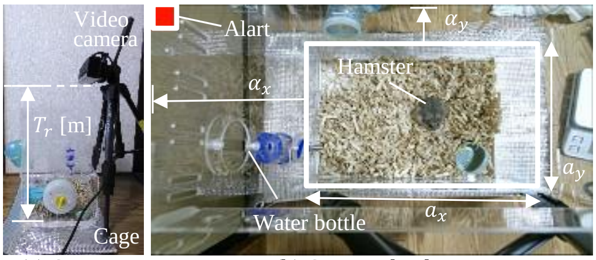


Fig. 1. Structure of the proposed system.

This paper proposes a monitoring system for pets based on video image analysis. The system can discriminate their state by extracting their internal information, which does not appear in their behavior, and its behavioral information simultaneously from a video camera. The proposed system extracts the heart rate variability from the heartbeats of the animals. In apartment buildings, other pets may also be kept, which can cause a great deal of stress [8]. The proposed system can detect states that do not appear in behavior by extracting the heart rate variability. In addition, the proposed system can alert when an unusual state is detected, and the owner can check the surrounding information from the video image, which may help in the early detection of injuries.



(a) Setup (b) System display

Fig. 2 Overview of the proposed system.

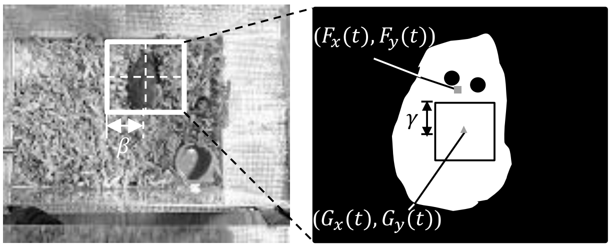


Fig. 3. Video image processing.

2. Method

In the proposed system, a video camera (Kinect V2, Microsoft) is used for real-time monitoring of a hamster in a cage, and alerts the owner when the system detects a state different from its daily behavior. Figure 1 shows the structure of the proposed system. The system consists of three parts: 1) signal measurement, 2) feature extraction, 3) state discrimination, and 4) display to present the information to the owner.

2.1 Signal measurement

The measurement environment of the system is shown in Fig. 2(a). A video camera is set up parallel to the ground without any tilt above a typical cage (T_r m), so that the daily life of a hamster can be measured. Various equipment such as a water bottle is placed (see Fig. 2(b)), and the height of the camera T_r is adjusted to maximize the hamster's view angle as much as possible.

In the signal measurement, the video image measured from the video camera is imported into the PC at a sampling frequency of f_s Hz. As shown in Fig. 3, a $a_x \times a_y$ rectangular area A with margins α_x and α_y pixels in each axis direction of the measurement area (see Fig. 2) is set from each frame to omit the living environment in the cage. The features for state discrimination were then extracted through the following signal processing.

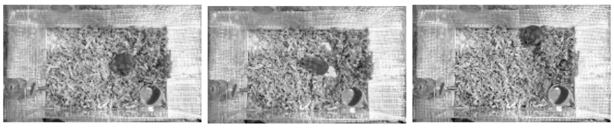

(a) t = 31.8 [s] (b) t =121.2 [s] (c) t = 300.0 [s]

Fig. 4 Scenes from the experiment.

After the measurement image is converted to the HSV color system, threshold values (the maximum and minimum for each component: $h_{\mathrm{th}}^{\max}$, $h_{\mathrm{th}}^{\min}$, $s_{\mathrm{th}}^{\max}$, $s_{\mathrm{th}}^{\min}$, $v_{\mathrm{th}}^{\max}$, and $v_{\mathrm{th}}^{\min}$) are set to extract only the hamster area, and the binary image is generated, where white (0) represents the hamster and black (1) represents the other area (see Fig. 3).

The body center coordinates of the hamster $G_x(t)$, $G_y(t)$ are obtained by calculating the center of gravity of the region of the pixels with value (0). A square region B with perimeter β pixels from $[G_x(t), G_y(t)]^{\mathrm{T}}$ is then defined, and positions estimating the hamster's eyes are also extracted as set threshold values. The coordinates of the center of gravity of the pixels in B with value (0) assuming the position of the eyes is tracked as the position of the hamster's head $[F_x(t), F_y(t)]^{\mathrm{T}}$ (see Fig. 3).

In addition, to measure heart rate variability as internal body information, a square region C of γ pixels around $[G_x(t), G_y(t)]^{\mathrm{T}}$ (the position of the hamster) in region A is defined, and the average value $H(t)$of the green component in region C is calculated.

2.2 Feature extraction

To extract the features that represent the motion and internal body information of the hamster, in this section, the change in the center of gravity $m(t)$, the change in body shape $[r(t), \theta(t)]$, and the heart rate information $h(t)$ is calculated.

2.2.1 Movement information

The change in body shape is defined using the hamster's center of gravity $[G_x(t), G_y(t)]^{\mathrm{T}}$ and head position $[F_x(t), F_y(t)]^{\mathrm{T}}$ as the features representing the hamster's body shape, referring to the method of Yuman *et al.*[4]. In this study, the tail of a rat, which is also a rodent, is used; however, it is difficult to obtain the tail of a hamster; therefore, in this paper, the following new equation is used for calculation.

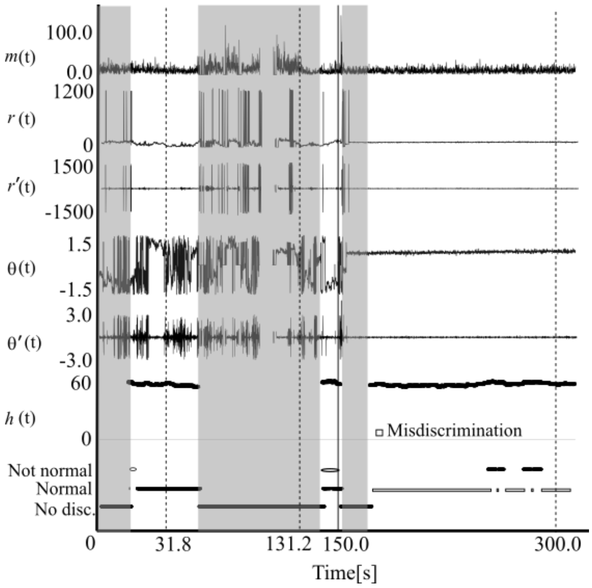


Fig. 5. Experimental result.

$$r(t) = \sqrt{\left(G_x(t) - F_x(t)\right)^2 + \left(G_x(t) - F_x(t)\right)^2} \quad (1)$$
$$\theta(t) = \tan^{-1}((G_y(t) - F_y(t))/G_x(t) - F_x(t))$$

where $r(t)$ is the distance between the head and the center of gravity of the hamster, and the larger the value, the more stretched the body. $\theta(t)$ is the angle between the center of gravity and the head. The value indicates the direction of the hamster, with clockwise being positive.

2.2.2 Internal body information

The heart rate information $h(t)$ is extracted from the heart rate during t_{H} s as a feature representing the internal state of the hamster. Tsumura *et al.* [9] improved the accuracy of heart rate information extraction by applying a band-pass filter to video images measured by five cameras and then performing independent component analysis. Here, because it is sufficient to calculate the heart rate, the peak is detected from $H(t)$, and the inverse of the time difference is multiplied by t_H. A second-order digital Butterworth filter with a cutoff frequency of (f_{l}, f_{h}) was applied to $h(t)$ to detect the peak easier. The mountain-climbing method was applied after applying a band-pass filter.

2.3 State discrimination

Threshold discrimination is used to classify the state of the hamster. First, the threshold value is determined based on the training data $\boldsymbol{Z}^n = [z_1(t)^n, \ldots, z_d(t)^n]^{\mathrm{T}}$

Table I. Confusion matrix.

	True	False
Positive	789	1975
Negative	110	90

$(n = 1,2, \ldots, N)$, obtained by measuring and extracting features of scenes containing K types of daily activities. The classification result $s(t)$ can be obtained by inputting a new feature vector. At this time, $s(t)$ becomes $s(t) = 1$ when the behavior is judged to be in the normal state (when values for all dimensions exceed the threshold), and $s(t) = 0$ when a state different from daily activity is obtained.

2.4 Display

In this section, the results of the system are presented on the display to inform the owner of the hamster's state. When the state of the hamster is different from that of daily life ($\mathrm{s}(t) = 1$), the area of $s_x \times s_y$ pixels in the upper left corner of the system screen is displayed in red, as shown in Fig. 3.

As described above, the state of hamster daily can be monitored with the system and notify the owner of any abnormalities. In addition, because the owner can check the video images around that time, he/she can visually understand the situation.

3. Experiment

To verify the effectiveness of the proposed system, the daily activities of the hamster was monitored using the system.

3.1 Experimental conditions

The experiment was conducted in the rearing environment shown in Fig. 1($T_r = 0.5$ m), with the cabin removed so that the hamster's movements could be observed. We tried to reduce the stress of the hamster by matching the measurement environment and the rearing environment as much as possible. To distinguish their daily conditions, the data were measured in the evening, when nocturnal hamster started their activities. In the experiment, we generated a sound by clapping hands after 150.0 s in the measurement of approximately 5 min. (4838 frame, $f_s = 16.16 \pm 0.38$ Hz) to add load to the hamster. In addition, inter-frame difference of $[r(t)$, $\theta(t)]$ is calculated and used to discrimination $[r'(t)$,

$\theta'(t)]$ ($\boldsymbol{Z}^n = [m(t)^n, r'(t)^n, \theta'(t)^n, h(t)^n]^{\mathrm{T}}$). Threshold values for discrimination were set as $-10 \le m \le 10$, $-20 \le r \le 20$, $-0.5 \le \theta \le 0.5$, $63 \le h$.

The other parameters were set by trial and error, and $\alpha_x = 680$ pixels, $\alpha_y = 150$ pixels, $a_x = 990$ pixels, $a_y = 650$ pixels, $\gamma = 21$ pixels, $s_x = s_y = 100$ pixels, $h_{\mathrm{th}}^{\max} = 95$, $h_{\mathrm{th}}^{\min} = 60$, $s_{\mathrm{th}}^{\max} = 30$, $s_{\mathrm{th}}^{\min} = 0$, $v_{\mathrm{th}}^{\max} = 75$, $v_{\mathrm{th}}^{\min} = 125$ for hamster's position estimation, and $h_{\mathrm{th}}^{\max} = 65$, $h_{\mathrm{th}}^{\min} = 30$, $s_{\mathrm{th}}^{\max} = 100$, $s_{\mathrm{th}}^{\min} = 50$, $v_{\mathrm{th}}^{\max} = 40$, $v_{\mathrm{th}}^{\min} = 10$ for eye position estimation. In addition, f_{l}= 4.76 Hz f_{h}=7.14 Hz is set because the heart rate of hamster was 280-412 bpm [10].

3.2 Results and discussion

Figure 4 shows the experimental scenes, and Figure 5 shows the results. The horizontal axis represents the time, and the vertical axis represents, from the top, the change in center of gravity $m(t)$, the change in body shape [$r(t)$, $\theta(t)$], their inter-frame difference [$r'(t)$, $\theta'(t)$], and heart rate information $h(t)$, and the result $s(t)$. In the section where the center of gravity could not be traced owing to the hamster being hidden under straw, no discrimination is made (shaded area in fig 5). From figure 4, it can be seen that 31.8 s and 121.2 s show a large change in the value of $\theta(t)$. These are the areas where the hamster moves such as curling up, and the system can track the hamster's behavior.

After 150.0 s (vertical line in Fig. 5), the extracted features fluctuated due to the sound. The results for the condition before the sound are also shown in Table I. Table I shows the confusion matrix for the state before and after the sound. Here, the state after the sound was considered to be a state different from daily life.

From these results, the proposed system can be used for the daily monitoring of a hamster. However, because only one example is shown in this paper, we will increase the measurement time and perform real-time monitoring for 24 h.

4. Conclusion

This paper proposes a monitoring system for a hamster using a video camera. The system can discriminate between different states from daily life by extracting features appearing in the behavior and internal states. In the experiment, the daily life of a hamster was monitored using the proposed system, and the change was evaluated in behavior before and after the sound was generated. Although the average discrimination rate was low (average discrimination rate: 30%), the system showed the possibility of immediately informing the owner of a situation where the hamster was overloaded.

In the future, we will conduct real-time monitoring for an extended period, and consider feature extraction and classifier to improve discrimination accuracy.

Acknowledgement

This work was supported by JSPS KAKENHI Grant Numbers: 17K12723 and 20K20212.

References

1. "Japan pet food association", Accessed December 2021,
2. Cabinet office government of Japan. Accessed December 2021.
3. Freeman, G. B.; Johnson, J. D.; Killinger, J. M.; Liao, S. C.; Davis, A. O.; Ruby, M. V.; Chaney, R. L.; Lovre, S. C.; Bergstrom, P. D. Bioavailability of arsenic in soil impacted by smelter activities following oral administration in rabbits. Fundam. Appl. Toxicol., (1993), Vol. 21, pp. 83-88.
4. Hettiarachchi, G. M.; Pierzynski, G. M.; Oehme, F. W.; Sonmez, O.; Ryan, J. A. Treatment of contaminated soil with phosphorus and manganese oxide reduces lead absorption by Sprague-Dawley rats. J. Environ. Qual. 2003, Vol. 32, pp. 1335-1345.
5. Yuman Nie, Takeshi Takaki, Idaku Ishii & Hiroshi Matsuda, Algorithm for Automatic Behavior Quantification of Laboratory Mice Using High-Frame-Rate Videos, SICE JCMSI, Vol. 4, No. 5, September 2011
6. Michael H. McCullough, Geoffrey J. Goodhill, Unsupervised quantification of naturalistic animal behaviors for gaining insight into the brain, Current Opinion in Neurobiology2021, Vol. 70, pp. 89-100.
7. Amruta Helwatkar, Daniel Riordan & Joseph Walsh, Sensor Technology For Animal Health Monitoring, International Journal on Smart Sensing and Intelligent Systems, International Journal Vol. 7, No. 5, pp. 1-6.
8. Robert D. Magrath, Tonya M. Haff, Pamela M. Fallow and Andrew N. Radford, Eavesdropping on heterospecific alarm calls: from mechanisms to consequences, Biological Reviews 90, 2015, pp. 560-586.
9. Tsumura, N, Non-contact Heart Rate Measurement Based on Facial Video Taken by RGB Cameras and Its Applications, Oleo Science, Vol. 21, No.5, 2021, pp. 3-10.
10. Japan veterinary medical group, Accessed December, 2021.

Authors Introduction

Mr. Yugo Yamazaki

He received the B.E. degree from Ibaraki University in 2021. He is a master course student with the Graduate School of Science and Engineering, Ibaraki University, Hitachi, Japan.

Dr. Taro Shibanoki

He received the B.E. degree from the University of Tokushima in 2008, and the M.E. and D.Eng. degrees from Hiroshima University in 2010 and 2012, respectively. He was a Research Fellow of the Japan Society for the Promotion of Science since 2011, an Assistant Professor (Special Appointment) in Hiroshima University since 2013 and an Assistant Professor, Lecturer, and an Associate Professor in Ibaraki University since 2014, respectively. He is currently an Associate Professor in Graduate School of Natural Science and Technology, Okayama University. His current research interests focus on medical robotics, biological signal processing, and human-machine interface.

Dr. Hideyuki Tonooka

He is a professor at Ibaraki University, Japan. His main research areas are calibration, ground processing, and applications in remote sensing. He has been a member of the project science team for the Advanced Spaceborne Thermal Emission and Reflection radiometer (ASTER), part of NASA's Earth Observing System (EOS), since 1995. He is also joining the Compact Infrared Camera (CIRC) project of the Japan Aerospace Exploration Agency (JAXA).

Extendable ICS Honeypot Design with Modbus/TCP

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

Jun-Hao Lin
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

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: ihliu@cans.ee.ncku.edu.tw, jhlin@cans.ee.ncku.edu.tw, hylai@cans.ee.ncku.edu.tw, jsli@mail.ncku.edu.tw**
www.ncku.edu.tw

Abstract

In order to protect the Cybersecurity of Industrial control system (ICS), we design a prototype of an ICS honeypot. All honeypots are controlled by a server, and using the description file to define honeypot's characteristics, to achieve our honeypot system with scalability and high interaction. We compare our honeypot system and Conpot. the results show that the responses of our honeypot system have more interaction. Even more, our honeypot obtained a perfect score in the honeypot scoring mechanism of Shodan.

Keywords: ICS, Honeypot, Cybersecurity, Shodan.

1. Introduction

With the advancement of communication technology and network transmission, the Industrial Control System (ICS) [1]is no longer limited to local control in order to improve convenience. The industrial control system has moved from a traditional closed operation mode to an Internet connection environment, making the boundary between Information Technology (IT) and Operational Technology (OT) gradually blurred. Although it has increased convenience and many application possibilities, it is bound to face the challenge of network security, especially in key infrastructure facilities that control water resources, electricity, transportation systems and other important fields that maintain the basic functions of the city. They are often the target of hacker attacks. It is important to understand the attack method and formulate the defense method.

To avoid internal attacks from industrial control systems, we assume that the attacker can enter the internal network, in this paper we take Modbus/TCP protocol [2] as an example. Design a set of extensible description files using JSON (JavaScript Object Notation) [3] format to define the characteristics of a honeypot which disguise as a programmable logic controller (PLC) or other ICS devices. And our honeypot system will attract the attacker's attention, record the entire process communication with the attacker, gather information about the attack, to help develop defense strategies.

2. Background

Industrial control system is often widely used in various automated factories and critical infrastructure.In recent decades, to improve the convenience, many industrial control equipment began to support the use of Ethernet, greatly increasing the remote operation and maintenance, but also added a lot of network security hazards.

To understand the attack patterns of attackers, some scholars [7] suggest deploying honeypots on the ICS network to protect the ICS environment.However, with the development of honeypot technology, the attacker's hiding space is oppressed, and the attacker is also in the attack and defense, constantly improve the ability to identify the honeypot, if the fingerprint of the honeypot is mastered by the attacker, the honeypot will not play a substantial role. The Internet of Things search engine-Shodan is a very common and effective tool for identifying honeypots [5] .

2.1. *ICS Honeypot*

ICS honeypot attempts to imitate an ICS device, like a programmable logic controller (PLC). It can deceive or trap attackers, build various baiting hosts, network services or simulation scenarios, capture and analyze the attack actions, and understand the attackers.

Conpot is a low-interactive open-source ICS honeypot developed by The Honeynet Project [4]. It is easy to implement and supports many ICS protocols (like Modbus/TCP, S7comm, EtherNet IP, etc.). However, its disadvantage is that its fingerprint features are obvious and easy to be identified by an anti-honeypot technology such as honeypot detection tools or IoT search engine.

2.2. *SHODAN IoT Search Engine*

Shodan is an IoT search engine commonly used by attackers to reconnaissance some device on the internet. Shodan have a big threat to the cybersecurity of ICS. It can also check whether it is a honeypot or a real control system by Shodan honeyscore [6]. For example, if a honeypot like Conpot uses a default configuration, Shodan can easily recognize it as a fake system, making the honeypot system worthless.

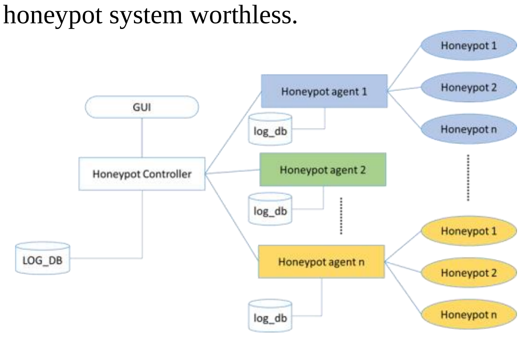


Fig. 1. Honeypot System architecture

3. ICS Honeypot System Design

To solve the common problem of lack of interaction in ICS honeypot, this paper tests and observes the network response of real PLC and tries to imitate its response characteristics. The PLC has a register space for storing sensor values or data processing. This numerical information may express the temperature, humidity, and air pressure of the current industrial environment. To increase the degree of reality, the prototype of this honeypot, taking Modbus/TCP as an example, designed a set of description files using JSON lightweight data exchange language as the definition of honeypot characteristics. The description file is the core of the honeypot behavior. It is used to imitate the network response behavior of real industrial control equipment. The description file defines the honeypot IP address, port number, register address and value, response method, etc., for the user can quickly and flexibly configure the honeypot in the industrial control environment.

3.1. *System Architecture*

The relationship between the honeypot controller, honeypot agent, and honeypots is shown in Fig. 1., There is a graphical user management interface on the honeypot controller, which can transmit honeypot description file to each honeypot agent, control opening or closing of honeypot in each honeypot agent, and view the honeypot log information, etc.

The honeypot agent parses the description files sent from the honeypot controller, generates the corresponding honeypots according to the feature of the description files and records the visitor information (such as IP address, port number, time, behavior, etc.) in the local database, and finally sends it back to the honeypot controller one by one.

The honeypots are application layer programs in OSI model created by a honeypot agent to open protocol services, which is used to interact with the attacker server honeypot.

3.2. *Honeypot Description File Design*

The description file is the core of the honeypot and is the place to define the characteristics of the honeypot. The original idea is that the communication behavior of the equipment in the industrial control environment is simple, and it is often only used to transfer the data between the devices. The communication process is nothing more than HMI reading the sensor value through PLC, or the sensor sending the signal to PLC. Under normal circumstances, the industrial control system flow law, data type is not complex, so it is not difficult to imitate the network communication behavior of industrial control equipment.

After the study of the real industrial control equipment, this paper takes Modbus/TCP, a common protocol in the current industrial control system, as an example, and imitates the surface observed by PLC in the network as the starting point. The simulated items, include PLC equipment model, IP address, port number, address and value of each register, implement of common standard function codes (read or write single or multiple registers), request the implementation of the exception code when an error occurs, and so on.

Finally, the description file is organized into four blocks to simulate the network state of industrial control equipment: Device Info, Pot Type, Default Config, Behavior, and reserve space for expansion fields in each block, can be used to make up for other special undefined conditions.

3.3. *Modbus/TCP Implementation Method*

Different from other ICS honeypots, our ICS honeypot not only provides industrial protocol services, but also focuses on improving the authenticity of its responses. According to the specifications of Modbus official document [2], we use Socket to reproduce the honeypot system that supports Modbus/TCP protocol.First, the content of the description file is parsed according to the response mode set by the user, and the honeypot service is opened at multiple sites. The second step is to determine whether it is a Modbus/TCP request. The third step is to determine whether the data is abnormal and whether the data exceeds the interval. These processes Afterwards, it will reply to the attacker based on the response content specified in the description file. If it does not meet the above, it will reply to the attacker's corresponding exception code, and store the log in the database.

4. Experiment and Result

We compare the network response differences between the MASTek environmental control model, Conpot, and the prototype of our industrial honeypot system, and using Shodan’s scoring mechanism for honeypots-honeyscore, can Shodan recognize the two honeypots as ICS or honeypot devices?

Then, the results of exploring our honeypot by Shodan can be seen that our honeypot is an equipment of industrial control system in the view of Shodan, and our honeypot can adjust the exposed equipment information on port 502. Using the Honeyscore scoring mechanism of Shodan, and the evaluation result indicated that the honeypot system might be real, and the Honeyscore was 0.3(honeypot would be determined if the Honeyscore was above 0.5).

Similarly, Shodan is used to actively scan Conpot’s IP. The features of Conpot have been completely mastered by Shodan and label it as a honeypot on the industrial control system. In the Honeyscore evaluation, the score even came to 1.0, that is, Shodan is fully confident that it is a honeypot device. In other words, this kind of trap has been difficult to lure an attacker.

5. Conclusion

In this paper, a honeypot description file framework was designed to imitate the characteristics of industrial control system’s equipment and implemented by taking Modbus/TCP protocol as an example. Our honeypot system has the characteristics of high expansion. The honeypot agents in different industrial environments can be centrally controlled by one honeypot controller, and more than one honeypot device can be deployed on each agent, which makes it a complete honeynet system.

Our honeypot system compared to Conpot is also very interactive in the network response. The response content is based on the characteristic response defined by description file, which makes it equivalent to the response of real industrial control equipment.

The authenticity of our honeypot system was verified on the Shodan Internet of Things search engine. The results show that our honeypot system can avoid the inspection mechanism of Shodan for ICS honeypot, Shodan label our honeypot system as ICS device. In terms of authenticity, our honeypot system can also perform well.

Acknowledgements

This work was supported by the Delta Electronics, Inc.' Industry and Academic cooperation project - Design and Prototype of Hybrid Industrial Control System Honeypot and the Ministry of Science and Technology (MOST) in Taiwan.

References

1. Kevin E. Hemsley, Dr. Ronald E. Fisher, History of Industrial Control System Cyber Incidents, Idaho Falls: Idaho National Laboratory, 2018.
2. ModbusOrganization, "Modbus_Application_Protocol_V1_1b3," 2021. [Online]. Available: https://modbus.org/docs/Modbus_Application_Protocol_V1_1b3.pdf
3. Wikipedia, "JSON," [Online]. Available: JSON - Wikipedia. [Accessed 26 05 2021].
4. A. Jicha, M. Patton and H. Chen, "SCADA honeypots: An in-depth analysis of Conpot," 016 IEEE Conference on Intelligence and Security Informatics (ISI), 2016.
5. R. Bodenheim, J. Butts, S. Dunlap and B. Mullins, "Evaluation of the ability of the Shodan search engine to identify Internet-facing industrial control devices," international Journal of Critical Infrastructure Protection, vol. 7, no. 2, pp. 114 - 123, 2014.
6. SHODAN, "Honeypot Or Not?," 2021. [Online]. Available: https://honeyscore.shodan.io/. [Accessed 27 05 2021].
7. Venkat Pothamsetty, Matthew Franz, "SCADA HoneyNet Project: Building Honeypots for Industrial Networks," 2004. [Online]. Available: http://scadahoneynet.sourceforge.net/ [Accessed 27 05 2021].

Authors Introduction

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. Jun-Hao Lin

He received his B.S. degree from the Department of Electrical Engineering, National Taipei University of Technology, Taiwan in 2019. He got the M.S. degree in National Cheng Kung University in Taiwan. His research focuses on network communication and cyber security.

Mr. Hsin-Yu Lai

He received his B.S. degree from the Department of Electrical Engineering, National Chung Cheng University, Taiwan in 2019. He is acquiring the master's degree in Department of Electrical Engineering / Institute of Computer and Communication Engineering, National Cheng Kung University in Taiwan.

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.

Industrial Control System Cybersecurity Testbed with TSN Feature

I-Hsien Liu [1]
Department of Electronic Engineering/Institute of Computer and Communication Engineering,
National Cheng Kung University, No. 1, Daxue Road, East District
Tainan, 701401, Taiwan

Li-Yin Chang [2]
Department of Electronic Engineering/Institute of Computer and Communication Engineering,
National Cheng Kung University, No. 1, Daxue Road, East District
Tainan, 701401, Taiwan

Jung-Shian Li* [3]
Department of Electronic Engineering/Institute of Computer and Communication Engineering,
National Cheng Kung University, No. 1, Daxue Road, East District
Tainan, 701401, Taiwan

Chuan-Gang Liu [4]
Department of Applied Informatics and Multimedia,
Chia Nan University of Pharmacy & Science, No. 60, Sec. 1, Erren Rd., Rende Dist., Tainan City 717301, Taiwan
E-mail: ihliu@cans.ee.ncku.edu.tw[1],lychang@cans.ee.ncku.edu.tw[2], jsli@cans.ee.ncku.edu.tw[3], chgliu@mail.cnu.edu.tw[4]
www.ncku.edu.tw[123], www.cnu.edu.tw[4]

Abstract

Due to the advent of Industrial automation and the Industrial Internet, information security attacks on the industrial environment have emerged one after another. In order to conduct better research and protection against industrial control attacks, we have built a test platform for industrial control networks for related research. We have built equipment that supports time-sensitive networks in this field to conduct research on Cybersecurity with time-sensitive networking.

Keywords: ICS, Networking, Cybersecurity, TSN.

1. Introduction

In terms of IT communication in Industrial Control System (ICS), Ethernet is commonly used in ICS and IT. However, due to the development of Industry 4.0 and smart manufacturing, the periodic requirements for the network are getting shorter and shorter. The traditional Ethernet system cannot meet the real-time requirements due to random media access and Best effort (BE) forwarding mechanism [1]. Therefore, It is difficult to ensure the timing behavior of critical traffic under these circumstances and to provide isolation from noncritical traffic. In order to ensure the security and real-time performance of critical traffic, we have established an industrial control system testbed [2] and support TSN(Time sensitive networking) equipment and technology and successfully utilize TSN standard to ensure the security and time of critical traffic. The devices are from Intel, Cisco, NI.

2. Background

In modern factories and smart manufacturing, ICS controls many devices and controllers, and there are many different communication protocols between controllers and devices. For example, MODBUS TCP Ethernet/IP, OLE in Process Control Unified Architecture (OPC-UA), IEEE 1722, Object Management Group (OMG) Real-time System Data Distribution Service (DDS) [3]. These protocols can support the extension of TSN to meet all the requirements of real-time Ethernet because of the characteristics of Ethernet, while making Ethernet transmission more reliable, reducing jitter and shortening delay.

2.1. *TSN-standard*

In order to solve the problem of ensuring that the delay behavior of critical traffic is isolated from general traffic, The IEEE 802.1 working group defined a new and enhanced set of standards, namely Time Sensitive Networking. It is an extension of IEEE 802.1 Ethernet, a series of new specifications established by the Time Sensitive Network Task Group of the IEEE 802.1 Working Group on the basis of existing standards as shown in Table 1 below.

Table 1. IEEE TSN Primarily Standard **[4]**.

TSN standard	Standard description
802.1Qcc	Network management
802.1Qbv	Scheduled traffic
802.1Qav	Credited based shaper
802.1Qcb	Frame replication
802.1AS	Timing and synchronization
802.1Qbu	Frame preemption
802.1Qca	Path control and reservation

In our ICS testbed, we mainly focus on the research and result analysis of 802.1Qbv and 802.1Qav. Below we will mainly introduce several protocols used on the testbed:

- 802.1AS: In the TSN system, time synchronization is the most important part. All devices must be synchronized to the same clock. 802.1AS is an enhanced version of the PTP time synchronization protocol. Compared with the general PTP, 802.1AS has only one central clock, and the rest are auxiliary clocks, and packets can only be transmitted in a synchronized time domain. [5].
- IEEE802.1Qbv: In order to achieve the coexistence of various priority flows in the same network and have available separate bandwidth and end-to-end delay specifications, 802.1Qbv defines the mechanism for packet forwarding in the switch, which uses Time Aware Shaper (TAS) to send packets in the different queue [6].Figure 1 shows the 802.1Qbv example.

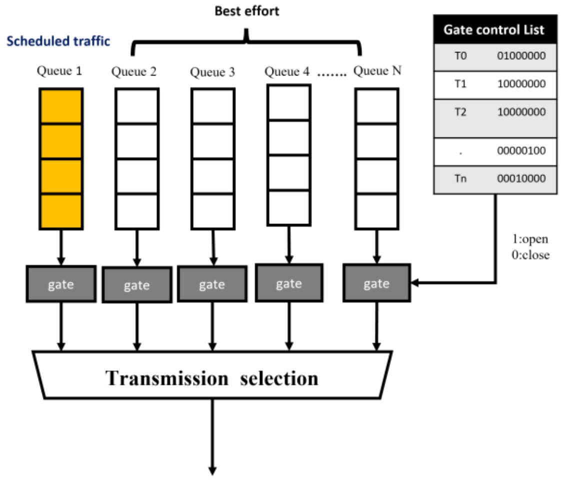


Fig. 1.802.1Qbv example [6].

- IEEE 802.1Qcc: 802.1Qcc defines the management configuration in the TSN network, which defines the distributed configuration and the centralized configuration, Figure 2 shows an example of the centralized configuration, The central network controller (CNC) will manage all the bridge ends in the network. The CNC mainly masters all network topology information, is responsible for calculating the transmission delay of the link, including the packet size and quantity, and then calculates a guarantee to meet Deterministic transmission schedule of traffic demand [7].
- IEEE 802.1Qav :Use credit based shaper to ensure that traditional asynchronous Ethernet data traffic will not interfere with AVB's real-time audio and video streams. it provides bounded latency per stream type. It has been developed for professional audio and video applications, and the major application can still be seen in such streams.

3. TSN Testbed Scenario

We setup a TSN testbed with the device from multiple vendors, we use cisco IE4000 as our bridge because ie4000 supports 802.1as and 802.1qbv. In the part of the speaker and the receiver, we use a computer with an I210 network interface card,because the support of 802.1Qbv, 802.1AS, and 802.1Qav. Figure 3 shows our testbed scenario. In our testbed, we use the ptp4l library in the linux to implement 802.1AS to sync all the device manually because cisco CNC can only sync the bridge.

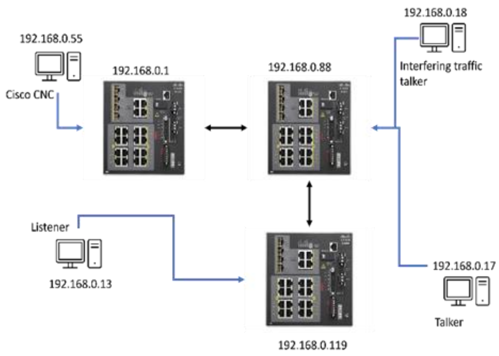


Fig. 2. TSN testbed scenario

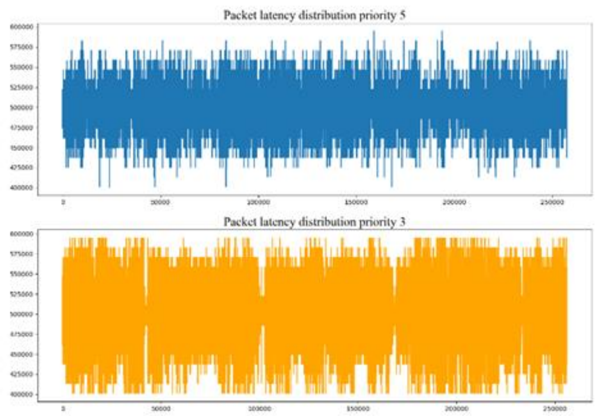


Fig. 3 Delay result without 802.1Qbv

4. Results

In our experiment, we analyze the traffic interference of UDP DOS attacks with Qbv and Qav flow. In the Qbv scenario, we set a cycle time period to 1 millisecond, cut into 10 time slices, the first and the fifth time slices we send BE traffic, and the second and sixth time slices send priority 5 traffic. The third and seventh time slices are priority 3 traffic, so we want the critical priority traffic latencies to be 0.5ms. During the experiment, we use UDP dos attack to interfere with traffic In the experiment, we use UDP dos attack to interfere with traffic.

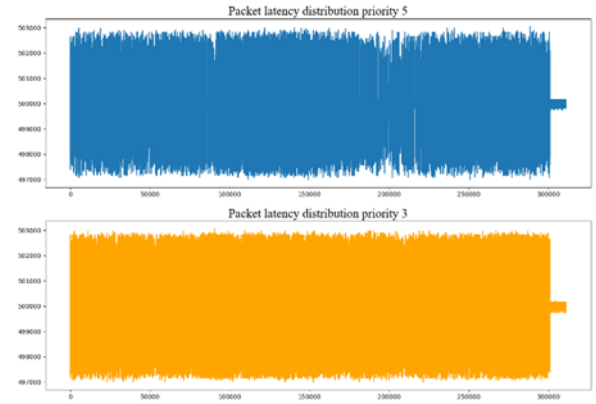


Fig. 4. Delay result with 802.1Qbv

Figure 4 shows the result of not using the 802.1Qbv flow isolation mechanism. It can be seen that jitter is very large even up to 100000ns, therefore the delay of critical traffic cannot be guaranteed.

In Figure 5, it can be seen that the effect of controlling jitter after using 802.1Qbv. Traffic isolation works perfect. Even if the malicious activities happened, jitter was still controlled under 3000ns.After 300,000 packets, we turn off the dos attack and we can see that the delay control can be within a few hundred ns.

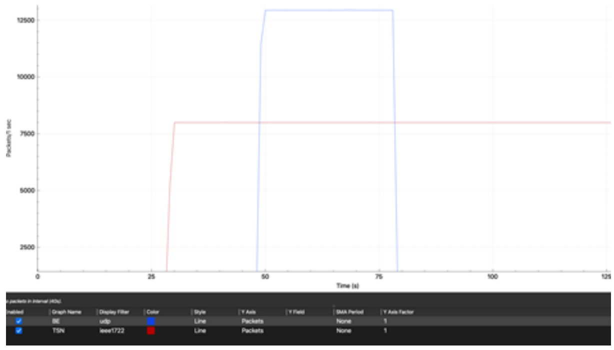


Fig. 5 Qav result with CBS

In the QAV experiment, we will produce IEEE1722 audio packets mixed with normal BE packets. We will transmit 8000 packet/second SR Class A audio frames as represented by the red line in the figure 5. Best effort traffic is in color blue. We generated interfering traffic using the iperf3 tool. As the Fig 5 display, the IEEE 1722 audio frames are consistently 125 µs delta time apart shows that the interference to BE traffic is well controlled.

5. Conclusion

From the experimental results, we can see the two abovementioned standards control the delay to protect critical traffic from jamming and exhibit an encouraging effect on DOS attacks. The actual delay is also related to the CPU computing power and kernel of the hardware device. Our testbed will also continue to be used for conducting more cyber security research on TSN.

Acknowledgements

This work was supported by the Ministry of Science and Technology (MOST) in Taiwan under contract numbers MOST 110-2218-E-006-013-MBK-.

References

1. J.-D. Decotignie, "Ethernet-Based Real-Time and Industrial Communications," Proceedings of the IEEE, vol. 93, no. 6, pp. 1102 - 1117, 6 JUNE 2005.
2. Sameer Chouksey, Hariram Selvamurugan Satheesh, Johan Åkerberg, "Coexistence, An Experimental Study of TSN-NonTSN," in 2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC), NV, USA, 2021.
3. V. GOLLER, "Time Sensitive Networks For Industrial Automation Systems," Analog Devices, 2016. [Online]. Available: [Accessed 11 12 2021].
4. IEEE, Time-Sensitive Networking (TSN) Task Group, IEEE, 2016.
5. "IEEE P802.1AS-Rev/D8.0, Draft Standard for Local and Metropolitan Area Networks—Timing and Synchronization for Time Sensitive Applications," 2019. [Online]
6. IEEE, IEEE Std 802.1Qbv - Enhancements for Scheduled Traffic, IEEE, 2016.
7. Cisco, Time-Sensitive Networking:A Technical Introduction, U.S.: Cisco, 2017.

Authors Introduction

Dr. I-Hsien Liu

He is a research fellow in the Taiwan Information Security Center @ National Cheng Kung Univ. (TWISC@NCKU) and Dept. of Electrical Eng., National Cheng Kung Univ., Taiwan. He obtained his PhD in 2015 in Computer and Communication Eng. from the National Cheng Kung Univ.. His interests are Cyber-Security, Wireless Network, Group Communication and Reliable Transmission.

Mr. Li-Yin Chang

He received his B.S. degree from the Dept. of Communiciation Eng., National Chung Cheng Univ., Taiwan in 2020. He is acquiring the master's degree in Dept. of Electrical Eng./Institute of Computer and Communication Eng., National Cheng Kung Univ. in Taiwan.

Prof. Jung-Shian Li

He is a full Prof. in the Dept. of Electrical Eng,, National Cheng Kung Univ., Taiwan. He graduated from the National Taiwan Univ., Taiwan, with B.S. in 1990 and M.S. degrees in 1992 in Electrical Eng.. 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 Univ.. He serves on the editorial boards of the International Journal of Communication Systems.

Prof. Chuan-Gang Liu

He is an Associate Prof. in the Dept. of Applied Informatics and Multimedia, Chia Nan Univ. of Pharmacy and Science. He received the B.Sc. degree from the Dept, of Electrical Eng., Tam Kang Univ., in 2000. Then he graduated from the National Cheng Kung Univ. with M.S. and PhD degrees in Electrical Eng.QQ. His research interests are in the areas of optical networks control, wireless networks, EPON, VANET, network security, cloud computing and TCP performance analysis.

Using the Modified Delphi Method to Construct the Quality Indicators of the Counseling Service System

Li-Min Chuang

The Department of International Business, Chang Jung Christian University,
No. 1, Changda Rd., Gueiren District, Tainan City, 711301, Taiwan

Hsiu-Hao Liu

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: liming@mail.cjcu.edu.tw, harry2356969@gmail.com

Abstract

The counseling service system is a case management tool used by professional guidance counselors and full-time teacher-counselor. However, the system satisfaction and willingness to use were generally low. This study used the modified Delphi method to construct the quality indicators of the counseling service system. The study found the quality indicators of the system can be divided into five major dimensions and 23 evaluation indicators to explore. There were three indicators most valued by experts and users, which were the functions are simple and convenient to operate, the project meets the needs of the work, and the efficiency and convenience of paper processing are improved. Based on the research results, this study puts forward substantive implications for academics and management.

Keywords: System Quality Indicator, Modified Delphi Method, Counseling Service System

1. Introduction

The counseling service system developed by the Tainan Students Guidance and Counseling Center is a case management website, which is intended to assist professional counselors and full-time counselors in the school's three-level counseling. Caring for students and tracking their developmental effects and achieve the role of transmitting information between the center and schools. The main function of the e-counseling service system is to record case data, consultation records, referrals and other operations. However, past studies have found that, Elementary and secondary teachers' satisfaction and willingness to use are both low. The main influencing factors come from the user's perceived usefulness of the system and poor system quality,[1,2] in view of the fact that the sample of relevant research in the past was based on elementary school teachers, there are still other relevant units' opinions that have not been presented. This study attempts to fill in the evaluation basis for the quality of the counseling service system by adding expert suggestions from different backgrounds. Based on the above background, this study conducted interviews with experts on the current status of the system and used the modified Delphi method (MDM) to perform indicator analysis, with a view to constructing quality indicators for the counseling service system for management units As a basis for improving system gaps, improving user satisfaction and willingness to use.

2. Literature Review

2.1 System quality

The service quality point of view regards the organization as the goal of providing customers with high quality services and has a collection of multiple processes.[3] Its definition is based on the comparison between the service expected by users and the service they feel,[4] and service quality can be applied to the information system functions, because the information system can be regarded as a service function to deal with the information needs of the organization. Therefore, this study defines system quality as "the information system provider delivers services to users through the website environment, and the user's overall evaluation of the service."

Zeithaml et al. proposed a conceptual framework for online service quality, which measures website service quality in 11 dimensions.[5] These 11 dimensions were readjusted into seven dimensions in 2002, including efficiency, reliability, completion, privacy, responsiveness, compensation, and contact. Among them, the four items of efficiency, reliability, completeness and privacy form the e-SERVQUAL scale, which mainly measures customer satisfaction with the online functions provided by the website, and the three items of responsiveness, compensation and contact form the recovery e-SERVQUAL scale is mainly used for the service that customers hope to get when they have questions or encounter problems.[6]

Since then, there have been many studies to modify the measurement indicators of online or website service quality. For example, the Electronic Retailing Quality Scale (E-TailQ) uses 14 items to evaluate the quality of e-tailing commerce;[7] there is also e-Core Service Quality Scale (ES-Qual) (including Efficiency, fulfillment, system availability and privacy etc. four dimensions) and e-Recovery Service Quality Scale (E-RecS-Qual) (including responsiveness, compensation and contact and other three dimensions), both of which are a scale that can measure the performance of website services.[8] Among them, the seven facets extracted by Parasuraman et al. [8] are roughly the same as Zeithaml et al.,[6] Parasuraman et al. [8] only changed reliability to system availability, but the definition of the two is the same.

In summary, whether it is the E-TailQ scale or the ES-Qual scale, both are based on rigorous statistical analysis and empirical research and have been verified in subsequent studies since their development,[9,10] At the same time, it is consistent with the theme of this research. According to this, this research adopts the structure of the two as the basis of the prototype of the system quality index.

2.2 The prototype of the system quality indicators

This study refers to the system service quality[9,10] and the related research of elementary school full-time teacher-counselor on the counseling service system,[1] and Initially summarize the system quality indicators into five dimensions: tangibility, reliability, responsiveness, caring and certainty, and develop 23 system quality assessment items (Table 1).

Table 1. A summary of the quality evaluation indicators of the counseling service system

Dimensions	No.	Indicators
Tangibles	A1	The proper layout
	A2	Information on all options is available
	A3	Easy to learn how to use
	A4	The function is simple and convenient to operate
	A5	It is easy to find the information you need
Reliability	B1	The project is rich and complete
	B2	The project meets the needs of the job
	B3	Improve the confidentiality mechanism
	B4	Stable use and operation
	B5	Web-to-web connections are fast and smooth
Assurance	C1	Helps communicate work matters
	C2	Helps improve performance
	C3	Improve the efficiency and convenience of paperwork
	C4	Reduce the cost of physical information or archives
Responsiveness	D1	Be able to quickly resolve questions and special needs
	D2	Can inform the service response time
	D3	Have good professional training
	D4	There are regular maintenance personnel
	D5	Website system failures are dealt with immediately
Empathy	E1	The expression of the web page is clear
	E2	Protection of information
	E3	Be able to understand the problem quickly
	E4	Have a good service attitude

3. Research Methods

3.1 Modified Delphi method

Murry and Hammons took to correct the typical German illegality is to use literature review to organize and develop a prototype questionnaire, instead of the typical Delphi method's open questionnaire as the first round of survey, that is the MDM.[11] Based on the use of the MDM, it can fully reflect the opinions of experts, brainstorm ideas, and have high accuracy.[12]

3.2 Opinion consistency and stability evaluation

In this study, two standards of quartile difference (Q.D.) and standard deviation (S.D.) were used to verify the consistency of expert opinions.[13] In the standard of interquartile range, this study adopted the suggestion of Holden and Wedman. When the interquartile range of the question is less than or equal to 0.6, the expert opinions are highly consistent, and between 0.6-1.00, the expert opinions are moderately consistent, and >1.00 it means that the expert opinions are not unanimous. In the standard deviation standard, when the standard deviation of the item is <1.0, it means that the expert opinions are consistent, and if the standard deviation is >1.0, it means that the expert opinions are not consistent. If the expert opinions of more than 85% of the items agree, the questionnaire will be completed.[14]

The stability of this study was based on the number of changes in the opinions of experts was less than 20%, as the distribution of opinions of the expert group on individual topics reaches the minimum standard of stability.

3.3 Indicator evaluation and suggested directions

This study was based on the prototype of the system quality indicators compiled in the literature, as the first revised Delphi questionnaire, and used the Likert five-point scale to measure the importance of the system quality elements. 1 is divided into "very unimportant", 5 is divided into "very important". According to the consensus of the first round of questionnaires, the second round of questionnaires will be compiled. Repeat the questionnaire by analogy until the questionnaire reaches a stable result. Finally, according to the average and mode to illustrate the importance of the indicators in this stage and ranking, if the scores are the same, ranking with stability.

4. Research Results

4.1 Selection results of the second round of the revised Delphi questionnaire

After completed the first round of questionnaire interviews with 16 experts in related fields, due to the inconsistencies in the indicator items, the second round of interviews was conducted. In the second round, the consensus has been reached according to expert opinions. Therefore, the expert interview phase was ended, and the indicators were formed (Table 2).

Table 2. Results of the first round and second round of interviews

Dimensions	First round						Second round						Overall ranking
	No.	Mode	Mean	Q.D.	S.D.	Rank	No.	Mode	Mean	Q.D.	S.D.	Rank	
Tangibles	A1	4	4.00	0.00	0.63	5	A1	4	3.88	0.38	0.81	5	22
	A2	4	4.19	0.50	0.75	4	A2	5	4.56	0.50	0.51	3	8
	A3	5	4.56	0.50	0.51	2	A3	5	4.63	0.50	0.62	2	6
	A4	5	4.75	0.38	0.45	1	A4	5	4.81	0.00	0.40	1	2
	A5	5	4.31	0.50	0.95	3	A5	5	4.56	0.50	0.63	4	8
Reliability	B1	4	3.88	0.38	0.81	5	B1	4	4.19	0.50	0.75	5	17
	B2	5	4.44	0.50	0.63	1	B2	5	4.75	0.00	0.58	1	3
	B3	5	4.44	0.50	1.03	2	B3	5	4.75	0.38	0.45	2	3
	B4	4	4.31	0.50	0.70	3	B4	5	4.69	0.50	0.48	3	5
	B5	4	4.00	0.88	1.03	4	B5	5	4.31	0.50	0.87	4	15
Assurance	C1	4	3.81	0.88	0.91	4	C1	4	4.00	0.75	0.73	4	21
	C2	4	4.06	0.50	0.93	3	C2	4	4.38	0.50	0.50	3	12
	C3	5	4.31	0.50	1.01	2	C3	5	4.56	0.50	0.51	1	8
	C4	5	4.44	0.50	0.63	1	C4	5	4.44	0.50	0.81	2	11
Responsiveness	D1	4	3.88	1.00	1.09	5	D1	4	4.19	0.50	0.66	3	17
	D2	3	3.50	0.50	0.73	4	D2	4	3.69	0.50	0.95	5	23
	D3	4	4.06	0.88	0.77	3	D3	4	4.06	0.38	0.68	4	20
	D4	4	4.13	0.50	0.72	1	D4	5	4.38	0.50	0.81	1	12
	D5	4	4.06	0.38	0.68	2	D5	4	4.25	0.50	0.58	2	16
Empathy	E1	4	4.50	0.50	0.52	2	E1	5	4.63	0.50	0.50	2	6
	E2	5	4.63	0.38	0.81	1	E2	5	4.88	0.00	0.34	1	1
	E3	4	4.06	0.50	0.85	3	E3	5	4.38	0.50	0.72	3	12
	E4	4	4.00	0.75	0.73	4	E4	4	4.13	0.38	0.62	4	19

After two rounds of questionnaire interviews, we obtained data with expert consistency and stability convergence, and obtained the results of the experiment. It will be evaluated according to five system quality dimensions. In "tangibility", "functional operation is simple and convenient" is the most important; in "reliability", "the project meets the needs of the work" is the most important; in "reliability", to "improve the efficiency of word processing and Convenience is the most important; in "responsiveness", "specialized maintenance" is the most important; finally, in "caring", "protection of information" is the most important, and it is also the highest score of all items. The experts attach great importance to the attitude of the system management unit or industry to protect student data.

From the experimental results, we can see the importance of the major dimensions. Overall, experts attach great importance to the security and rigor of the confidentiality and rigor of the information of the system itself and the administrator, and secondly, whether the system is actually helpful to the work. Since the e-counseling system used by Tainan City at this stage has not been universally recognized by users, the results of this study can be used by its management and maintenance units as a direction for improvement.

5. Conclusion

In addition to the e-counseling system being used in Tainan City, counselors in other counties, cities, or institutions also use similarly functional student counseling record systems, but the designs are different, and the usage situation are also different. This study used the MDM to collect the opinions of experts from various units and obtains an indicator of system quality. In addition to providing the system management unit in Tainan City as a reference indicator for regular maintenance, it can also be used by other government agencies or school institutions to construct and the basis for maintaining the relevant system of student consultation records, in order to take into account the user's point of view and the importance of subsequent maintenance and management in the system development process.

References

1. H.-H. Liu. A Study of the Counseling Service System Satisfaction for Teacher-Counselor in Tainan Elementary School. *Journal of Robotics, Networking and Artificial Life*, **8**(3), 2021, pp. 197-200.
2. H.-H. Liu. Assess the Critical Factors for the Counseling Service System Usage Intention, *Proceedings of the 2021 International Conference on Artificial Life and Robotics*, Japan, January 2021.
3. B.L. Myers, L.A. Kappelman, V.R. Prybutok. A Comprehensive Model for Assessing Quality and Productivity of the Information Systems Function: Toward a Theory for Information Systems Assessment. *Information Resources Management Journal*, **10**(1), 1997, pp. 6-25.
4. A. Parasuraman, V.A. Zeithaml, L.L. Berry. A Conceptual Model of Service Quality and Its Implications for Future Research. *Journal of Marketing*, **49**(4), 1985, pp. 41-50.
5. V.A. Zeithaml, A. Parasuraman, A. Malhotra. A Conceptual Framework for Understanding E-Service Quality: Implications for Future Research and Managerial Practice. Mass: Marketing Science Institute, Cambridge, 2000.
6. V.A. Zeithaml, A. Parasuraman, A. Malhotra. Service quality delivery through web sites: A critical review of extant knowledge. *Journal of the Academy of Marketing Science*, **30**(4), 2002, pp. 362-375.
7. M. Wolfinbarger, M.C. Gilly. eTailQ: Dimensionalizing, measuring and predicting etail quality. *Journal of Retailing*, **79**(3), 2003, pp. 183-198.
8. A. Parasuraman, V.A. Zeithaml, A. Malhotra. E-S-QUAL: A multiple-item scale for assessing electronic service. *Journal of Service Research*, **7**(3), 2005, pp. 213-233.
9. M. Laxmi. A Study of Reviewing Country and Industry Specific E-Service Quality Dimensions. *Journal of Critical Reviews*, **7**(16), 2020, pp. 3741-3748.
10. P. Rita, T. Oliveira, A. Farisa. The impact of e-service quality and customer satisfaction on customer behavior in online shopping. *Heliyon*, **5**(10), 2019, e02690.
11. J.W. Murry, J.O. Hammons. Delphi: A Versatile Methodology for Conducting Qualitative Research. *The Review of Higher Education*, **18**(4), 1995, pp. 423-436.
12. S. Thangaratinam, C. Redman. The Delphi Technique. *The Obstetrician & Gynaecologist*, **7**(2), 2005, pp. 120-125.
13. K.W. Brooks. Delphi Technique: Expanding Applications. *North Central Association Quarterly*, **53**(3), 1979, pp. 377-385.
14. M.C. Holden, J.F. Wedman. Future Issues of Computer-mediated Communication: The Results of a Delphi Study. *Educational Technology Research and Development*, **41**(4), 1993, pp. 5-24.

Authors Introduction

Dr. Li-Min Chuang

He received the Ph.D. degree from the National Cheng Kung University. He is currently an associate professor and department head at the Department of International Business at Chang Jung Christian University, Taiwan. His research interests focus on strategic management, innovation management, and service innovation.

Mr. Hsiu-Hao Liu

He received his Master's degree from the Department of Psychology, National Chung Cheng University, Taiwan in 2015. He is currently a Doctoral course student in Chang Jung Christian University, Taiwan. His research interests include the application of information technology to education and counseling.

The key success factors of introducing ERP system in Taiwan's manufacturing industry

Li-Min Chuang
The Department of International Business, Chang Jung Christian University, No. 1, Changda Rd., Gueiren District, Tainan City, 711301, Taiwan

Yu-Po Lee
The Ph.D. Program in Business and Operations Management, College of Management, Chang Jung Christian University, No. 1, Changda Rd., Gueiren District, Tainan City, 711301, Taiwan
Email: limimg@gmail.com, davidlitw@yahoo.com

Abstract

This study examines the key success factors of ERP introduction in Taiwan's manufacturing industry as a reference for enterprises to reduce the high and unforeseeable financial and time costs of ERP introduction. In the second phase, a questionnaire study was conducted using the Analytic Hierarchy Process to extract the relative weights of the distance between primary and secondary dimensions, and 5 primary dimensions were derived, including "Management/Organization", "Introduction Process", "Technical Support", "Documentation", and "Personnel", as well as 15 secondary dimensions.

Keywords: Manufacturing industry, Enterprise Resource Planning (ERP), Likert scale, Analytic Hierarchy Process (AHP)

1. Introduction

Taiwan has an island-based foreign trade economy, and its manufacturing industry is faced with energy deficiency, rising wages, and mounting environmental awareness, as well as competition from other countries, which all lead to operational difficulties; as an essential component in the transformation of enterprises, information management helps to increase productivity, improve product quality, and boost profitability.

The manufacturing sector is the mainstay of Taiwan's economy during the Covid-19 pandemic, with manufacturing output reaching NTD 3.5581 trillion in Q1 2021, which is an annual increase of 14.62%, the largest rise since Q2 2020, and positive growth for 2 consecutive quarters[12].

In recent years, business management systems have realized digitalization, and its application in information systems (IT) has drawn considerable attention in research works. One of these systems, the Enterprise Resource Planning (ERP) system, has become an important instrument for senior managers to control and conclude transactions[5]. Therefore, the smooth and rapid introduction of ERP systems in the manufacturing industry has developed into a key subject for the Taiwanese manufacturing industry.

This study combined the relevant literature on ERP introduction and established five primary dimensions and three items of ERP introduction for Taiwan's manufacturing industry with a total of 15 sub-dimensions. It is intended that the findings of this study will serve as an important reference for ERP introduction in Taiwan's manufacturing industry.

2. Literature Review

Enterprise Resource Planning (ERP) systems are often defined as a key managerial tool for planning the resources and transactions of an enterprise. The essential feature of this system is the storage, administration, and application of information to the plans and outcomes of

business operations. Weill and Ross observed that the financial performance of enterprises with excellent Information Technology Governance (ITG) are superior to those with poor IGT[11], which suggests that informatization has a positive effect on the financial performance of firms, and literature review has confirmed that the introduction and operation of ERP have been a key factor in the success of enterprises[3].
According to statistics, nearly half (45%) of the existing ERP systems will be replaced by the new ERP systems available in the market today[9]. The findings of this study can provide an important basis for the initial introduction, and serve as a reference when upgrading or updating systems.
During the introduction of an ERP system, there are a number of issues that enterprises may encounter. In addition to the support and active involvement of the enterprise's senior management[6], the selection of consultants and system vendors is also required to ensure that all operators have the relevant skills to safeguard the effective and continuous operation of the system. Enterprises should continue to educate their employees[1] and arrange training for key personnel to acquire the relevant knowledge[2]. Meanwhile, appropriate software should be selected to avoid unpredictable damage and allow room for future updates. Furthermore, managers should encourage all employees to participate in the introduction of the system, in order that all employees can contribute and be united in its application and use[10].

3. Key Model Building for ERP Introduction

This study adopted a quantitative approach, conducted a literature review of five journal papers on ERP introduction in enterprises, constructed 5 primary and 30 secondary dimensions, and took enterprises that had successfully introduced ERP systems as the subject[1,4,5,7,8]. A five-point Likert scale was developed as a prediction questionnaire to extract three out of the five primary dimensions, a total of 15 secondary dimensions were obtained, and then, the weights of each dimension were calculated using AHP to obtain the key success factors for the introduction of the ERP system in Taiwan's manufacturing industry.

Table1. Results of the Overall Assessment of Key Success Factors for the Introduction of Enterprise Resource Planning (ERP) in Taiwan's Manufacturing Industry

Dimension	Weight	Rank	Evaluation Indication	Weight	Rank	Weight	Rank
A Management/ Organization	0.1961	3	A1. Commitment of Business Owners and Executives	0.3635	2	0.0672	6
			A2. Awareness of Business Owners and Executives	0.1568	3	0.0288	10
			A3. Involvement of Business Members	0.4798	1	0.1011	3
B Introduction Process	0.2203	2	B1. Organization of Resources	0.1649	2	0.0412	8
			B2. Completeness of Documentation	0.1318	3	0.0251	12
			B3. Appointment of Professional Consultants	0.7033	1	0.1541	2
C Technical Support	0.1206	4	C1. Hardware Equipment Support	0.0908	3	0.0126	14
			C2. Software Equipment Support	0.2993	2	0.0292	9
			C3. System Administration Adequacy	0.6100	1	0.0789	4
D Documentation	0.0413	5	D1. Primary Document	0.1084	3	0.0042	15
			D2. Completeness of Data Structure	0.3464	2	0.0135	13
			D3. Maintenance and Completeness	0.5452	1	0.0236	11
E Personnel	0.4207	1	E2. Training	0.6804	1	0.2849	1
			E5. Project Lead	0.1699	2	0.0712	5
			E6. Clear Objectives	0.1496	3	0.0627	7

4. Empirical Analysis

With the three highest scoring sub-dimensions and primary dimensions of the questionnaire as the actual measurement samples, 10 AHP questionnaires were returned from the middle and senior level managers in charge of ERP introduction in 10 enterprises. After consistency testing, the C.I. values are all $\leqq 0.1$, thus, the 10 AHP questionnaires are all valid.Regarding each "primary dimension" and "secondary dimension", this study applied Microsoft Excel to calculate the relative weights. Based on the result, the connotation of each weight indicator is explained.

5. Conclusion and Suggestions

5.1. Conclusions and Propositions

The findings of the study "Key Success Factors for the Introduction of Enterprise Resource Planning (ERP) in Taiwan's Manufacturing Industry" are summarized below in order of weighting:
Proposition 1: In the "Key Success Factors for ERP Introduction in Manufacturing Industry", the "Personnel" dimension is more crucial than the "Introduction Process", "Management/Organization", "Technical Support", and "Documentation" dimensions, thus, it is recommended that business owners and executives should implement various staff education and training programs to develop capable and responsible employees.
Proposition 2: In the "Management/Organization" dimension, the "Involvement of Business Members" dimension is more crucial than the "Commitment of Business Owners and Executives" and "Awareness of Business Owners and Executives" dimensions, which demonstrates that the involvement of all members of the enterprise, and their perseverance, are key issues during the introduction process, as well as critical indicators for the success of the organization.
Proposition 3: In the "Introduction Process" dimension, the "Appointment of Professional Consultants" dimension is more crucial than the "Organization of Resources" and "Completeness of Documentation" dimensions, thus, a competent professional consultant is an essential resource for planning, application, and system use experience.
Proposition 4: In the "Technical Support" dimension, the "System Administration Adequacy" dimension is more crucial than the "Software Equipment Support" and "Hardware Equipment Support" dimensions, as a stable and fit-for-purpose system can prevent unpredictable failures during the introduction process, and a secure and stable system can also bring competitive advantage.
Proposition 5: In the "Documentation" dimension, the "Maintenance and Completeness" dimension is more crucial than the "Completeness of Data Structure" and "Primary Document" dimensions, which indicates that the maintenance of documentation has a significantly positive impact on the system's introduction.
Proposition 6: In the "Personnel" dimension, the "Training" dimension is more crucial than the "Project Lead" and "Clear Objectives" dimensions, which corresponds to the statement that well-trained staffs are the most important asset of an enterprise, thus, routine education and conducting professional training designed for each project are also key factors in the introduction of the system.

5.2. Recommendations for Subsequent Researchers

This study examined relevant literature and employed a Likert scale as a pre-test to determine the content of the questionnaire. Then, AHP was applied to construct the evaluation method to gather the opinions of industry executives and generalize the key success factors for ERP introduction as a reference for those interested in implementing the system:

5.2.1. *Identifying the subject of the study for more accurate findings*

The subject of this study was the enterprises that introduced the system in the manufacturing industry. This study outlined the respective dimensions and weights, and the subjects were mainly corporate executives; although the success of ERP introduction is very much related to the enterprise itself, it is recommended that subsequent researchers can select other subjects for study by including other related parties (e.g., MIS of the enterprise, consultants of the counseling services, or relevant academic experts).

5.2.2. *Exploring the impact on business continuity with the findings of this study*

For enterprises, ERP introduction is just the inception, meaning, it is the introduction outcome that makes the real contribution to the enterprise. In order to achieve continuity in system performance, enterprises should continue to introduce new skills through employee education and training[11]. Depending on the philosophy and ambition of the business owner, the scale of subsequent ERP introduction and development will vary, as will the value created for the enterprise. It is recommended that subsequent researchers should conduct follow-up studies on the effectiveness of ERP introduction on business operations. As system introduction is only the beginning, practical application reference value is expected to deliver more results.

References

1. Albert Y.T. Sun., Abe Yazdania., John D. Overendb. Achievement assessment for enterprise resource planning (ERP) system implementations based on critical success factors (CSFs). Int. J. Production Economics (2005) 189–203.
2. D. Sedera, G.G. Gable, Knowledge management competence for enterprise system success, J. Strat. Inform. Syst. 19 (4) (2010) 296–306.
3. H.C. Lai, S.Y. Hwang, T.P. Liang, S.Y. Hung, H.L. Wu, A survey of journal publications of MIS departments in Taiwan, Acad. Res. Fields Depart. Human. Soc.Sci. 12 (2) (2011) 58–68 [in Chinese].
4. H.-J. Li., S.-I Chang., D. C. Yen. Investigating CSFs for the life cycle of ERP system from the perspective of IT governance. Computer Standards & Interfaces ,50 (2017) 269–279.
5. Karoline B. Osnes, Julie R. Olsen, Polyxeni Vassilakopoulou, Eli Hustad. ERP Systems in Multinational Enterprises: A literature Review of Post-implementation Challenges. Procedia Computer Science. 138 (2018) 541–548.
6. Lorraine Jean Staehr. Understanding the role of managerial agency in achieving business benefits from ERP systems, Inform. Syst. J. 20 (3) (2010) 213–238.
7. Mahmood Ali., Lloyd Miller., ERP system implementation in large enterprises - a systematic literature review. Journal of Enterprise Information Management. 30 (4) (2017) 666-692.
8. Mustafa Ağaoğlu., E. Serra Yurtkoru, Aslı Küçükaslan Ekmekçi. The effect of ERP implementation CSFs on business performance: an empirical study on users' perception. Procedia - Social and Behavioral Sciences 210 (2015) 35–42.
9. A Panorama Consulting Solutions Research Report, 2013 Erp Report: Organizational Change and Business Process Management. 2013 Panorama Consulting Solutions.
10. P.B. Seddon, C. Calvert, S. Yang, A multi-project model of key factors affecting organizational benefits from enterprise systems, MIS Q. 34 (2) (2010) 305–328.
11. S.V. Grabski, S.A. Leech, P.J. Schmidt, A Review of ERP Research: a future Agenda for accounting information systems, J. Inform. Syst. 25 (1) (2011) 37–78.
12. Statistics Department of the Ministry of Economic Affairs, Executive Yuan, Taiwan's first quarter manufacturing output statistics. (2021)

Authors Introduction

Dr. Li-Min Chuang

He received the Ph.D. degree from the National Cheng Kung University. He is currently an associate professor and department head at the Department of International Business at Chang Jung Christian University, Taiwan. His research interests focus on strategic management, innovation management, and service innovation.

Mr. Yu-Po Lee

He works as the manager in the Wei Shin Aluminum Company for more than 30 years. He received his Master's degree from the EMBA, Chang Jung Christian University in 2012. He is currently studying for a doctoral program at Chang Jung Christian University in Taiwan.

The Fuzzy AHP approach for intelligent building assessment model

Li-Min Chuang

The Department of International Business, Chang Jung Christian University,
No. 1, Changda Rd., Gueiren District, Tainan City, 711301, Taiwan

Yu-Po Lee

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

Chien-Chih Kuo

Chien Chang Construction Co., Ltd.
No. 88, Wencheng Rd., North Dist., Tainan City, 702029, Taiwan
Email: liming@gmail.com, davidlitw@yahoo.com, kenkuo68@hotmail.com

Abstract

The main objective of this thesis is to probe into how Taiwanese building investment and development companies rate the analytical framework and weights of artificial intelligence buildings. Document Analysis, the Delphi method, and the Fuzzy Analytic Hierarchy Process (FAHP) are applied to conduct a FAHP questionnaire survey among 20 building investment and development companies in Tainan. Based on the calculation of composite weights, the findings are: (1) The most crucial evaluation indicator for security and hazard prevention is the "access control system". (2) The most crucial evaluation indicator for energy-saving management is "energy-saving technology". (3) The most crucial evaluation indicator for health and comfort is the "interior comfort system". (4) The most crucial evaluation indicator for intelligent innovation is the "intelligent innovation concept".

Keywords: Intelligent Building, Fuzzy Analytic Hierarchy Process, Expert Decision Evaluation

1. Introduction

In the 21st century, humankind is facing the problems of global warming, climate change, urban overdevelopment, and the greenhouse effect, which have led to high global temperatures, depletion of forests, destruction of the ozone layer, and the frequent occurrence of extreme weather phenomena. As humankind is unable to resist the relentless forces of nature, in order to mitigate the damage to the earth's environment and pursue the goal of global sustainable development, many developed countries have embarked on the construction of intelligent and sustainable cities, and further developed the concept of Intelligent Buildings (IB). Therefore, the main reason for the growing trend of intelligent buildings in Taiwan and abroad is the transition of the human living environment.

This thesis proposes two connotations for intelligent buildings. One is to create a human living space based on a combination of green architecture and green building materials, which must provide users with safer, healthier, more convenient, comfortable, and energy-efficient living environments. The second is to introduce the concept of Intelligent Innovation into the planning and design of buildings to create a new value of intelligent buildings. The introduction of intelligent systems and devices helps maintain and extend the life cycle of buildings, and provides the best solution to the

energy consumption of buildings.
Objectives of this study are to:

(i) Construct an initial hierarchical structure of intelligent building evaluation models, as based on the theories described in the Intelligent Building Evaluation Manual.

(ii) Adopt an expert decision-making approach to evaluate the analytical framework of intelligent buildings and construct a definitive hierarchical structure of intelligent building evaluation models.

(iii) Apply AHP to calculate the respective weights of the 4 primary dimensions and 12 evaluation indicators of the intelligent building evaluation model.

2. Literature Review

The term "Intelligent Building" originated in the late 1970s and early 1980s with the emergence of an intelligent building that was converted from an old financial building in 1984 in Hartford City, Connecticut, USA. The following year, the Aoyama Building in Tokyo, Japan further improved the functionality of building intelligence, and maintained that a highly functional building is an intelligent building. These intelligent concepts of creating human living spaces through automated technology have not only created a new vision of architecture, they have also given rise to new architectural terms, such as intelligent buildings, intelligent houses, intelligent communities, and intelligent cities.

Based on the requirements of the main occupants and building facilities, intelligent buildings can be classified as automated buildings, intelligent homes, green buildings, efficient buildings, and energy efficient buildings that exchange with the grid[1-3]. Automated buildings focus on building electrics and automated mechanical facilities, while intelligent homes emphasize the design of a user-friendly environment for the occupants. Green or sustainable building study is centered on the creation of a productive interior environment and an environmentally friendly journey through the life cycle of a building, from its design, construction, and operation, to maintenance, renovation, and demolition. The study of energy-efficient buildings also focuses on the full life cycle of a building, in order to minimize energy consumption over that period.

Dounis et al. reviewed the control systems built up to 2008, and proposed a framework to analyze intelligence-led energy and comfort control systems[4]. Shaikh et al. reviewed the building control systems optimized up to 2013, which were divided into intelligent controllers and intelligent methods of managing energy and comfort calculations[5]. Nguyen et al. reviewed, analyzed, and classified the different building optimization problems, as well as the algorithms, tools, and operations used to optimize building energy management systems until 2013[6]. Evins et al. encapsulated the construction, operation, and energy production of intelligent and sustainable buildings for design and control systems until 2012[7].

For building developers, the key reason for introducing intelligent systems and devices is cost, hence, it is essential to understand what design requirements can satisfy an intelligent home. The 2016 edition of the Intelligent Building Standards Evaluation Manual includes eight evaluation criteria, incorporates convenience features into health and comfort, and adds intelligent innovation.

3. Evaluation Model Building for Intelligent Buildings

3.1. *Hierarchical Structure*

This thesis combines the results of two rounds of Delphi questionnaires to construct a definitive hierarchical structure for the intelligent building evaluation model, which consists of 4 primary dimensions and 12 evaluation indicators, as shown in Figure 1.

3.2. *Subject of this study*

In order to investigate how Taiwanese building investment and development companies rate the analytical framework and weights of intelligent buildings, this thesis took the building developers in Tainan as the study subject, and 8 experts were selected for the Delphi method and 20 for the FAHP questionnaire. In addition, prior to the FAHP questionnaire survey, the researcher first explained the objective of this thesis to the FAHP questionnaire subjects, and conducted the FAHP survey among building developers who were willing to participate.

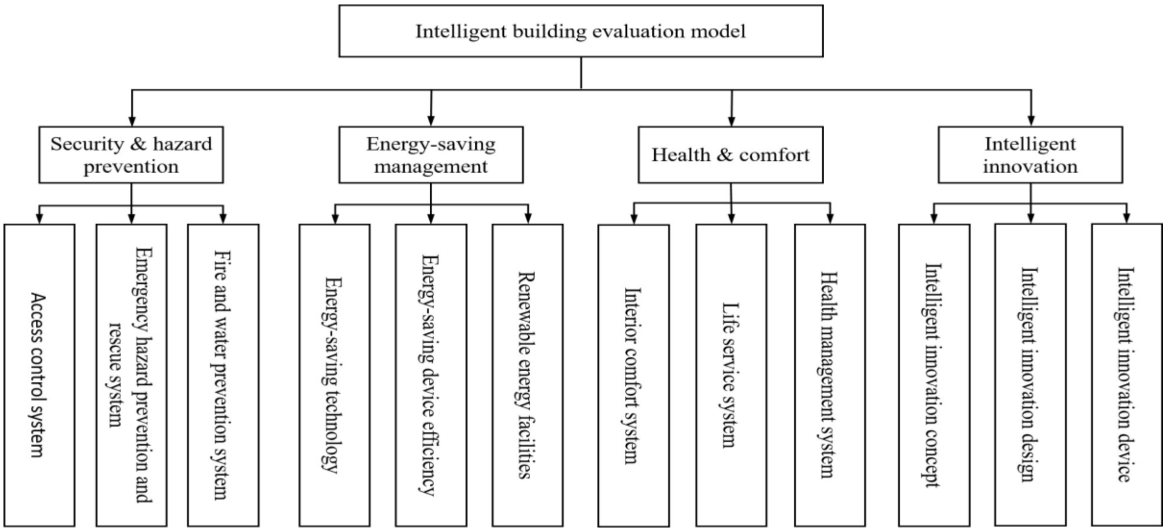


Fig. 1. Intelligent building evaluation model

4. Empirical Study of the Intelligent Building Evaluation Model

In this thesis, 9 evaluation indicators are multiplied by their respective dimensions to obtain the composite weights. Table 1 presents the composite weights of the intelligent building evaluation model.

As indicated in Table 1, the most crucial evaluation indicator, as perceived by the 12 building developers, is the "access control system" (with the composite weight of 0.235), and "energy-saving technology" (with the composite weight of 0.175) is ranked 2nd. The 3rd indicator is the "emergency hazard prevention and rescue system" (with the composite weight of 0.102), 4th place is the "interior comfort system" (with the composite weight of 0.085), and 5th is the "intelligent innovation concept" (with the composite weight of 0.073). Ranked from 6th to 12th place, respectively: "energy-saving device efficiency" (with the composite weight of 0.064), "fire and water prevention system" (with the composite weight of 0.061), "renewable energy facilities" (with the composite weight of 0.060), "life service system" (with the composite weight of 0.053), "intelligent innovation design" (with the composite weight of 0.034), "intelligent innovation device" (with the composite weight of 0.030), and "health management system" (with the composite weight of 0.028).

Table. 1. The composite weights of the intelligent building evaluation model

Dimension	Weight	Evaluation Indicator	Weight	Composite Weight
Security & hazard prevention	0.375	Access control system	0.593	0.235
		Emergency hazard prevention and rescue system	0.257	0.102
		Fire and water prevention system	0.150	0.061
Energy-saving management	0.303	Energy-saving technology	0.588	0.175
		Energy-saving device efficiency	0.232	0.064
		Renewable energy facilities	0.180	0.060
Health & comfort	0.164	Interior comfort system	0.495	0.085
		Life service system	0.237	0.053
		Health management system	0.178	0.028
Intelligent innovation	0.158	Intelligent innovation concept	0.622	0.073
		Intelligent innovation design	0.246	0.034
		Intelligent innovation device	0.132	0.030

5. Conclusions

The four key findings of this study are listed below:

5.1. *The most crucial evaluation indicator for security and hazard prevention is the "access control system"*

Building developers considered "Security and hazard prevention" as the most crucial dimension, with "access control system" being the most significant. This finding is aligned with the results of a survey conducted by the Taiwan Architecture & Building Center and the Taiwan Intelligent Building Association in 2013 regarding the "Application of Intelligent Building Adoption in

Taiwan". For the public, whether they choose a traditional or intelligent building, the primary evaluation factor is whether life and property are protected. It is also worth noting that the added value of intelligent homes is ex-ante proactive prevention of hazards by using various modern technologies to make homes intelligent, which allows for both burglary and hazard prevention, thereby reducing the chance of hazards occurring in homes, as well as minimizing the loss of life and property.

5.2. *The most crucial evaluation indicator for energy-saving management is the "energy-saving technology"*

Building developers considered "energy-saving management" as the second most crucial dimension, with "energy-saving technology" being the most significant, which implies that the majority of building developers agree that the application of energy-saving management systems in intelligent homes is both an important design trend and a major marketing pitch. This is probably due to the fact that energy-saving is an important indicator for evaluating whether an intelligent home can achieve significant energy savings after the introduction of energy-saving devices, such as intelligent air-conditioning and lighting with intelligent energy control technology. In other words, in addition to the need for ex-ante proactive prevention of burglary and hazard, the main consideration for intelligent homes includes the benefits of using energy-saving devices, such as economic efficiency and energy-saving effectiveness.

5.3. *The most crucial evaluation indicator for health and comfort is the "interior comfort system"*

Building developers considered "health and comfort" as the third most crucial dimension, with the "interior comfort system" being the most significant, which is probably because the main objective of intelligent buildings is to create a safe, comfortable, and healthy home environment. This objective relies heavily on the integration and linkage between the respective automated systems and energy-saving devices; for example, through intelligent technology, people can remotely activate or set a fixed time to switch on the power, which allows intelligent air conditioning to create a comfortable living environment, and people can enjoy cool air when they return home.

5.4. *The most crucial evaluation indicator for intelligent innovation is the "intelligent innovation concept"*

Building developers considered "intelligent innovation" as the least crucial dimension, with the "intelligent innovation concept" being the most significant. The main reason for this is the recent surge of intelligent cities across the world, and the increasing popularity of intelligent systems and devices, which indicates that many people have a certain level of understanding of the concept of intelligent buildings, including security, energy saving, health, and comfort.

References

1. J. Wong, H. Li, S. Wang, Intelligent building research: a review, Autom. Constr. 14 (1) (2005) 143–159.
2. M. Wigginton, J. Harris, Intelligent Skin, Architectural Press, United Kingdom, 2002.
3. U.S. Department of Energy Office of Energy Efficiency and Renewable Energy, Grid-interactive Efficient Buildings Technical Report Series- Whole-Building Controls, Sensors, Modeling, and Analytics, US Department of Energy, December 2019.
4. A.I. Dounis, C. Caraiscos, Advanced control systems engineering for energy and comfort management in a building environment - a review, Renew. Sustain. Energy Rev. 13 (2009) 1246–1261.
5. P. Shaikh, N. Nor, P. Nallagownden, I. Elamvazuthi, T. Ibrahim, A review on optimized control systems for building energy and comfort management of smart sustainable buildings, Renew. Sustain. Energy Rev. 34 (2014) 409–429.
6. A.T. Nguyen, S. Reiter, P. Rigo, A review on simulation-based optimization methods applied to building performance analysis, Appl. Energy 113 (2014) 1043–1058.
7. R. Evins, A review of computational optimisation methods applied to sustainable building design, Renew. Sustain. Energy Rev. 22 (2013) 230–245.

Authors Introduction

Dr. Li-Min Chuang

He received the Ph.D. degree from the National Cheng Kung University. He is currently an associate professor and department head at the Department of International Business at Chang Jung Christian University, Taiwan. His research interests focus on strategic management, innovation management, and service innovation.

Mr. Yu-Po Lee

He works as the manager in the Wei Shin Aluminum Company for more than 30 years. He received his Master's degree from the EMBA, Chang Jung Christian University in 2012. He is currently studying for a doctoral program at Chang Jung Christian University in Taiwan.

Mr. Chien-Chih Kuo

He received his Master's degree from Chang Jung Christian University, Taiwan in 2015. He is a chairman of Chieh Chang Construction Co., Ltd., and Tainan Judicious Creative Architecture Association.

Blockchain-based Verification Mechanism for Industrial Control System

Yao-Chu Tsai [1]
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 [2]
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 [*3]
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: yctsai@cans.ee.ncku.edu.tw [1], ihliu@cans.ee.ncku.edu.tw [2], jsli@mail.ncku.edu.tw [*3]*
www.ncku.edu.tw

Abstract

Industrial control system (ICS) and critical infrastructure have become increasingly dependent on communication network and cyber-physical systems. Since infrastructure is vulnerable to adversarial attacks, the research on cyber security is vital in Industry 4.0. In order to secure the integrity of data in ICS, this paper proposes a blockchain-based network architecture implemented on physical industrial equipment. By arrangement of blockchain transaction process in the specialized client-server network model, industrial control signal transmission can be verified and recorded based on authority.

Keywords: Cyber Security, Blockchain, Verification, ICS Security

1. Introduction

With the development of automation and network communication in modern manufacturing technology, the concept of Industry 4.0 has been proposed in recent years, also known as the Fourth Industrial Revolution. The Industrial Internet of Things (IIoT) is one of the widely researched issues. The purpose is to innovate in the industrial economy, construct a smart-conscious industrial environment, and develop smart factories with adaptability, resource efficiency and human-machine collaborative engineering.

Industrial Control System (ICS) contains various kinds of IIoT and critical infrastructure. Security and reliability of ICS has become a concern of industry and government. In 2019, The Department of Homeland Security, United States published "A Guide to a Critical Infrastructure Security and Resilience" to regard critical infrastructure security as national security [1].

As a subgroup of ICS, Supervisory Control and Data Acquisition (SCADA) system is composed of hardware and software related to the data storage, processing, and communication. For instance, Programmable logic controller (PLC) is an operating component in SCADA, which transmits signal among manufacturing equipment and machinery, serving as a fundamental part in cyber-physical system [2].

PLC is basically a terminal device that can control low-level I/O devices, such as sensors or motors, and usually

connects to the PC via Ethernet to embed program. As a kind of embedded system on the Industrial Internet, it is exactly at the risk of encountering network attacks. In 2021, a real incident happened in a water treatment plant in Florida. A hacker tried to remotely hack into the water purification control system with an attempt to fill a potentially harmful chemical [3].
In order to ensure the integrity of data transmission and information security, this paper designs a blockchain-based network architecture for the data transmission in industrial control system.

2. Background

In this section, we mainly discuss the adaptability of blockchain technique in ICS. Building in the field of industrial networks, the performance of computing equipment in IIoT cannot be compared with that in IT environment. The restriction on IIoT is strict due to transmission time and memory space. Most business owners tend to avoid changing or upgrading hardware because of cost and the possibility of production interruptions [4].

2.1. *Modbus TCP*

Modbus is one of the standard communication protocols among PLCs, and it is also commonly used in the current industrial field. Furthermore, Modbus TCP can transmit data through Ethernet TCP/IP. Because Modbus is a plaintext protocol belonging to the application layer, it is easy to be interpreted or tampered by hackers, able to conduct man-in-the-middle attack (MitM) and other malicious behaviors.

2.2. *Private Chain*

It is vital to design a secure and efficient network architecture in ICS. Among various types of blockchain, although private chain is a more centralized system, it is quite suitable as a transmission medium for the transfer of confidential value within a single company or organization. Different from public chain and consortium blockchain, only the holder in private chain can participate in the recording of ledgers and data. This factor causes streamlined structure and higher transaction speed of private chain. Therefore, private chain is a better solution for specific institution.

2.3. *Proof of Authority*

Proof of Authority (PoA) is a consensus algorithm in blockchain. To obtain the authority, the nodes in the system need to pass the identity verification first. If verified, certain blockchain nodes are set to have the authority to participate in transactions and finally decide whether to add new blocks to the blockchain. In private chain, PoA can makes a positive effect. Compared with proof of work (PoW), PoA does not rely too much on the computing power of network nodes and is more suitable for applications in the environment of industrial control network.

3. System Architecture

We make explanation of proposed method and overall framework in this section. There are some adjustments in traditional network configuration and verification mechanism to make blockchain operation more compatible in ICS.

3.1. *Network Model*

For the purpose of implementing blockchain in ICS environment, we design a network architecture presented in Fig. 1. It is a client-server network model imported in ICS [5].

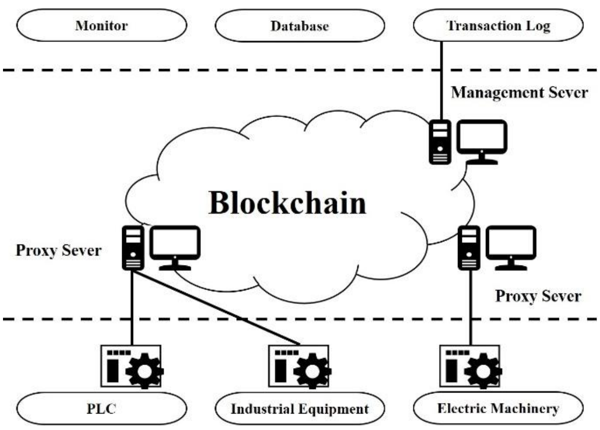


Fig. 1 Blockchain-based ICS network model

The network model explains the ICS devices, such as PLCs, industrial equipment, and electric machinery, are clients of the proxy server. The proxy server accesses to blockchain, so that it can communicate in a protected environment., allocating resources to the ICS devices in the LAN. PC is usually used to be a proxy server, which is more suitable for fast executing blockchain programs in terms of performance.

We propose a method to wrap the Modbus protocol in a block when sending signal using blockchain. Block is a JSON file containing some transaction information, and there is an “extraData” field in it. We decide to put the content of a Modbus packet in this field for transmission in blockchain.

Combined with the immutability property of blockchain, we set a management server to efficiently manage all account information on the network and record the transaction process. If the transaction logs are sealed in the blocks, the log content cannot be easily tampered.

3.2. *Blockchain-based Transaction Process*

Having finished registering, we can launch transactions among accounts. The blockchain-based transaction process includes six elements, PLC A, proxy server A, PLC B, proxy server B, and management server. And the whole process can be divided into two phases.

Phase 1: Request from domain A to domain B

- Step 1: PLC A sends Modbus TCP packets to proxy server A to request authentication to access the resources of PLC B.
- Step 2: Proxy server A and proxy server B synchronizes the account information of each other.
- Step 3: Proxy server A initiates a transaction through the blockchain and transfers the Modbus TCP message to proxy server B in the extraData field.
- Step 4: Proxy server B accepts the transaction and signs the certificate by its private key.
- Step 5: Proxy server B takes out the Modbus TCP message in extraData field and sends it to PLC B.

Phase 2: Response from B to A

- Step 6: PLC B sends proxy server B a Modbus TCP packets to make a reply to PLC A.
- Step 7: Proxy server B initiates a transaction through the blockchain and transfers the Modbus TCP message to proxy server A in the extraData field.
- Step 8: Proxy server A accepts the transaction and signs the certificate by its private key.
- Step 9: Proxy server A takes out the Modbus TCP message in extraData field and sends it to PLC A.
- Step 10: Proxy server A and proxy server B synchronize their transaction information to the management server for transaction record backup.

4. Experiment

It is practical to choose an appropriate platform according to the required functions in blockchain. Ethereum is used for blockchain development in this paper. The experimental environment setup is shown in Fig. 2.

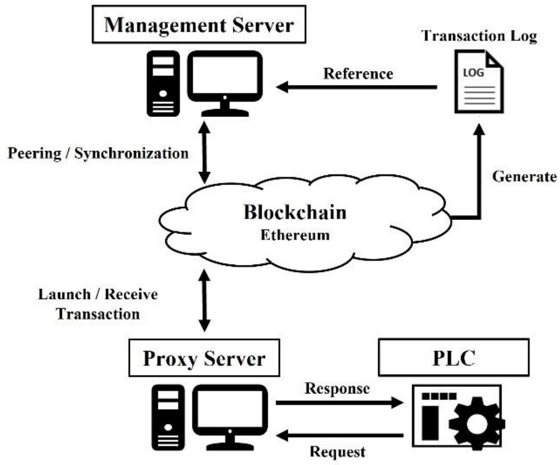


Fig. 2 Experiment environment setup

During online transactions among proxy servers, the Modbus data benefits from the characteristic of security in blockchain. When PLC is ready to receive the verified message, the data needs taking out of the blockchain. As a result, we make an Application Programming Interface (API) to fetch the Modbus information in the block. Likewise, the interface program can also receive signals from PLC. In our experiment, we use the PLC WinPAC WP-8428-CE7 produced by ICP DAS, Taiwan as the industrial control equipment.

Adjusting the block generation period in genesis block, transmission time measurement can be implemented, and the result is shown in Table 1. We found that the period dominates entire transmission time. It can be inferred that when the period is large, the time interval between block generation is also relatively large, which leads to a higher standard deviation of the transaction time.

Table 1. Transmission time with different periods

Period (s)	Number of Transactions	Average Time (s)	Standard Deviation (s)
1	100	1.5305	0.058719
5	100	3.5838	0.096524

One of application in the management server is transaction log. At the end of each transaction, we arrange for the proxy servers to peer with the management server, synchronizing their transaction logs, so that network administrator can examine all transaction records in convenience by using management server.

5. Conclusion

Since the development of information technology is quite rapid, the number and types of IoT devices are increasing. Being a widely used domain of IoT, ICS is constructed by cyber-physical systems. ICS cybersecurity has been a research focus nowadays.

For the purpose of ensuring the transmission integrity in ICS, this paper applies an emerging technology, blockchain, to physical industrial network. Running blockchain in the designed client-server network model, it would improve the adaptability of resource-intensive defense mechanism in ICS.

This paper also plans the transaction process in the blockchain framework, implementing on physical industrial equipment as well. The transaction logs are highly credible and eligible to provide digital evidence since it is not easily tampered. According to the limit and requirement of the environment, we adjust the parameter of blockchain to comply with the coordination of ICS device and IT network management system.

References

1. F. Enayaty-Ahangar, L. A. Albert, and E. DuBois, "A survey of optimization models and methods for cyberinfrastructure security", IISE Transactions, vol. 53, no. 2, pp. 182-198, 2020
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. CNN, "Florida water treatment facility hack used a dormant remote access software, sheriff says", https://edition.cnn.com/2021/02/10/us/florida-water-poison-cyber/index.html
4. G. Bonney, H. Höfken, B. Paffen, and M. Schuba, "ICS/SCADA Security Analysis of a Beckhoff CX5020 PLC", 2015 International Conference on Information Systems Security and Privacy (ICISSP), 2015.
5. C. Wu, J.Lu, W.Li, H.Meng, and Y.Ren, "Master-slave Blockchain Based Cross-domain Trust Access Mechanism for UPIOT", 2020 5th International Conference on Computer and Communication Systems (ICCCS), 2020.

Authors Introduction

Mr. Yao-Chu Tsai

He was born in Pingtung, Taiwan in 1997. He received his B.S. degree from the Department of Systems and Naval Mechatronic Engineering, National Cheng Kung University, Taiwan in 2020. He is acquiring the master's degree in Department of Electrical Engineering/Institute of Computer and Communication Engineering, National Cheng Kung University in Taiwan.

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.

Research on Path Planning Algorithms of Multiple Mobile Robots in Intelligent Warehousing

Jiwu Wang
School of Mechanical and Electronic Engineering, Beijing Jiaotong University
Beijing, Haidian District, China

Shilong Zheng
School of Mechanical and Electronic Engineering, Beijing Jiaotong University
Beijing, Haidian District, China
E-mail: jwwang@bjtu.edu.cn, 20126112@bjtu.edu.cn
www.bjtu.edu.cn

Abstract

In recent years, with the rapid development of e-commerce industry, people's demand for intelligent automatic storage system has become more and more urgent. Therefore, the research on intelligent storage system based on multi mobile robots has turned into a hot spot and development direction of intelligent manufacturing industry. Among them, the collaborative path planning between multi mobile robots has become a key problem to be solved. In order to solve this problem, this paper will propose an improved A^*algorithm based on reservation table. Firstly, the reservation table is generated through centralized control; Then, the collision problem between robots is prevented by reservation table and directed graph; Finally, through the distributed control method, the multi robot cooperative path planning is realized by using the improved A^* algorithm. The effectiveness of the algorithm is verified by simulation.

Keywords: multi-robot system; path planning; intelligent warehouse; A^* algorithm

1. Introduction

With the rapid development of e-commerce industry and industrial automation, the demand for efficient intelligent automatic warehouse system is becoming more and more urgent. Intelligent warehouse system has become one of the hot spots in the field of intelligent system research. Therefore, the automatic warehouse system based on multi-mobile intelligent robot has become the research hotspot of intelligent warehouse system.

Collaborative path planning among multi-robots is the key to realize intelligent warehouse based on multi-mobile robots. This paper will solve the above problems through the reservation table and the improved A^* algorithm under the directed graph, so as to build a safe and stable intelligent warehouse system.

At present, the main control methods of multi-robot system include centralized control method and distributed control method.[1] The centralized control method is coordinated by the central processing system to find the global optimal path of the robots. However, with the increase of the number of robots in the warehouse, the computational complexity of the centralized control method will be greatly increased. Distributed control method is a kind of path planning method which can adapt to the rapidly increasing number of robots and the rapidly changing warehouse environment. However, due to the limited information obtained, the distributed control method can only find the local optimal path for the robot, and may cause conflicts among machines. In this paper, the advantages of centralized control method and distributed control method are combined. On the one hand, the centralized control method is used to obtain the global information of the warehouse (such as the status and location of the robot, etc.), and the global information is used to construct the reservation table to provide real-time traffic information for the robot in the intelligent warehouse. On the other hand, the robot system adopts distributed control mode, and each robot carries out path planning according to the information provided by the central control system, which not only reduces the burden of the central control system, but also provides global information for multi-robot path planning.

2. Intelligent Warehouse Model and Task Analysist

Because the warehouse environment is a structured environment, this paper uses grid map to represent the warehouse environment. As shown in Figure 1. The grid map mainly includes three parts: 1. The picking table $S = \{s_1, \dots, s_m\}$ (m is the number of picking tables) on the left side of the grid map. The staff pack at the picking table, which is located at the edge of the warehouse to facilitate the distribution of goods; 2. The path continuously composed of adjacent grids in the grid map; 3. The shelf $L = \{l_1, \dots l_e\}$ (e is the number of shelves). A group of robots $R = \{r_1, \dots, r_n\}$ (n is the number of robots) walk and carry goods in the warehouse at a uniform speed v. [2]

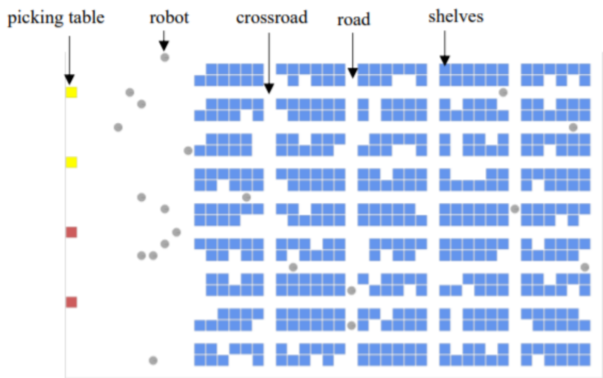


Fig. 1. Intelligent warehouse model.

The process of muti-robot transporting goods in intelligent warehouse is shown in Figure 2. The central control system arranges the order information and sends it to the idle robot. The order information mainly includes the shelf position information and the position information of the picking table. The robot first goes to the shelf position to carry the shelf for picking (the robot that does not carry the goods can move under the shelf during the movement), transports the shelf to the corresponding picking table, does not carry out the picking work, and then sends the shelf to the original position after picking, then the order task ends. The state of the robot changes from working to idle and continues to execute the next order task.

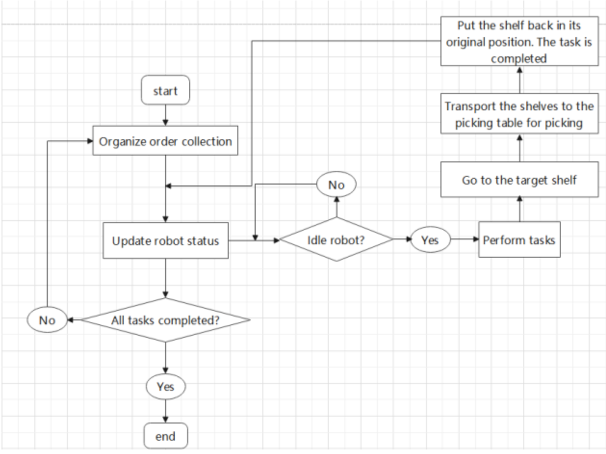


Fig. 2. Task flow.

3. Improved A^* Algorithm

The A^* algorithm is called the best first heuristic algorithm, which is simpler and faster than the unknown exploration algorithm.[3] The algorithm steps are as follows: starting from the starting grid, each time the robot selects the grid to be expanded in the next step, it will estimate the grid path cost adjacent to its own grid. Path cost refers to the shortest estimated value of the robot from the starting grid to the target grid after passing through the grid to be expanded. After obtaining the estimated cost of the surrounding grid, put it into a candidate table with expansion nodes, and then select a grid with the lowest estimated cost in the table for expansion. The robot repeats the above steps until it extends to the target grid. The estimated cost per grid is:

$$f^*(n) = g(n) + h^*(n) \tag{1}$$

Where: $g(n)$is the real cost of the robot moving from the starting grid to the current grid n:

$$g(n) = \frac{d}{v} \tag{2}$$

v is the speed at which the robot travels at a constant speed, and d is the real distance from the starting grid to the current grid n.

$h^*(n)$ is a heuristic equation to calculate the estimated cost of the robot moving from the current grid n to the target grid:

$$h^*(n) = \frac{d_n}{v} \quad (3)$$

d_n is the shortest estimated distance from the current grid n to the target grid, i.e. Manhattan distance:

$$d_n = abs(n.x - goal.x) + abs(n.y - goal.y) \quad (4)$$

$n.x$ and $n.y$ are the horizontal and vertical columns of the current nodes in the raster map, $goal.x$ and $goal.y$ are the horizontal and vertical columns of the target nodes in the raster map.

Distance, and time can be used to estimate $f^*(n)$. The traditional A^* algorithm mainly uses Manhattan distance, Euler distance, etc., as the cost of the estimation function in the formula, while this paper uses time as the cost of estimation.

When the A^* algorithm is used in the warehouse there are still some problems, as shown in figure 3, starting from the same starting grids, the robot can choose a few different paths to reach the same target grid, called path a and b.

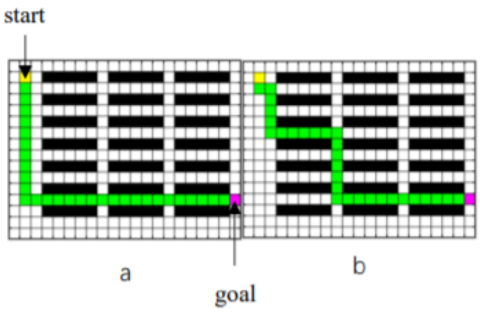


Fig. 3. The paths.

Respectively, in a structured intelligence warehouse, the same length of the path, the traditional A^* algorithm will randomly select a path to extend. However, in a real warehouse environment, the robot would spend more time turning around corners, so it would take different time for the robot to walk along different paths of the same path length. As shown in the figure, there is one turn when the robot walks along path 1, and there are five turns when the robot walks along the path and walks. Therefore, the turning cost time t_{turn} is proposed in this paper. When the robot decides to turn at the intersection, t_{turn} is added into the estimation function, and "Eq. (1)" is modified as

$$f^*(n) = g(n) + h^*(n) + t_{turn} \quad (5)$$

Using the improved method, a path with shorter distance and less time consumption can be selected, thus improving the efficiency of path planning for a single robot.

4. Reservation Table and Directed Graph

Collisions and even deadlocks may occur in robots running continuously in A dynamic environment, which cannot be solved by the improved A* path planning algorithm alone. Figure 4 shows four typical collision scenarios.

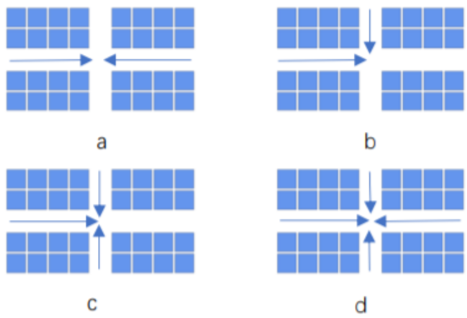


Fig. 4. Several collision types.

If the collision is in case b, the robot can be resolved by waiting for other robots with higher weights to pass first. When the other three collision situations occur, it will take a lot of time to avoid collision and avoidance among robots, and with the increase of the number of machines, collision and avoidance will cause deadlocks. The above three situations can be avoided by building a directed graph. It is stipulated that each road between shelves in the warehouse is a one-way street, and the direction of the two adjacent roads is opposite.[4] When the robot walks in the warehouse, it can only move forward in the direction of the road, thus avoiding collision and deadlock to a large extent. The robots move in order according to the traffic rules in the intelligent warehouse. As shown in Figure 5, the horizontal arrow indicates that the robot can only move from left to right in accordance with traffic rules; the longitudinal arrow indicates two one-way streets in opposite directions; the position of the shelf is a two-way arrow, and the two-way arrow indicates that the robot can move in both directions. That is, the direction of the directed graph should be followed when the improved A* algorithm is used for global path planning.

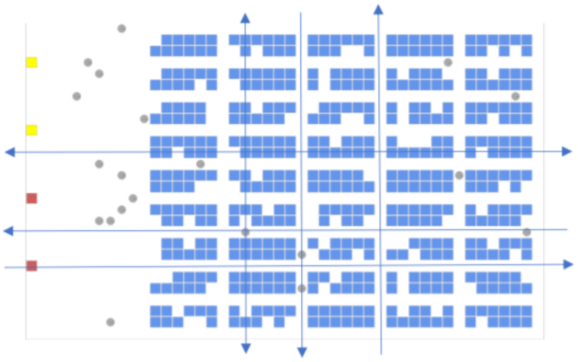


Fig. 5. Directed graph.

Even with traffic rules, there will still be a type b collision in the robot world as shown in Figure 4. Therefore, each grid in a raster map can only be occupied by one robot at a time.[5] If the grid that the robot intends to enter is occupied, it must wait until the grid becomes idle. In this paper, the reservation table is used to record the state of each grid in the grid map. Figure 6 shows the reservation table from the current time t.

col / row	col_1	col_2	…	col_{n-1}	col_n
row_1	$RobotID_1$		…	$RobotID_3$	
row_2			…		$RobotID_{k-1}$
⋮	⋮	⋮		⋮	⋮
row_{m-1}		$RobotID_2$	…		
row_m			…		$RobotID_k$

Fig. 6. Reservation table.

After each robot path planning, the central control system will update the reservation table once. The reservation table is a two-dimensional table the size of a raster map, and the Spaces in the table record the usage of each grid in the corresponding raster map. If the grid is occupied, the ID of the robot occupying the grid and the time of the robot occupying the grid will be recorded in the corresponding position in the table. Other robots can decide whether to move forward or wait for another robot to pass by querying the reservation table at the current time. If multiple robots occupy a grid at the same time, the order of occupancy will be determined according to the weight. Robots that start early have higher weights.

5. The Simulation Results

In this chapter, the warehouse model proposed above is used to simulate and verify the effectiveness of the path planning algorithm. Set the number of robots to 20 and the number of orders to 100.

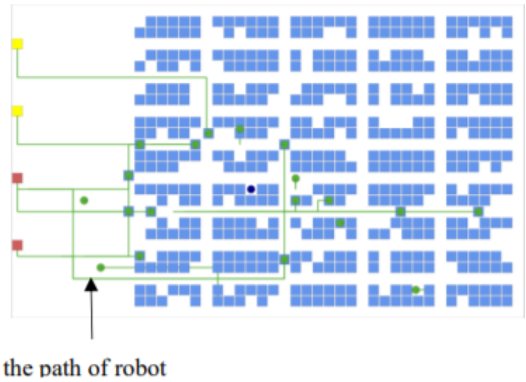


Fig. 7. Improve A^* algorithm simulation process.

As shown in Figure 7 of the simulation process, the total time spent by the robots to complete the task was 13 minutes and 47 seconds, and the total distance traveled was 4221 meters. No collision or deadlock occurred between the robots. The proposed algorithm is effective.

References

1. Tai Ruochen. Multi-robot Path Planning Algorithm in Warehousing and Logistics Environment [D]. Shanghai Jiaotong University,2019.
2. Zhang Danlu. Research on key technologies of multi-robot dynamic path planning based on intelligent warehouse [D]. University of Chinese Academy of Sciences (Chongqing Institute of Green and Intelligent Technology),2017.
3. Jiang Zhongyu. Research on path Planning of multi-handling Robot under Moving Obstacles [D]. Beijing Jiaotong University,2020.
4. Xv tong, Tang Zhenmin. Dynamic Collision Avoidance Planning in multi-robot system [J]. Computer Engineering,2003,29(17):79-81.
5. Zhang Ximei. Research on Picking Operation Optimization and Algorithm Based on Kiva System [D]. Beijing University of Posts and Telecommunications,2015.

Authors Introduction

Dr. Jiwu Wang

He is an associate professor, Beijing Jiaotong University. His research interests are Intelligent Robot, Machine Vision, and Image Processing.

Mr. Shilong Zheng

She is a postgraduate in Beijing Jiaotong University. His research interest is Intelligent Robot.

Research on the Effectiveness of Monocular Visual SLAM Depth Estimation Base on Improved ORB Algorithm

Jiwu Wang
School of Mechanical and Electronic Engineering, Beijing Jiaotong University
Beijing, Haidian District, China

Weipeng wan
School of Mechanical and Electronic Engineering, Beijing Jiaotong University
Beijing, Haidian District, China
E-mail: jwwang@bjtu.edu.cn, 13014641154@163.com
www.bjtu.edu.cn

Abstract

The application of monocular vision to measure the depth information of image feature points is one of the important tasks of monocular vision slam. Triangulation is a method often used to measure the depth information of feature points, but in the actual application process, the uncertainty of feature point matching will cause greater depth uncertainty. This paper proposes an improved ORB feature point extraction strategy, combined with the quad-tree model to achieve the homogenization of feature points. The feature points matching method uses the brute force matching method, RANSAC filters the matching point pairs, and obtains better matching results for depth estimation. Experiments show that the improved feature point extraction and matching method can effectively obtain camera pose estimation value and depth estimation value. And in this way, the accuracy of the estimated value is improved .

Keywords: orb；monocular vision；depth estimation; vision slam.

1. Introduction

Measuring the depth information of image feature points is a very important part of orb-slam or other monocular vision slam. In the visual odometer based on feature point method, the motion of the camera is estimated with the extracted image feature points.[1] The basic process includes image feature point extraction, image feature point matching between frames, camera motion solution by epipolar constraint . The camera rotation matrix R and translation matrix T obtained at this time are used to estimate the image feature point depth. The feature point extraction and matching in the above process affects the estimation of the rotation matrix R and the translation matrix T, and also affects the estimation of the image depth information.

In view of the above problems, we can find that a good feature point extraction and matching method will bring better pose estimation and better depth estimation. Among the many feature point extraction strategies, the ORB algorithm has the advantages of smaller calculation amount and simple feature point information description. It is widely used in slam based on the feature point method, but the accuracy and stability of traditional ORB detection algorithm in slam need to be further improved.

2. Principle

2.1. *The principle of ORB feature point extraction*

The ORB algorithm is improved on the basis of the FAST algorithm. ORB adds a description of the direction to FAST to ensure rotation invariance and uses the Steer BRIEF descriptor to describe the feature points.[2] The principle steps of the algorithm are as follows:

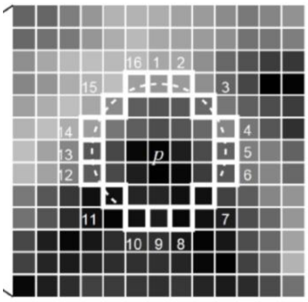


Fig. 1. FAST key points

Step1:Extract FAST key points. As shown in Figure 1. We arbitrarily select a pixel point p in the image, and the gray value of point p is Ip. The gray value of pixel p is compared with the gray value of 16 pixels on a circular window with a pixel radius of 3. Set a threshold t. If there are consecutive N (set to 9 or 12) pixels on the circle whose gray value is greater than Ip+t or less than Ip-t, then the pixel p can be considered as a FAST key point.
Step 2:Calculate the direction of feature points. Redefine the coordinate system with the key point p as the center O as shown in Figure 2.

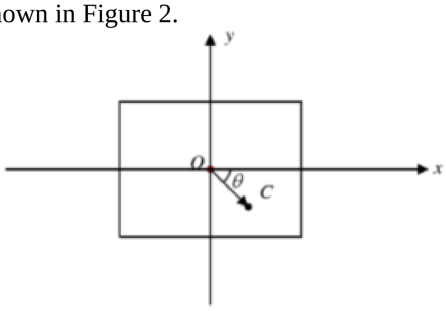


Fig. 2. Key point direction

The direction of the key point can be calculated according to the following equation. In the equation, x and y represent the coordinates of the pixel, I represents the gray value of the pixel, B represents pixel block, C represents the centroid position of the pixel block near the feature point p, and the vector **OC** represents the direction of the feature point p.

$$m_{pq} = \sum_{x,y\epsilon B} x^p y^q \, I(x,y), \;\; p,q = \{0,1\}. \quad (1)$$

$$C= \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}}\right). \quad (2)$$

$$\theta = arctan(\frac{m_{01}}{m_{10}}). \quad (3)$$

Step3:Information description of key point P. Use the "Steer BREIEF" descriptor for information description. Randomly select 256 pairs of pixel points in the pixel block near the key point P, each pair of points is represented as (pi, qi), pi and qi represent the pixel points, if the gray value of pi is greater than qi, the value is 1; otherwise, it is 0."Steer BREIEF" uses the rotation information to calculate the descriptor. 256 pairs of pixels form a matrix Q. After the image is rotated by an angle α, the matrix used to describe the feature point information becomes Q1.[3]

$$Q1 = R_\alpha Q. \quad (4)$$

2.2. *Camera pose estimation and feature point depth estimation*

Assuming that a feature point P can be detected in the space, the projection of point P on the image at camera position 1 is p1, and the projection on the image at camera position 2 is p2. From the epipolar constraint, the following formula can be obtained:

$$x_2^T t^{\wedge} R x_1 = 0. \quad (5)$$

$$x_1 = K^{-1} p_1 \, , x_2 = K^{-1} p_2 \, . \quad (6)$$

x1 and x2 are the projected coordinates of point P on the normalized plane of the camera, T is the translation matrix, and R is the rotation matrix.

$$s_2 x_2 = s_1 R x_1 + t. \quad (7)$$

$$s_2 \hat{x}_2 x_2 = 0 = s_1 \hat{x}_2 R x_1 + \hat{x}_2 t. \quad (8)$$

s1 and s2 represent the depth of point P at camera position 1 and camera position 2. Observing the following equations 5 and 6 and Figure 3, we can find that the position error of the points p1 and p2 on the image will lead to the estimation error of the R and T matrices, as well as the depth estimation error.

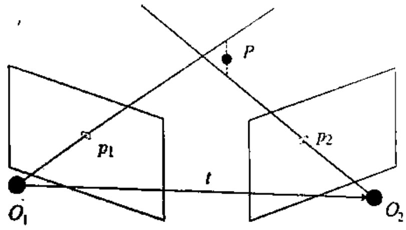


Fig. 3. Projection model

3. Improved ORB feature extraction and matching method

This article considers that the selection of the key point P is not only related to the peripheral pixels of the circular window, but also needs to pay attention to the pixel information inside the circle, and believes that the pixels closer to the key point need higher weights.

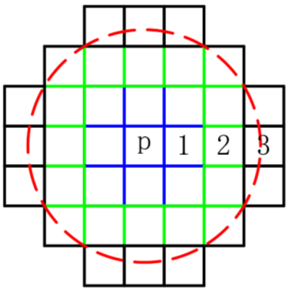


Fig. 4. Improve key point extraction

Set a judgment threshold t, divide the pixels in a circular window with a pixel radius of 3 into three layers, and subtract the gray value of p point from the gray value of each pixel on the inner to outer layer. The grayscale difference of the pixels of the first layer is denoted as vi, the difference of the second layer is denoted as vj, and the difference of the third layer is denoted as vk. Calculate the sum of the absolute value of the weighted difference of each layer. If there are N (take 18) differences whose absolute value is greater than the threshold t and satisfy that sum/37 is greater than t, then point p can be considered as the key point.

$$sum = w1 * |vi| + w2 * |vj| + w3 * |vk| \quad (9)$$

Use the gray-centroid method to calculate the feature point direction, and use "Steer BRIEF" to determine the feature point descriptor.

In the process of feature point extraction, the quad-tree model is combined to realize the homogenization of feature points and avoid the problem of accuracy reduction caused by the clustering of feature points. In the matching process, combined with the RANSAC algorithm to remove mismatches.

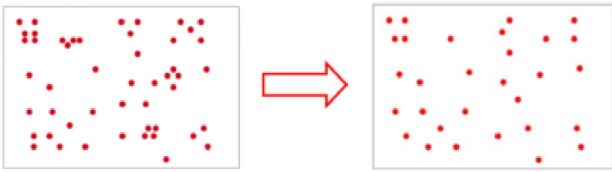


Fig. 5. Quad-tree homogenization

4. Experimental results and analysis

4.1. *Comparison of feature point extraction strategy*

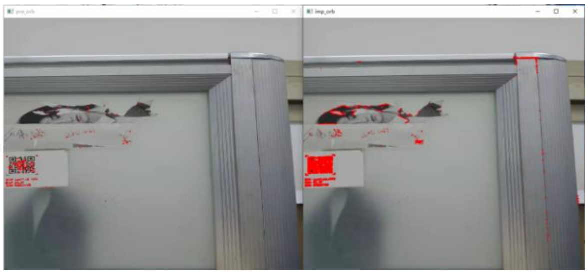
In this section, the following experiments are carried out to verify the effectiveness of the improved algorithm. The experimental results are shown in Table 1

Table 1. The planning and control components.

	ORB	Improved ORB
Feature point extraction number	503	5042
Maximum Hamming Distance	103	99
Number of matches with Hamming distance less than 40	51	671

In the feature point extraction process, the threshold t is set to 40. The original feature point extraction strategy can obtain 503 feature points. Among the matching results, the maximum Hamming distance is 103, and there are 51 matching results with Hamming distance less than 40. The improved feature point extraction strategy can get 5042 feature points. Among the matching results, the maximum Hamming distance is 99, and there are 671 matching results with hamming distance less than 40.

The two feature point extraction methods do not use non maximum suppression.



(a) Feature point extraction results

(b) Feature point matching results

(c) Matching results with Hamming distance less than 40

Fig. 6. Comparison of feature point extraction effects

4.2. *Quad-tree homogenization and RANSAC mismatch removal experiment*

Experimental results show that the improved orb feature point extraction strategy, combined with quad-tree and RANSAC algorithm can effectively manage feature points, quad-tree homogenization can effectively improve feature point matching results, and RANSAC can filters out most of the mismatched results.[4]

Fig. 7. Homogenization of feature points

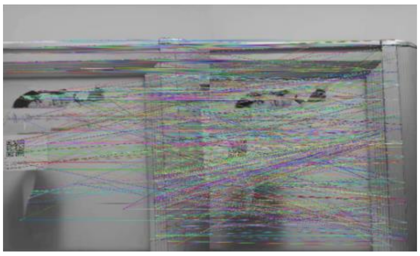


Fig. 8. Homogenization matching results

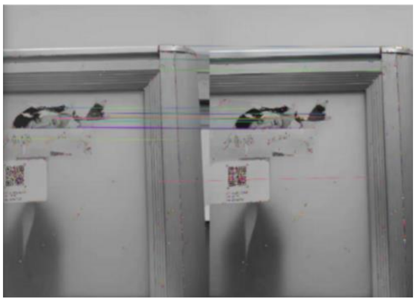
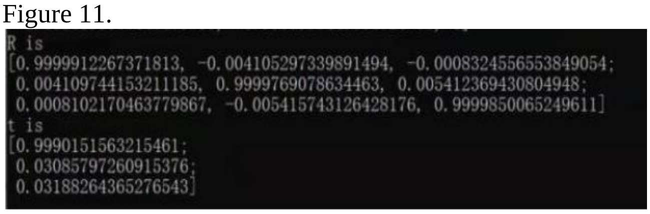


Fig. 9. RANSAC removes mismatches

4.3. *Camera pose and feature point depth estimation*

The results of rotation matrix R, translation matrix T and feature point depth obtained by using the improved feature point extraction strategy combined with quad-tree and RANSAC algorithm are shown in Figure 10 and Figure 11.



```
R is
[0.9999912267371813, -0.004105297339891494, -0.0008324556553849054;
 0.004109744153211185, 0.9999769078634463, 0.005412369430804948;
 0.0008102170463779867, -0.005415743126428176, 0.9999850065249611]
t is
[0.9990151563215461;
 0.03085797260915376;
 0.03188264365276543]
```

Fig. 10. depth information

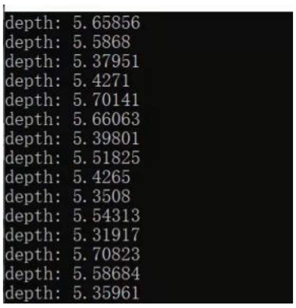


Fig. 11. Rotation matrix R and translation matrix T

The results without any optimization strategy are shown in the figure 12 and figure 13.

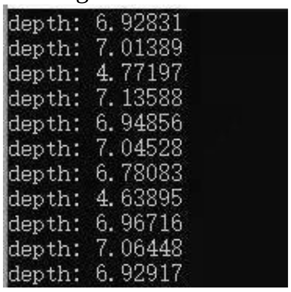


Fig. 12. Without optimized depth information

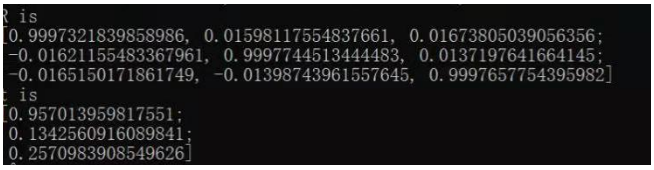


Fig. 13. Without optimized Rotation matrix R and translation matrix T

In the image acquisition process, only let the camera translate along the x direction, and the feature points are basically distributed in the same plane. The experimental results show that the optimized rotation matrix R is similar to the original result. But the result of the optimized translation matrix T is [0.999,0.031,0.032], and the result of the Original translation matrix is [0.957,0.134,0.257], which shows that the optimization effect is better. And the optimized feature point depth is basically distributed around 5 unit distances, which is more in line with the experimental environment.

5. Conclusion

This paper proposes an improved ORB algorithm for monocular vision slam to measure the depth information of image feature points and estimate the pose of the camera. Experimental results show that this method can meet the measurement accuracy requirements. Compared with the unimproved method, the method proposed in this paper improves the estimation accuracy of pose and depth. This method can provide good support for the posture tracking of the robot. However, due to the increase in the computational complexity of the algorithm, it takes longer to run. In order to meet the real-time requirements of slam, the speed of the algorithm needs further improved.

References

1. Guanci Yang, Zhanjie Chen, Yang Li, etc.Rapid Relocation Method for Mobile Robot Based on Improved ORB-SLAM2 Algorithm [J].Remote Sens, 2019,11,149.
2. Yang Hong-fan, LI Hang, etc.Image feature points extraction and matching method based on improved ORB algorithm [J].JOURNAL OF GRAPHICS, 2020, 41(4) : 548-555.
3. Leng Xue,Yang Jinhua. Improvement of ORB algorithm [J]. AER-Advances in Engineering Research, 2012, 28:903-906.
4. Jiangguo She, Rentong Xu, Ning Chen. Image stitching technology based on orb and improved RANSAC algorithm. Journal of Jiangsu University of Science and Technology (Natural Science Edition), 2015,29(2):164-169.

Authors Introduction

Dr. Jiwu Wang

He is an associate professor, Beijing Jiaotong University. His research interests are Intelligent Robot, Machine Vision, and Image Processing.

Mr. Weipeng Wan

He is a postgraduate in Beijing Jiaotong University. His research interests are vision slam and image processing.

Corner Detection Algorithm Based on Edge Contour in Automatic Loading Positioning

Jiwu Wang
School of Mechanical and Electronic Engineering, Beijing Jiaotong University
Beijing, Haidian District, China

Junwei Fu
School of Mechanical and Electronic Engineering, Beijing Jiaotong University
Beijing, Haidian District, China
E-mail: jwwang@bjtu.edu.cn, 20125999@bjtu.edu.cn
www.bjtu.edu.cn

Abstract

Corner detection is a basic task in the field of image processing and computer vision. Aiming at the problem that the existing corner detection algorithms based on curvature scale space (CSS) can not obtain the required corners accurately, a corner detection algorithm based on angle is proposed. Firstly, the appropriate threshold is set by CSS algorithm to extract the candidate corners, and then the angle value of candidate corners is calculated at long-distance pixels, and the threshold is used to filter to obtain the final required corners. Experimental results show that the algorithm can accurately extract the required corners.

Keywords: Corner detection; CSS; IPAN; Angle calculation; Corner positioning.

1. Introduction

As a part of the supply chain, logistics takes warehousing as the center to promote production and keep pace with each other; Logistics is to meet the needs of customers and realize the process of relevant goods from origin to consumption at the lowest cost through transportation, storage and distribution. As an important link in the logistics process, the efficiency and automation of loading and unloading process greatly affect the production cost in the transportation process. Therefore, the realization of full automation in the process of loading and unloading is of great significance to improve the efficiency of logistics transportation and reduce the transportation cost. In the process of loading, the accuracy of vehicle parking pose measurement is a very important link to realize automatic production. Corner detection is an indispensable step in the process of vehicle pose measurement through visual technology. Corner is an important feature in two-dimensional image. Corner recognition provides important information for many applications in vision based fields, such as target detection, motion detection , scene analysis. They point out the existence and location of objects, so as to narrow the scope of searching objects and make it easier to interpret images. Therefore, the corner detection is also particularly important. In the process of vehicle loading and positioning, the accuracy of corner detection is an important problem that must be solved, which has a vital impact on the calculation of vehicle posture.

2. Corner detection algorithm

Corner detection is a basic subject in the field of machine vision and computer vision. At present, there is no precise mathematical definition of corner. The following three kinds of points are usually called corner points: the first is the intersection of more than two edges, the second is the point where the brightness change in all directions on the image is large enough, and the third is the point of maximum curvature on the edge curve . The existing corner detection algorithms are mainly divided into three categories: model-based method, gray-scale method and contour based method. The corner detection algorithm based on contour has the advantage of small positioning error. Before corner detection of the image, edge detection is carried out first, and then corners are detected on the detected contour rather than on the whole image. By identifying the significant points on the contour curve (generally selected as the maximum curvature points). The key of contour based corner detection algorithm is to construct an efficient corner response function.

Mokhtarian et al. proposed a corner detection algorithm based on CSS, which is also a landmark algorithm in this field. This algorithm has good robustness, and introduced a tracking method in the detection process to ensure that corner detection is not affected by noise. Subsequently, more scholars proposed corner detection algorithm based on CSS, and Zhong et al. proposed DCSs algorithm, which reduces the amount of calculation and improves the efficiency of the algorithm. Zhang [6] and others proposed a robust multi-scale curvature product detection algorithm by analyzing the curvature behavior of multi-scale space. The corner detection algorithm based on CSS technology usually faces two problems: ① the algorithm is sensitive to the local change of curve and noise, which may lead to poor detection effect. ② Select appropriate Gaussian smoothing parameters. In the process of practical application, there are more problems. Although the concept of angle seems to be very clear literally, there is no generally accepted mathematical definition at present.

2.1. *CSS algorithm*

CSS (curvature scale space) corner detection algorithm is one of the most classic corner detection algorithms based on contour. CSS technology is suitable for extracting a continuous scale of curvature features from the input contour. In the detection process, Canny edge detection algorithm is used to extract the edge contour. Then parameterize the contour curve C:

$$C(u) = (x(u), y(u))$$

An evolved version C_σ of C can then be computed. C_σ is defined by:

$$C_\sigma = (X(u,\sigma), Y(u,\sigma))$$

where

$$X(u,\sigma) = x(u) \otimes g(u,\sigma)$$
$$Y(u,\sigma) = y(u) \otimes g(u,\sigma)$$

Then calculate the curvature through the curvature formula, which is defined as:

$$K(u,\sigma) = \frac{X_u(u,\sigma)Y_{uu}(u,\sigma) - X_{uu}(u,\sigma)Y_u(u,\sigma)}{[X_u(u,\sigma)^2 + Y_u(u,\sigma)^2]^{1.5}}$$

where

$$X_u(u,\sigma) = x(u) \otimes g_u(u,\sigma)$$
$$X_{uu}(u,\sigma) = x(u) \otimes g_{uu}(u,\sigma)$$
$$Y_u(u,\sigma) = y(u) \otimes g_u(u,\sigma)$$
$$Y_{uu}(u,\sigma) = y(u) \otimes g_{uu}(u,\sigma)$$

$\otimes$ is the convolution operator and $g(u,\sigma)$ denotes a Gaussian of width σ.

2.2 . IPAN algorithm

IPAN (image and pattern analysis group) is a very famous corner detection algorithm, which is based on the physical angle of points on the curve. The algorithm draws a triangle on the curve, in which one point is a candidate corner and the other two points are points with the same distance from the candidate corner. As shown in Figure 2.1, {P1, P2, P3... PN-1, PN} is a group of continuous points on the curve. The algorithm determines the final corner in two steps. Firstly, the triangle angle formula is used to calculate the angle α Compared with the threshold value, PI in the figure is used as the candidate point, pi-k and PI + k are used as the other two points of the triangle. If the triangle meets the following formula, this point is used as the candidate corner point.

$$d_{\min}^2 \le |P_i - P_{i+k}| \le d_{\max}^2$$
$$d_{\min}^2 \le |P_i - P_{i-k}| \le d_{\max}^2$$
$$\alpha \le \alpha_{\max}$$

where, $|P_i - P_{i+k}| = |a| = a$ and $|P_i - P_{i+k}| = |b| = b$ are the distance P_i to P_{i+k} an P_i to P_{i-k} ,the angle can be calculated by cosine formula α,Although the default value of $d_{\min}^2$, $d_{\max}^2$ and is given $\alpha_{\max}$ in this way, the value of the equally important K is not given. Similarly, only one K value can not well estimate the angle of Pi point.

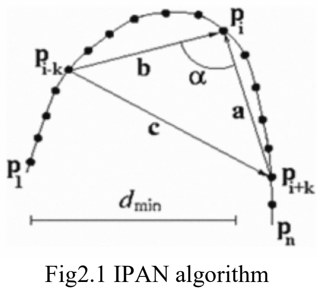


Fig2.1 IPAN algorithm

3. Paper Algorithm

The algorithm steps are as follows:

Step1:Use Canny operator to detect the edge of the image, and use Gaussian kernel function to smooth the curve.

Step2: Extract the contour of the edge detected in step 1, fill the contour gap, and find the T-shaped corner of the contour
Step3:Calculate the curvature value of edge points on a large scale, filter false corners and false corners with appropriate threshold, and get all candidate corners.
Step4:Improve the positioning by tracking corners in a small scale to obtain the final corners under the CSS algorithm
Step5:The final corner extracted by CSS algorithm is taken as the candidate corner, and the Ipan algorithm is used to take a larger value for the threshold K, so as to eliminate the corner generated by local noise to obtain the final corner.

3.1. *Canny edge detector*

This method uses Canny edge detector to extract image edges.Canny proposed three criteria for the evaluation of edge detection operators:

1. High detection rate: Edge operators should only respond to edges. The detection operator does not miss any edges and should not mark non-edges as edges.
2. Position accurately: The distance between the detected edge and the actual edge should be as small as possible.
3. Clear response: only one response for each edge, only one point.

3.2. *Obtaining candidate corners*

After using the Canny algorithm to extract edge contours, gaps may occur at some contiguous edges, and the CSS method may not be able to find T-shaped corners due to contour gaps (Figure 3.1). When the edge extraction method reaches the end of the contour, it performs two checks:

1. If the endpoint is nearly connected to another endpoint, fill the gap and continue the extraction.
2. If the endpoint is nearly connected to an edge contour. but not to another endpoint, mark this point as a T-junction corner.

In the process of extracting corners with CCS algorithm, these corners are defined as the local maximum of the absolute value of curvature. On a very fine scale, there are many such maxima due to the noise on the digital contour. With the increase of scale, the noise is smoothed, leaving only the maximum value corresponding to the real edge angle. CSS corner detection method finds corners at these local maxima. With the change of the edge contour after smoothing, the actual position of the corner will also change. If the detection is implemented in a large range, there will be a great deviation in the positioning of diagonal points. CSS algorithm overcomes this problem by tracking corners to ensure that corner detection is not affected by noise. Then take the smaller one σ value and check the same corner at the low scale. Therefore, the position at the corner may be updated until the scale is very low and the operation is very limited. This improves the positioning effect and the calculation cost is also very low, because it is lower than σ_{high} The curvature value of high scale does not need to be calculated at each contour point, but only in a small area near the detected corner.

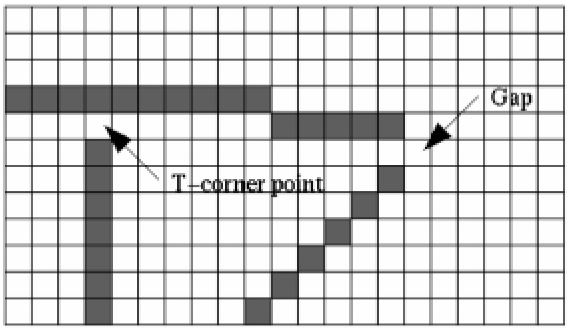


Fig 3.1 The two cases of gaps in the edge contours.

3.3. *Final corner determination*

As shown in Figure 3.2, among the candidate corners obtained by the CSS corner detection algorithm, because the algorithm is sensitive to local changes and noise on edge curves, and because of environmental factors, it is impossible to extract only the edge contours that we need according to the threshold setting during edge extraction. So only a few of the candidate corners extracted by the CSS algorithm are what we need. Analysis of a large number of images shows that the required corner points are concentrated on a large contiguous edge curve. Based on the inspiration of IPAN algorithm, a larger threshold K is selected to remove some corner points and get the final required corner points.

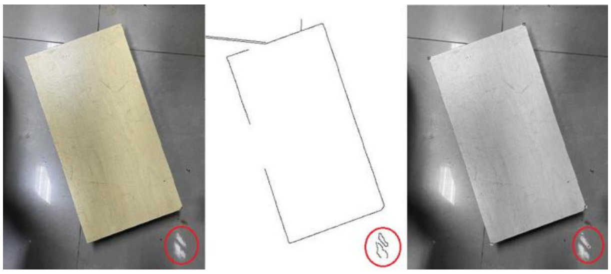


Fig 3.2 Image of CSS algorithm corner detection

4. Experimental Results

This algorithm will detect different images, and compare the effect of CSS algorithm, so as to intuitively understand the function of the algorithm.

(First column left is original picture,Fist column right is edge picture;Second column left is CSS,Second column right is the algorithm results in this paper)

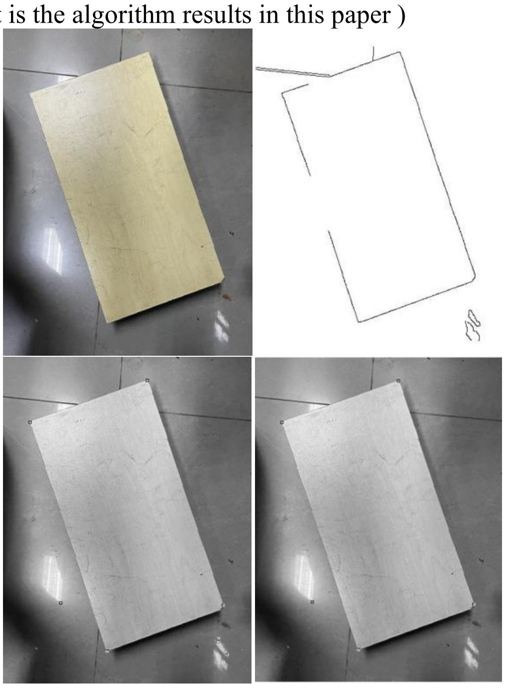


Fig 4.1

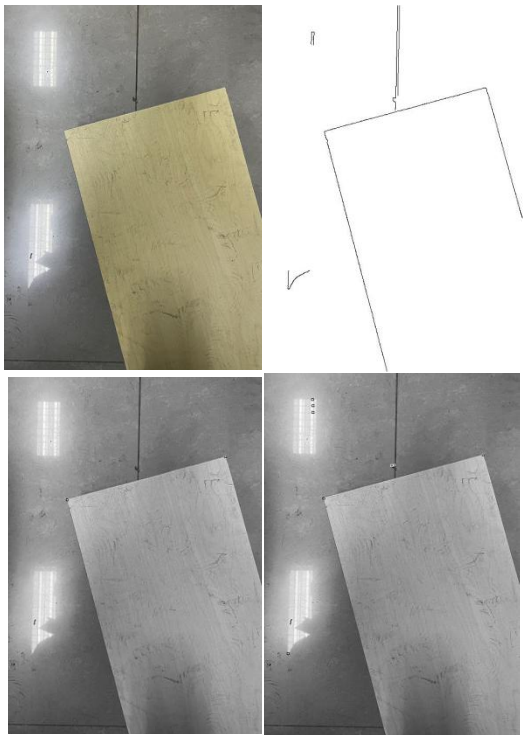


Fig 4.2

Fig 4.3

5. Conclusion

This paper introduces an angle-based corner detection algorithm. Firstly, the candidate corner points are obtained by using the CSS corner detection algorithm, and the corner points generated by the image edge are proposed to ensure that the required corner points are in the set of candidate corner points and minimize the unnecessary corner points. Then, the angle of each candidate corner point is measured at a high threshold through the angle threshold. Corners that do not meet the requirements are eliminated. The experimental results show that the algorithm can extract the desired corner points accurately. However, it does not have good reliability in complex background. Further research on more reliable corner detection algorithm in complex background and its application in practice will be carried out in the future.

Acknowledgements

This work has been supported by my mentor and thanks for his help in the research process.

References

1. Lebeda K, Hadfield S, Matas J, et al. " Texture-independent long-term tracking using virtual corners," IEEE Transactions on Image Processing,359-371, Vol25, No 1,2016.
2. Peng Lu, Yu Liang; Shuwei Chen, "Abnormal Behavior Detection of Video Groups Based on Corner Kinetic Energy"[J], Journal of Zhenzhou University(Engineering Edition). Vol.36, No 3,20-24, 2015.
3. Farzin Mokhtarian, Riku Suomela, " Roubust Image Corner Detection Through Curvature Scale Space, "IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 20,No 12,1998.
4. D. Chetverikov. A simple and efficient algorithm for detection of high curvature points in planar curves.2756:746-753,2003.
5. Rafi Md. Najmus Sadat, Zinat Sayeeda, et al."A Corner Detection Method Using Angle Accumlation," ICCIT2011,22-24.

Authors Introduction

Dr. Jiwu Wang

He is an associate professor, Beijing Jiaotong University. His research interests are Intelligent Robot, Machine Vision, and Image Processing.

Mr. Junwei Fu

He is a postgraduate in Beijing Jiaotong University. He research interest is visual positioning and image processing.

Motion Planning to Retrieve an Object from Random Pile

Shusei Nagato
Graduate School of Engineering Science, Osaka University, 1-3, Machikaneyama, Osaka, Japan
E-mail: nagato@hlab.sys.es.osaka-u.ac.jp

Tomohiro Motoda
Graduate School of Engineering Science, Osaka University

Keisuke Koyama
Graduate School of Engineering Science, Osaka University

Weiwei Wan
Graduate School of Engineering Science, Osaka University

Kensuke Harada
Graduate School of Engineering Science, Osaka University

Abstract

It is challenging to retrieve a target object from a randomly stacked pile by using a robot due to the occlusion of the target object. In this study, we propose a novel retrieval method in which a robot selects the viewpose to observe the occlusion part of the target object using the RGB-D images, and then selects the motion of grasping/dragging to retrieve the object depending on the configuration of the pile. We experimentally confirm that a robot effectively observes a pile with a complex configuration and successfully retrieves a target object.

Key Words: viewpose selection, motion planning, object recognition

1. Introduction

In recent years, robots have been expected to replace human workers in logistics warehouses. However, since many items are usually crammed onto a shelf in a logistics warehouse, it becomes challenging for a robot to pick an item from the shelf. Even if a robot tries to pick an item from the shelf, it may not be able to see the target product since the product is occluded by other items, or even if it can see the target product, its hand may not be able to grasp the product since the products are close together.

Robotic manipulation of a daily object in clutter has been studied by many researchers [1-6]. Most of the studies have focused on picking the topmost object [1][2], picking an object while avoiding its surrounding objects [3][4], and the push–grasp strategy [5][6]. However, these studies did not consider the problem of directly picking an object on which other objects are placed. If other objects are placed on top of the target object, it becomes difficult for a robot to identify the pose of this object because it is heavily occluded. In addition, even if the pose can be detected, the graspable area of the target

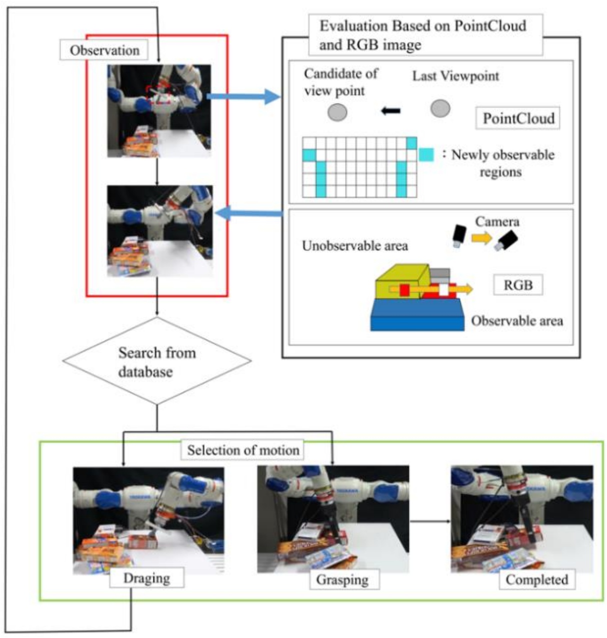


Fig1 Overview of this work

object is heavily limited owing to the occlusion. However, if a robot can directly pick the target object from among other objects placed on it, the robot does not need to remove those objects first.

In this study, we propose a motion planning method in which a robot repeatedly captures the image of a pile until it can grab the configuration of objects surrounding the target object (Fig. 1). Once the robot recognizes the configuration, it determines the manipulation strategy for picking the target object. If the robot assesses that the target object is not graspable, it drags the object before grasping it. In our proposed method, the robot first observes the workspace using an RGB-D sensor attached to its arm. At this time, the occlusion of the object is assessed using RGB-D information. When only point cloud information is used, it becomes difficult to reduce the occlusion of a specific target object. When the RGB image information is used, the viewpose can be moved to reduce the occlusion of a specific target object. Based on this result, we have searched for an effective grasping posture from the database to grasp the target object. If no effective grasping posture is found, the target object is dragged in the direction to reduce the occlusion by assessing the occlusion using RGB image features. After the extraction, we observe the object again and repeat the process until the target object is retrieved using the grasping operation. In this way, the target object can be retrieved even when it is difficult to recognize the state of the object or grasp the target object.

2. Related Works

Changjoo et al.[7] proposed a path generation method to retrieve a target object in a cluttered environment by once relocating objects out of the workspace where they introduced a graph expressing the collision possibility among objects to find a path. Sang et al. [8] used this method to retrieve the target object by determining the order of the objects to be relocated. However, in these studies, the target object cannot be retrieved when the objects cannot be relocated.

As for the research on planning a viewpose of vision sensor, Harada et al. [9] proposed a method to determine the sensor pose that maximizes the visibility of piled objects stored in a box. In addition, Motoda et al. [10] used this method to maximize the visibility [9] as an indicator to understand the pile's state. Although the viewpose selections used in these studies can reduce the occlusion included in the image of the pile, they are not always efficient because they are not optimal for observing the target object.

On the other hand, this research proposes a novel approach where a robot selects the grasping and dragging motions to select the target object to retrieve the target object from the pile depending on the situation of the pile. In addition, viewpose selection to reduce the occlusion of the target object is performed using both an RGB image and point cloud.

3. Proposed Method

This study proposes a motion planning method for a robot to grasp and retrieve a single target object which shape is known in advance. We use a dual-arm robot where one arm is equipped with a sensor to acquire RGB-D images from an arbitrary viewpose and a finger to drag an object, and the other arm is equipped with a two-fingered gripper to grasp the target object.

Fig. 1 shows the outline of our proposed method. First, we capture the RGB-D image of the target object by selecting a viewpose that can focus on its occluded part. Then, according to the observation results, a robot selects either the grasping or the dragging operation. If the dragging operation is selected, then the above process is repeated until the robot retrieves the target object by selecting the grasping motion. In the following sections, we describe the details of this method. We first explain the criteria used to select the action according to the state of the pile. Next, we explain the method used to recognize the state of the target object using the RGB images. Finally, we explain the selection of viewpose for better observation of the target object.

3.1. *Manipulation selection based on state*

Fig. 2 shows a sequence of events that leads the target object to be extracted. First, the occlusion of the target object is detected based on the feature matching of the RGB images of the pile. Next, we select a viewpose with which the robot can observe the occlusion existing in the pile's RGB image [9]. Subsequently, we search for an effective grasping pose from the grasping pose database that was created in advance of executing the planner where the grasping pose database stores the pose of the hand relative to the object to stably grasp the object. If a grasping pose is successfully found, the target object is grasped using the identified grasping pose. If the target

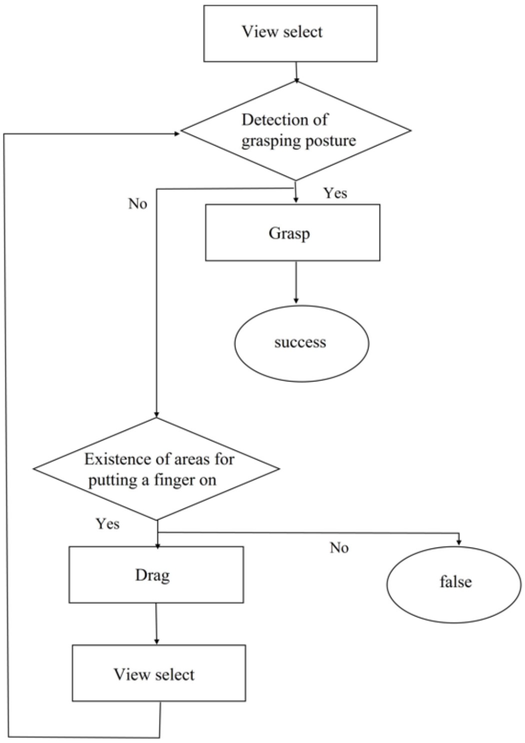


Fig.2 Flowchart of the proposed method

object is successfully retrieved, the task is completed. On the other hand, if a grasping pose cannot be found, we check whether or not there is enough area on the target object to place a robot's finger. If there is no enough area to place a finger, the task fails; however, if a finger can be placed on the target object, the robot performs the dragging operation. After the dragging operation, the occlusion of the target object is detected again based on the feature matching. If the occlusion area is still larger than the threshold, the observation is performed using the same aforementioned method. Otherwise, the target object is observed with the current viewpose. Based on the observation results, we repeat the sequence of operations until the robot successfully extracts the target object.

3.2. *State recognition of target object*

To extract the target object from the pile, we need to know the occlusion of the target object. As shown in Fig. 3, we acquire the RGB image of the pile using the current viewpose and extract the local features of the image by

Fig. 3 Flow of state recognition

using A-KAZE [11], and the image of the target object that we have prepared in advance. Next, feature matching is performed between two images, and the corresponding points included in the two images are obtained. In this process, random sample consensus (RANSAC) [12] is used to eliminate matching with outliers. Next, we perform the homographic transformation to align the target object in the original and its transformed images. The image in the center of Fig. 3 shows the result of the homography transformation. Subsequently, a match is detected based on the average of RGB values among nine pixels, including the pixels of interest and neighboring pixels. This is shown in the right figure of Fig. 3, where the area which is not drawn in yellow is occluded.

3.3. *Viewpose*

In this study, we use both RGB and point clouds where RGB is used to detect the target object while point cloud is used to identify occlusions in the pile’s image. This information is used to reduce the occlusions of the target object caused by other objects and move the sensor to the position where the target object is less occluded. The viewpose is selected from multiple candidates explained in the following.

We now explain how to construct candidates of viewposes and how to select one from multiple candidates. The view direction is determined such that the camera faces the center of the table. We assume a regular polyhedron at the center of the table and a set of lines passing through both the center of each face and the center of the polyhedron. We assume that the vision sensor is placed along this line with facing the center of the table. Here, the position of the vision sensor is selected to minimize the occlusion of the target object. First, the workspace is divided into three-dimensional grids [13], and the occupancy grid map [9] is created for each grid based on the point cloud information acquired with the current viewpose. The next position of the vision sensor is selected such that the largest number of

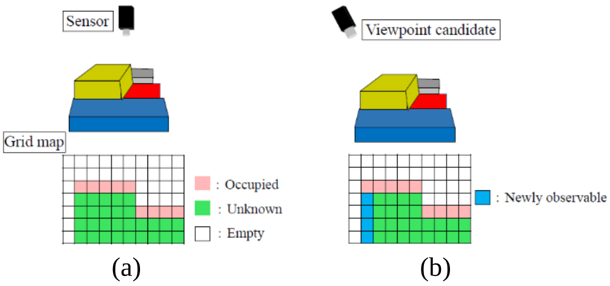


Fig.4 Next viewpose exploration by points clouds: (a)State of grids by last observations, and (b) Evaluation method in viewpose candidates

unobserved grids can be observed as shown in Fig. 4.

If only the point cloud is used, the effect of viewpose selection minimizing the occlusion of a specific object is limited[9]. Therefore, we select a viewpose making the occlusion of only the target object be visible by using RGB information. First, we determine the occlusion of the target object using the current viewpose. In this way, the center of the observable and unobservable part of the target object can be identified. Let the center of the observable and unobservable parts of the target object be O and A, respectively. In addition, let the ith horizontal position of the viewpose candidate be $B_i = (b_{xi}, b_{yi})$ (i = 1,2,⋯ ,n). In this case, it is expected that the unobserved part can be observed by moving the sensor following the vector from the unobserved part to the observed part with imposing the limitation on the viewpose candidates:

$$distance(O, A, (b_{xi}, b_{yi})) < l \quad (1)$$

$$\left| arccos \frac{\overrightarrow{OA} \cdot \overrightarrow{OB_i}}{|\overrightarrow{OA}||\overrightarrow{OB_i}|} \right| < \theta \quad (2)$$

where the $distance(O, A, (b_{xi}, b_{yi}))$ denotes the Euclidean distance between the line including both O and A and the ith position of the viewpose candidate. Equation (1) limits the viewpose candidates by the distance between the line connecting the two unobserved points and the observed point, whereas Equation (2) limits the direction of movement by the arrow shown in Fig. 5(a) where an example of movement is shown in Fig. 5(b). The overall flow of the shifting position of the vision sensor is explained in the following. The initial viewpose is selected randomly from viewpose candidates. Feature matching by A-KAZE is performed from this

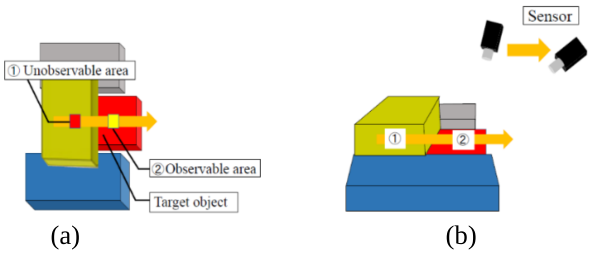


Fig.5 Exploration of next camera position by RGB image: (a)Determining the direction of the following viewpose, and (b) example of viewpose selection

viewpose. If the number of matches is small, it is assumed that the target object has not been found, and feature matching is performed again by using a viewpose with less occlusion based on the point cloud information. We assume that the target object is found when the number of matches exceeds a specific threshold, and the target object is directly observed. From the method shown above, the occlusion of the target object is checked, and the visible area can be obtained. When the observable area is larger than 80%, the viewpose selection is terminated, and action selection is performed. When more than 80% of the target object cannot be observed, the viewpose is selected based on the evaluation value in Equations (1) and (2).

4. Experiments

In this section, we describe the environment used in the experiments. We used a Motoman-SDA5F manufactured by Yaskawa Electric Corporation. For observation, we attached the Intel Realsense SR305 at the tip of the left hand. The left hand was used for the dragging operation, and the right hand was equipped with Robotiq 2F-140 Adaptive Grippers for the grasping operation. A workspace was set in front of the robot, and various objects were placed around the target object whose size and appearance were known.

4.1. *Results of the target object extraction*

Figs. 7(a)-(f) show the results of the proposed method In Fig. 7(a), the observation is performed by randomly selecting a viewpose from multiple candidates. The observed target object is shown in Fig. 7(b), where occlusion is evaluated by the RGB image. Since less than 80% of the target object was observed, the point cloud is

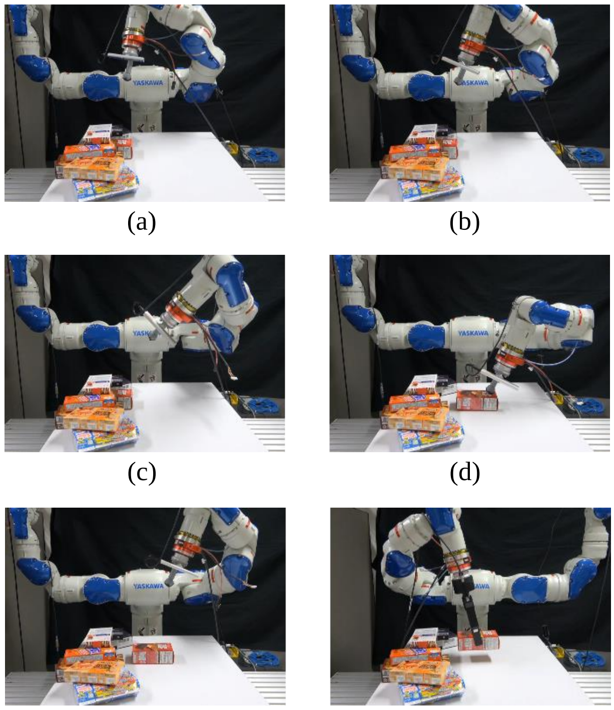
(a)Situation (b) Target object

Fig.6 Experimental environment

acquired, and the occlusion in the workspace is calculated. In Fig. 7(c), the target object is observed from the perspective of reducing the occlusion. There is no grasping posture in the initial state; therefore, the dragging operation is performed from the observation result of the current viewpose. In the dragging motion, the boundary between the observable and non-observable points is determined based on the result of the occlusion assessment; the finger is placed at the center of the observable point, and the dragging motion is performed in the direction perpendicular to the boundary. Fig. 7(d) shows the actual dragging operation. Because the position of the target object changes depending on the dragging, the viewpose candidates are calculated again. In Fig. 7(e), the occlusion of the target object is assessed by observing it again, and the observable area is more than 80%. Therefore, there is no need to reduce the occlusion area, and the viewpose selection is terminated. The grasping motion is performed as shown in Fig. 7(f). Because the grasping posture exists, the retrieval of the target object is completed.

4.2. *Consideration*

As shown in the results, we succeeded in grasping and extracting the target object from a pile using the proposed method. In this study, we confirmed that we could observe the target object using RGB-D information to position the sensor to observe the part of the target object that could not be observed using the previous viewpose. The experiment showed that it is possible to retrieve the target object after dragging it. However, the overall success rate was approximately 20%. This was because the estimation of the target object's occlusion was sometimes incorrect. Because only RGB information was used to estimate the occlusion of a target object, incorrect results were obtained when an object with an RGB value close to that of the target object was an obstacle or when the RGB value of the target object differed from that of the original image because of the influence of light condition. In the future, we improve the accuracy of the occlusion assessment of the target object using not only RGB images but also depth information.



(a) (b) (c) (d) (e) (f)

Fig.7 Flow of action: (a) Observation from initial viewpose, (b) Observation from above, (c) Observation of the occlusion area, (d) Drag, (e) Observation from initial viewpose, (f)Grasp

5. Conclusion

In this study, we proposed a method for selecting a viewpose to reduce the occlusion of a target object using RGB-D information of the pile. According to the observation results, we confirmed that the target object could be retrieved by combining the dragging and grasping motions where the oject could not be retrieved only by using the grasping motion.

As for future work, we plan to improve the accuracy of occlusion assessment by using RGB image by simuntaneously using the point cloud information. In addition, we estimate the posture of the target object, assess obstacles based on the observed information, and verify whether the target object can be retrieved even in an unknown environment. Furthermore, we plan to determine the direction and amount of dragging by considering the situation of objects piled up around the target object.

References

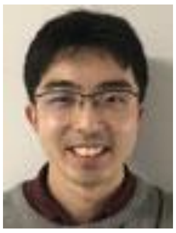
1. H. Zhang, X. Lan, X. Zhou, Z. Tian, Y. Zhang, and N. Zheng, "Visual Manipulation Relationship Network for Autonomous Robotics," in Proc. of 2018 IEEE-RAS 18th Int. Conf. on Humanoid Robots (Humanoids), 2018, pp. 118--125.
2. Y. Domae, H. Okuda, Y. Taguchi, K. Sumi, and T. Hirai, "Fast graspability evaluation on single depth maps for bin picking with general grippers," in Proc. of IEEE Int. Conf. on Robotics and Automation (ICRA), 2014, pp. 1997--2004.
3. D. Berenson, S. S. Srinivasa, D. Ferguson, A. Collet, and J. J. Kuffner, "Manipulation planning with Workspace Goal Regions," in Proc. of IEEE Int. Conf. on Robotics and Automation (ICRA), 2009, pp. 618-624
4. D. Berenson, S. S. Srinivasa, D. Ferguson, and J. J. Kuffner, "Manipulation planning on constraint manifolds," in Proc. of IEEE Int. Conf. on Robotics and Automation (ICRA), 2009, pp. 625-632
5. L. Berscheid, P. Meißner and T. Kröger, "Robot Learning of Shifting Objects for Grasping in Cluttered Environments," in Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS), 2019, pp. 612-618.
6. E. Huang, Z. Jia and M. T. Mason, "Large-Scale Multi-Object Rearrangement," in Proc. of Int. Conf. on Robotics and Automation (ICRA), 2019, pp. 211-218.
7. C. Nam, J. Lee, S. Hun Cheong, B. Y. Cho, and C. Kim, "Fast and resilient manipulation planning for target retrieval in clutter," in Proc. of IEEE Int. Conf. on Robotics and Automation (ICRA), 2020, pp. 3777--3783.
8. S. Hun Cheong, B. Y. Cho, J. Lee, C. Kim, and C. Nam, "Where to relocate?: Object rearrangement inside cluttered and confined environments for robotic manipulation," in Proc. of IEEE Int. Conf. on Robotics and Automation (ICRA), 2020, pp. 7791--7797.
9. K. Harada, W. Wan, T. Tsuji, K. Kikuchi, K. Nagata, and H. Onda. Experiments on learning based industrial bin-picking with iterative visual recognition," Industrial Robot: the international journal of robotics research and application, Vol. 45, No. 4, 2018, pp. 446--457.
10. T. Motoda, W. Wan, and K. Harada. Probabilistic action/observation planning for playing yamakuzushi," in Proc. of IEEE/SICE International Symposium on System Integrations (SII 2020), p. 150--155,2020.
11. P. Alcantarilla and T. Solutions, "Fast explicit diffusion for accelerated features in nonlinear scale spaces," in British Machine Vision Conference (BMVC), 2013.
12. R. Raguram, O. Chum, M. Pollefeys, J. Matas, and J. Frahm. Usac: A universal framework for random sample consensus. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 35, No. 8, 2013.
13. K. Nagata, T. Miyasaka, D. N. Nenchev, N. Yamanobe, K. Maruyama, S. Kawabata, and Y. Kawai, "Picking up an indicated object in a complex environment," in Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, 2010, pp. 2109--2116

Authors Introduction

Mr. Shusei Nagato

He received the BS degree from Engineering Science Department of Osaka University in 2021. From 2021, he has been a master student in Graduate School of Engineering Science, Osaka University.

Mr. Tomohiro Motoda



He received the MS degree from Graduate School of Engineering Science, Osaka University in 2020. From 2020, he has been a Ph.D. student in Graduate School of Engineering Science, Osaka University.

Prof. Keisuke Koyama

He received the Ph.D. degree from Graduate School of Informatics and Engineering, The University of Electro-Communications in 2017. During 2017-2019, he worked as a Project Assistant Professor, Graduate School of Information Science and Technology, The University of Tokyo. From 2019, he has been working as an Assistant Professor, Graduate School of Engineering Science, Osaka University.

Prof. Weiwei Wan

He received the Ph.D. degree from Graduate School of Information Science and Technology, The University of Tokyo in 2013. During 2013-2015, he was a postdoc researcher at CMU. During 2015-2017, he was a research scientist at AIST. From 2017 he has been an associate professor at Graduate School of Engineering Science, Osaka University.

Prof. Kensuke Harada

1. He received Ph.D. degrees in Mechanical Engineering from Kyoto University in 1997. During 1997-202, he worked as a research associate in Hiroshima University. During 2002-2016, he worked in AIST. From 2016, he has been working as a Professor at Graduate School of Engineering Science, Osaka University.

Design of Two-sided Gripper for Bin Picking

Maike He
Division of Mechanical Science and Engineering
Kanazawa University
Kakuma, Kanazawa, Japan

hemaike@stu.kanazawa-u.ac.jp

Tokuo Tsuji
Tatsuhiro Hiramitsu
Hiroaki Seki
Institute of Science and Engineering
Kanazawa University
Kakuma, Kanazawa, Japan

{tokuo-tsuji,thiramitsu,hseki,}@stu.kanazawa-u.ac.jp
www.kanazawa-u.ac.jp

Abstract

This paper presents the gripper and system for Bin picking. We have designed a two-sided gripper, which can grab from the inside of the part, and it can also grab from the outside of the part. Our gripper can grab parts with holes by inserting the finers to the holes. This gripper has a wide range of applicability.

Keywords: Bin picking, Gripper, Double-sided grabbing, Grab from inside, Various parts.

1. Introduction

Bin picking is a core problem in industry automation. The goal is that a robot pick-up known objects with random poses out of a bin using a gripper. Progress of bin picking technology has the good effect in solving the current shortage of labor.

The current bin picking has the following problems. [1] It is common to use one gripper for one kind of part, which will increase the cost. To increase the use in the production of many parts, a gripper which is capable of handling many parts is in demand. This research is dedicated to designing and manufacturing a double-sided gripper that can be applicable to pick up many kinds of parts. This kind of gripper can grab not only from the inside of the part, but also from the outside of the part.

Different parts can be grasped by this new gripper. The gripper can grab objects with hole by inputting the fingertip into the holes. It is effective for pick up only one object from bin. It can grab object of various sizes.

Among the target objects of this experiment, the smallest inner diameter is 6mm, the minimum outer diameter is 9mm, the maximum outer diameter is 62mm. So the target object is all parts with an inner diameter of 6mm or more, or an outer diameter of 9mm-62mm.

The research repeats two phases, design and experiment. We redesign the gripper according to the experimental results. In this way, the problems of the previous design can be found and improved, so that the final design can meet the requirements as much as possible.

2. Related works

Several robotic hands have been developed. There are many soft robotic hands been designed.[2] There are also many grippers being manufactured. For example, parallel gripper, biologically inspired grippers, three finger grippers, compliant gripper.[3] The above-mentioned various grippers have a good effect on grasping the target object. Our gripper focuses the object with holes. The part with holes does not refer to a particular part, but a different part with holes. Our gripper has a wide range of applicability.

3. Design Gripper

The double-sided gripper is a gripper that can be grabbed from the inside of the finger and the outside of the finger, as shown in Figure 1.

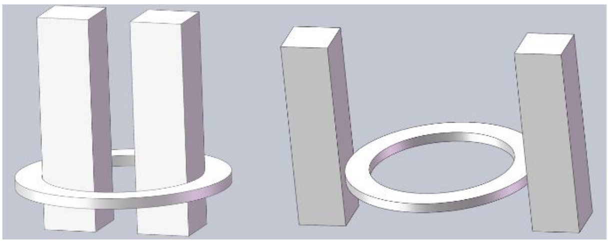


Fig.1. Schematic diagram of double-sided gripper

3.1. *Design*

In order to solve the problem of the gripper grabbing various parts, we consider adding the front end to the original shape of the gripper.[4] The design of the front end is for small parts with holes, and it has good performance for small parts. Because of the small design of the front end, it can enter the holes of parts and grab the parts from the inside to ensure that only one part can be grabbed at a time.

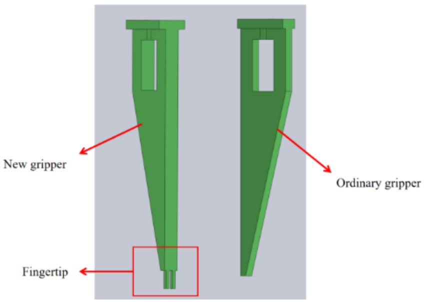


Fig.2. Gripper comparison

We design the structure shown in Figure 2. Compared with general grippers, there is an extra front-end design on the gripper. And the target parts of this research are mainly parts with small holes, so this gripper works well for small parts.
The front-end design has many parameters, such as front-end size, front-end material, front-end quantity, etc.[5,6]

4. Experiment

A total of 8 different parts were used in this research, as shown in Figure 3.

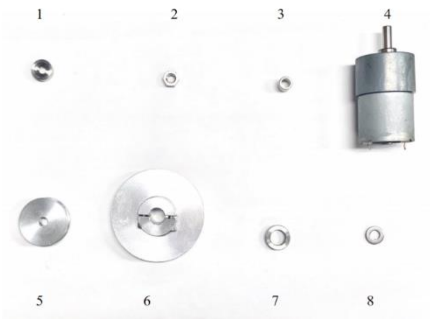


Fig.3. Parts

We tested the performance of 3 different front-ends. Then analyze and improve the design based on the results. Figure 4 is the three front-end models tested in this experiment. The difference between these three models is the size and angle. The front end of model A is 3mm in size. Models B and C are 2mm. Model A is a 60-degree triangular section, model B has no angle, and model C is a 60-degree quadrilateral section.

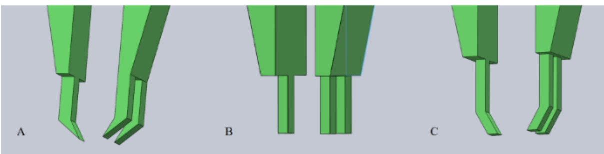


Fig.4. Fingertip model

The real Bin Picking process is tested in the experiment. We put many parts in a messy box and conduct a grasping experiment. As shown in Figure 5.

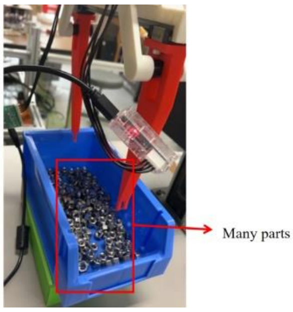


Fig.5. Actual experiment

The experimental results are shown in Figure 6. Four parts are grabbed from outside. The part1 is grabbed by the model C from inside. The other models cannot grab object from inside because the front end is too large to go deep into the hole. Each part carried out 10 grasping experiments, and the number of grasping was recorded. Success here refers to picking up parts, no matter how many. The average is the total number of crawls divided by the number of times. We analyze the data, take part 1 as an example. The success rate of models A and C is higher than that of model B, and the probability of picking one of model C is higher than that of model A, so for part 1, model C is the optimal model. After analyzing all data, it can be concluded that model C is the best model of them. It should be noted that the best model mentioned here is only a conclusion drawn based on the experimental data, not the final conclusion. Among the experimental results, model B has never succeeded in grabbing part 4, and it can be seen that the angle of front end is important. Because if the finger surface is straight, the part will fall off very easily due to lack of supporting force.

Model	Parts	Success rate	Average	Grabbed only 1 rate
Model A	1	70%	1.2	30%
	2	100%	2.2	40%
	3	80%	1.1	50%
	4	10%	0.1	10%
Model B	1	20%	0.2	20%
	2	100%	2.4	40%
	3	50%	0.8	30%
	4	0%	0	0%
Model C	1	70%	0.9	60%
	2	90%	1.4	50%
	3	60%	0.6	60%
	4	50%	0.5	50%
	1 (inside)	50%	0.5	50%

Fig.6. Actual experiment date

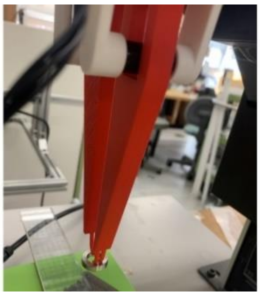


Fig.7. Grab from the inside

The size of the front end was too large, which prevented it from extending into the holes of many parts. Because the gripper is made of resin, the model below 2mm cannot be manufactured, so the experiment of grasping from the inside only tested the No. 1 part. Grabbing from the inside is shown in Figure 7.

4.1. *Improved front end*

Through experiments and improvements, We solved the problems that appeared before and designed a model as shown in Figure 8. In order to make a smaller front end and ensure strength, we changed the material. The material of the metal is SUS303. The size of the front end is 2mm, 2mm, 1mm. The size of the front end also has a gradient reduction, and the smallest place is 1mm. This way even the smallest hole with an inner diameter of 6mm can be entered. And there is enough contact area when grabbing large parts. Quadrilateral with an angle of 60 in section. The smallest positions are rounded on the outside of the front end. This helps to secure the parts.

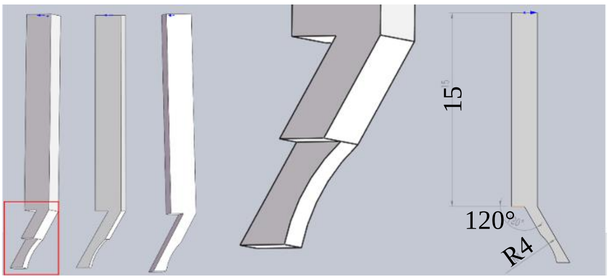


Fig.8. Fingertip design

This new model can grab all parts from the outside and all parts with holes from the inside. And the comprehensive success rate can reach 80%. Success here means grabbing 1 part. As shown in Figure 9, it shows how this model grabs the smallest part from the inside and the heaviest part from the outside.

Fig.9. Gripper grabs parts

5. Conclusion

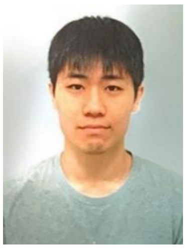
In this research, we designed a gripper that can hold various objects. This gripper can be gripped inside and outside the gripper. This gripper can grab small parts.

References

1. Fujita M, Domae Y, Noda A, et al. What are the important technologies for bin picking? Technology analysis of robots in competitions based on a set of performance metrics[J]. Advanced Robotics, 2020, 34(7-8): 560-574.
2. Nishimura T, Mizushima K, Suzuki Y, et al. Variable-grasping-mode underactuated soft gripper with environmental contact-based operation[J]. IEEE Robotics and Automation Letters, 2017, 2(2): 1164-1171.
3. Zhang B, Xie Y, Zhou J, et al. State-of-the-art robotic grippers, grasping and control strategies, as well as their applications in agricultural robots: A review[J]. Computers and Electronics in Agriculture, 2020, 177: 105694.
4. Tajima S, Wakamatsu S, Abe T, et al. Robust bin-picking system using tactile sensor[J]. Advanced Robotics, 2020, 34(7-8): 439-453.
5. Fraś J, Maciaś M, Czubaczyński F, et al. Soft flexible gripper design, characterization and application[C]//International Conference on Systems, Control and Information Technologies 2016. Springer, Cham, 2016: 368-377.
6. Lu Y, Xie Z, Wang J, et al. A novel design of a parallel gripper actuated by a large-stroke shape memory alloy actuator[J]. International Journal of Mechanical Sciences, 2019, 159: 74-80.

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.

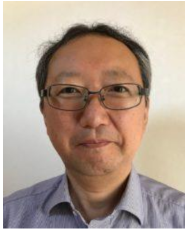
Dr. Tokuo Tsuji

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.

Dr. 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.

Dr. 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.

Training Data Augmentation for Semantic Segmentation of Food Images Using Deep Learning

Takayuki Yamabe
Department of mechanical engineering, Kanazawa University, Address Kanazawa, Ishikawa, Japan

Tatsuya Ishichi
Tokuo Tsuji
Tatsuhiro Hiramitsu
Hiroaki Seki

E-mail: yamabetakayuki@stu.kanazawa-u.ac.jp
http://as.ms.t.kanazawa-u.ac.jp

Abstract

We propose a training data generation method for semantic segmentation of food ingredients. Training data for semantic segmentation requires a large amount of effort for a human to carefully paint the boundaries of objects. Therefore, we propose a method to automatically augment appropriate training data by adding image composition processes for images of only one type of food. In our experiments, we confirmed the effectiveness of each data augmentation.

Keywords: data augmentation, food recognition, semantic segmentation, deep learning

1. Introduction

Food recognition from image is in demand for health care and food product quality control. The recognition of general food categories has reached practical accuracy and is being used in several applications[1]. On the other hand, the recognition of quantitative balance of food ingredients is an unresolved issue. One of the methods for recognizing food balance is semantic segmentation based on deep learning, which we focus on in this study. Semantic segmentation is a process that determines the region of an object in pixels for an image.

In order to achieve highly accurate image recognition, it is necessary to have a large amount of training data which is generated by human hand. In response to this, there has been research into making data creation more efficient [2], and research into data expansion, which ensures data diversity by transforming a small number of data [3]. In this study, we propose a training data generation method that is effective for semantic segmentation of food ingredients.

We define an image of mixed food as an image of several kinds of food, and an image of one kind food as an image of only one kind of food. For food balance recognition, we need to identify the images of mixed food, and the images needed for training are also images of mixed food. The labeling of images of mixed food requires manual and careful painting of the boundaries of the food ingredients, which is a large labor-intensive process. On the other hand, the images of one kind food can be labeled automatically based on the difference in color between the background and the food ingredients. Therefore, the labeling process can be easier if images of one kind food are combined to produce composite images like images of mixed food. This method of generating training data using composite images has been used in binarization tasks such as " Training Data Augmentation for Hidden Fruit Image Segmentation by using Deep Learning" [4], but has not been applied to semantic segmentation. Therefore, we propose and evaluate a data augmentation method using image composite for learning semantic segmentation.

In this study, we combine image composition with basic data augmentation methods such as cropping, flipping/rotation, and color manipulation to generate images suitable for learning multiple food ingredients. In the experiment, we checked the effectiveness of each processing.

2. Data augmentation method

2.1. *Cropping*

The purpose of this process is to generate various variations of food arrangement from a single image. A random area of 1000x1000 pixels is cropped from the input image.

2.2. *Horizonal Flipping and Rotation*

The purpose of this process is to increase the variation in the direction of rotation of the food. In the flipping process, the input image is flipped left to right. In the rotation process, the input image is rotated by affine transformation in the range of 0 to 360 degrees around the image center. In a series of processing, the input image is flipped with a probability of 50 percent, followed by rotation.

2.3. *Composition*

The purpose of this process is to deal with overlapping of foodstuffs. First, we cut out only the region of food parts of the input image. Next, we shrink the region so as not to introduce noise at the boundary with the background, and merge it into another input image. After merging, unnatural unevenness is created in the boundary area, so finally, blurring is applied only to the boundary area. Since the actual image to be recognized can be assumed to show only one type of food or all types of food, the compositing process is performed random times for one training image.

2.4. *Color Jittering*

The purpose of this process is to capture essential features such as patterns rather than the color of the food. Each of the three RGB values of the input image is multiplied by 0.8 to 1.2.

3. Experiments in food recognition

3.1. *Experiment Preparation*

Carrots, cabbage, sprouts, pork, green peppers, onions, and shimeji mushrooms were prepared and stir-fried over medium heat for 5 minutes before being photographed.

The shooting conditions were as follows: Panasonic DMC-FZ200 camera, f-number 8.0, SS 1.0, ISO sensitivity 100, distance between the camera and the plate 60cm, zoom 4x, under fluorescent light, and in forward light.

After cooking the food, we divided the food into two groups: one for images of one kind food and the other for images of mixed food. 28 images of one kind food were taken, 4 for each type of food, and 8 images of mixed food were taken, 4 for training data and 4 for test data. For the images of mixed food, 8 images were taken, 4 for training data and 4 for test data. The same food ingredients were not photographed more than once.

3.2. *Experimental method*

3.2.1. *Learning and Recognition*

In training, the model is trained on a training image of 1000x1000 pixels after data augmentation.

In recognition, test data is input to the trained model for recognition. The image input as test data is 4000x3000 pixels, which is different from the trained 1000x1000 pixels training image. Therefore, we split the input image, perform semantic segmentation on each of them, and then merge them into the original form to output the recognition result.

3.2.2. *Evaluation function*

As a metric for evaluating the performance of a multi-class classification mode, there is an evaluation function called categorical accuracy provided by the machine learning library keras [5]. The categorical accuracy is the number of pixels where the correct label matches the label with the highest output value of the model, divided by the total number of pixels. Here, since the images prepared in this study are easy to recognize the background. The more the background area of the test image is, the higher the categorical accuracy will be. Therefore, it may not be possible to properly evaluate the percentage of correct answers for the food ingredients. For this reason, we use our own evaluation function, food accuracy, which evaluates the correctness rate using the same method as categorical accuracy for only the pixels of foodstuff excluding the background.

3.3. Experimental conditions

The common training conditions are as follows
Model: U-Net [6], Number of training data: 10000, Number of epochs: 30, Batch size: 8, Loss function: Categorical Cross Entropy, Optimizer: Adam [7], Learning rate: 0.001.

As shown in Table **1**, The experiment was conducted in six different training conditions. Each condition is shown below.

- Training condition A
  Training with training data generated by color jittering, flipping/rotation, cropping.
- Training condition B
  Training with training data generated without color jittering in the condition of training A.
- Training condition C
  Training with training data generated without flipping/rotation in the condition of training A.
- Training condition D
  Training with training data generated without cropping in the condition of training A.
- Training condition E
  Training with training data generated without compositing in the condition of training A.
- Training condition F
  Training with training data generated without compositing using images of mixed food instead of images of one kind food in the condition of training A.

By comparing the results of training condition A with those of the other conditions, we can see the effects of each data augmentation process.

Table 1. Training conditions

Training	**Image type for training**	**Color jittering**	**Flipping and rotation**	**Cropping**	**Composition**
A	**One Kind**	○	○	○	○
B	**One Kind**		○	○	○
C	**One Kind**	○		○	○
D	**One Kind**	○	○		○
E	**One Kind**	○	○	○	
F	**Mixed**	○	○	○	-

Table 2. Food accuracy after Training (%)

Training	Food Accuracy
A	86.0
B	82.7
C	86.6
D	77.5
E	48.3
F	85.7

Fig. 1. Input Image

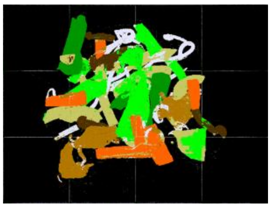


Fig. 2. Output Image

3.4. Experimental results

The food accuracy after training is shown in Table 2. An input image and the output image of the model of training A for it are shown in Fig. 1-Fig. 2.

4. Discussion of the experimental results

In Table 2, the food accuracy of training B, D, and E was lower than that of training A. This confirmed the increase of food accuracy by data augmentation of color jittering, cropping, and composition. On the other hand, there was almost no difference in food accuracy between training A and C. Therefore, we could not confirm the increase of food accuracy by data augmentation of inversion and rotation. This is because there was enough variation in the training data in terms of food orientation that flipping and rotation became meaningless data augmentation.

Since there was almost no difference in Food Accuracy between training A and F, it was confirmed that the combined image of images of one kind food can be used as training data equivalent to the image of mixed food.

5. Conclusions

We proposed the method for the semantic segmentation of food images allows the model to learn efficiently by composing images of one kind food that can be easily labeled. As for the data augmentation methods, we confirmed the increase of Food Accuracy by the data augmentation of color jittering, cropping, and composition.

References

1. Keiji Yanai. "Research Trends on Food Image Recognition." journal of the Japanese Society for Artificial Intelligence 34.1 (2019): 41-49.
2. Hirotaka Tanaka, and Hiroyuki Shinnou. "Efficient Construction of training data for Object Detection." The 34th Annual Conference of the Japanese Society for Artificial Intelligence. The Japanese Society for Artificial Intelligence, 2020.
3. Mikołajczyk, Agnieszka, and Michał Grochowski. "Data augmentation for improving deep learning in image classification problem." 2018 international interdisciplinary PhD workshop (IIPhDW). IEEE, 2018.
4. Ryoma Takai, and Kazuki Kobayashi. "Training Data Augmentation for Hidden Fruit Image Segmentation by using Deep Learning." The 33rd Annual Conference of the Japanese Society for Artificial Intelligence. The Japanese Society for Artificial Intelligence, 2019.
5. Francois Chollet, et al. Keras. https://keras.io, 2015.
6. Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015.
7. Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." arXiv preprint arXiv:1412.6980 (2014).

Authors Introduction

Mr. Takayuki Yamabe

He received his B.S. degree in Engineering in 2021 from the School of Mechanical Engineering, Kanazawa University in Japan. He is acquiring the M.E. in Kanazawa University.

Mr. Tatsuya Ishichi

He is acquiring the B.S. degree in Kanazawa University.

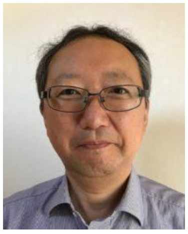
Dr. Tokuo Tsuji

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.

Dr. Tatsuhiro Hiramitsu

He is assistant professor of 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.

Dr. 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.

Suitable Error Recovery Process using Combined Evaluation Standards in Robotic Manufacturing Plant

Akira Nakamura
Department of Information Systems
Faculty of Engineering
Saitama Institute of Technology
1690 Fusaiji, Fukaya, Saitama 369-0293, Japan

and Kensuke Harada
Robotic Manipulation Research Group
Systems Innovation Department
Graduate School of Engineering Science, Osaka University
1-3 Machikaneyama, Toyonaka 560-8531, Japan

E-mail: akira-nakamura@sit.ac.jp, harada@sys.es.osaka-u.ac.jp
https://www.sit.ac.jp

Abstract

The number of manufacturing plants where industrial robots work is increasing. Therefore, errors during work are likely to occur. For big errors, it is often necessary to go back to the previous step and resume work. There are two issues: which step to return to and what kind of work to do from the point of return. In this paper, we will show that it is good to use a combination of multiple evaluation standards to decide the planning.

Keywords: error recovery, task stratification, error classification, robotic manufacturing plant

1. Introduction

Robotic automation plants have been increasing more and more in recent years. However, on the other hand, errors are more likely to occur. It is also important to optimize the recovery process after an error occurs.[1-5] For this reason, we have continued to study error recovery.

For several years, we have been studying the systematization of the error recovery theory. We proposed a new error recovery method based on the concepts of both task stratification and error classifications.[6-9] The main part of this method consists of fundamental elements with sequences of sensing, modeling, planning, and execution (Fig. 1). If an error occurs here, the process goes to the recovery part. In this part, the error cause is estimated, error is classified, system is corrected, and process is re-executed using the corrected system with an improved reliability.

Currently, deciding both the past step that the process should return to and the recovery planning after returning has become problematic. For this, we proposed a planning of the error recovery by deriving these two factors in consideration of cost.8 In this study, we proposed a planning method for error recovery derived using various evaluation standards.

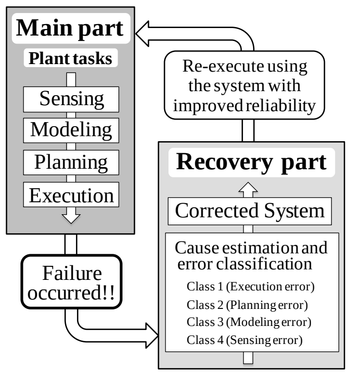
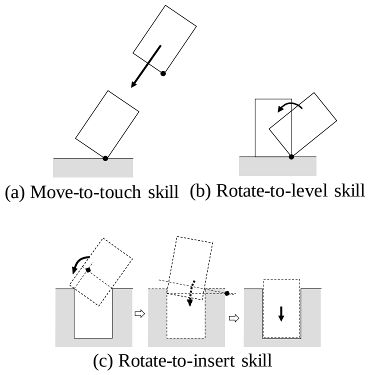


Fig. 1 Robot task system with an error recovery function



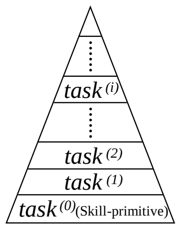


Fig. 2 Three fundamental skills

Fig. 3 Hierarchy of tasks

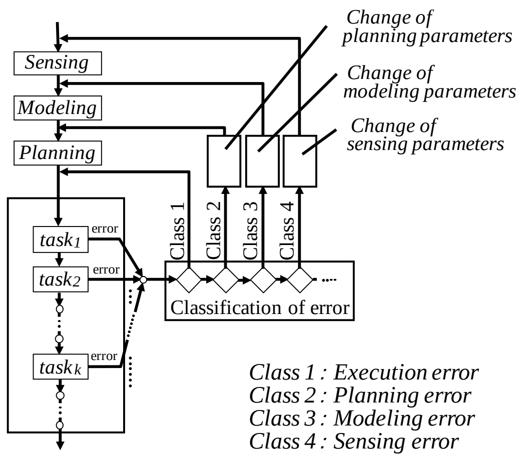


Fig. 4 Fundamental process flow with error recovery

The concept of skills, which are motion primitives, is described in Section 2. The error recovery technique is showen in Section 3. A method of recovery planning using plural evaluation standards is proposed in Section 4, and finally, a effective sample is presented in Section 5.

2. Concept of Skill

In this paper, the unit of motion is called a skill. This section explains skills which are components of human behavior and robotic motion.[10-12]

2.1. *Skill primitives*

A task such as an assembly consists of several motion primitives. Here, by analyzing human movements, we consider three skill primitives: "move-to-touch", "rotate-to-level" and "rotate-to-insert" shown in Fig 2. For tasks such as assembly, it is considered to be composed of the above-mentioned three important skill primitives and the resemble skill primitives.

2.2. *Stratification of tasks*

Figure 3 shows the hierarchy of robotic manipulation tasks. If we can ignore the servo layer, the first skill layer consists of behavioral units such as the important skill primitives mentioned above. Since tasks are composed by stratification, tasks have several layer in hierarchy in general.

3. Error Recovery

In the actual environment, unlike the ideal environment, various factors can cause errors in robotic machine performance. In this section, a concept of an error classification and our error recovery technique are described.

3.1. *Error classification*

Manipulation failures can be attributed to several types of errors. We have classified error conditions into four classes according to their possible causes: Execution, Planning, Modeling, and Sensing. Correcting the system based on the cause of these errors does not always solve the problem. For example, if the work environment changes significantly due to an error, it is necessary to return to the previous step.[6-9]

3.2. *Error recovery based on classification*

When an error occurs in an automated plant, the cause is first estimated and the equipment system makes appropriate corrections according to the estimated cause. The execution process then returns to the previous step and the task is restarted from the step (Fig. 4). As the equipment behavior is corrected, the probability of the same error occurring is reduced.

If a small error occurs, the process returns to the previous step at the lowest level (Fig. 4, Fig. 5). If a large error occurs, the process returns to the previous step in the upper hierarchy. In both cases, the process restarts from the previous step (Fig. 5).

3.3. *Candidate processes for recovery*

Consider candidates for possible error recovery processes. In the previous subsection, we showed that the steps that a process returns after an error occurs depend on its size. However, it is possible to back larger than the minimum required regression step.

Figure 6 shows a sequence of error recovery. It can be seen that the task consists of many subtasks from the start S to the goal G. In Fig. 6, errors occur along the way and how recovery is performed is shown. For more details, please refer to the Proceedings of ICAROB 2021.

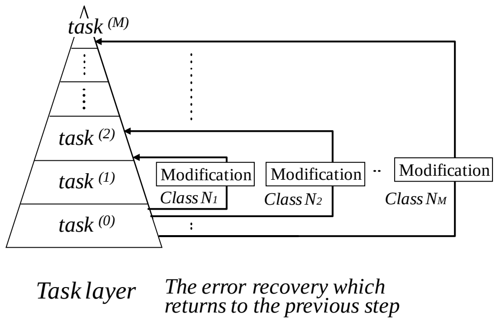


Fig. 5 The expression of task stratification and the process flow of the error recovery

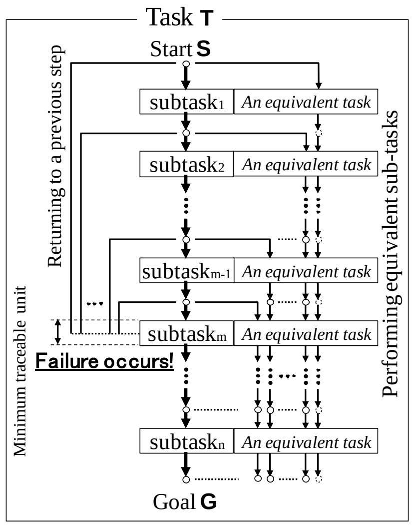


Fig. 6 Various processes of error recovery considered for a failure occurred in $subtask_m$

4. Evaluation Standards for Selection of a Recovery Process

The previous sections have shown that many recovery processes may be generated. Therefore, it is important to limit the route by determining both the previous steps to be returned and the recovery process after returning.

At ICAROB 2020, we considered practical costs as evaluation standards and proposed an appropriate return planning method.

In this paper, we will propose a method to derive the optimal return process using various evaluation standards.

4.1. *Evaluation standards*

At ICAROB 2021, we have considered the following eight evaluation standards to select a recovery process. Here is a brief explanation. Please see ICAROB 2021 for details.

(i) Cost
Material costs, part costs, electricity bills, and planning expenses

(ii) Time
Time required for the work

(iii) Reliability
Operations for Reliability

(iv) Safety
Operations for Safety

(v) Finishing
The beauty of outward appearance

(vi) Recovery data
A sufficient amount of data

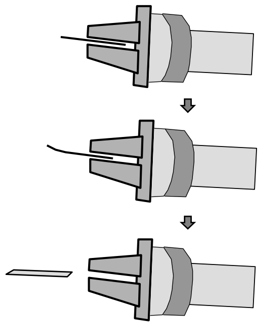


Fig. 7 The sticker sticking task

(vii) Tool
Unique tools for special tasks

(viii) Operator skill
Skills of expert craftsmen

4.2. *Simultaneous use of evaluation standards*

In ICAROB 2020, only cost was considered as an evaluation standard; in ICAROB 2021, the number of evaluation standards was increased to eight. However, it was supposed to be used independently for process evaluation. In this paper, several evaluation standards are used at the same time to show their advantages.

5. Precedence of Recovery Processes Based on Combined Evaluation Standards of Recovery

5.1. *Specific tasks to consider recovery*

Figure 7 shows the sticker sticking by the manipulation robot. Here, check whether the product tag sticker is attached in the correct position, and reattach it if necessary.

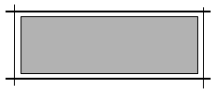


Fig. 8 Pasted exactly in place

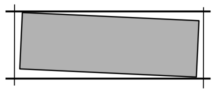


Fig. 9 Pasted slightly out of position

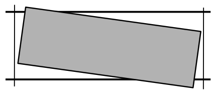


Fig. 10 Pasted far out of position

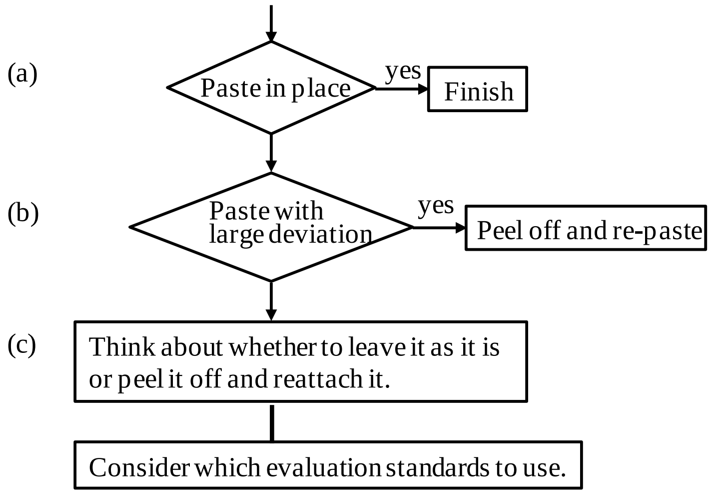


Fig. 11 The error recovery path for pasting position of the tag sticker

Figure 8 shows the case where it can be pasted in place. Figure 9 does not deviate significantly from the fixed position, but it is a case where it is necessary to confirm the actual situation and consider whether to leave it as it is or to reattach it. Figure 10 is pasted with a large deviation from the fixed position, and it is necessary to re-paste it.

Figure 11 summarizes the entire error recovery related to the task of sticking the product tag sticker in place in one figure.

5.2. *Task selection using evaluation standards*

Figure 11 shows the error recovery path for attaching the product tag sticker. In fact, in most cases, the fixed position pasting is successful (Fig.11(a)). If it is significantly misaligned, peel it off and reattach it (Fig.11(b)). The case between them is difficult (Fig.11(c)). There are two choices: leave it as it is or peel it off and reattach it. Here, we use three evaluation standards: cost Jc, time Jt, and finishing Jf.

Use a scraper to peel off the seal. A scraper is a very good tool when used lightly. Let's assume that the virtual unit of the evaluation standard is J_0. Based on this, (Case 1) and (Case 2) are compared. (Case 1) $J_c = 0$, $J_t = 0$, $J_f = 0$, Total $J = 0 + 0 + 0 = 0$, (Case 2) $J_c = K_cJ_0$, $J_t = K_tJ_0$, $J_f = K_fJ_0$, Total $J = K_cJ_0 + K_tJ_0 + K_fJ_0$. From the comparison between (Case 1) and (Case 2), we can see that it is better to leave (Case 1) as it is. (Case 2) is difficult to execute because it involves the risk of finishing.

6. Conclusion

If an error occurs in Robotic automated plants, the process advances to the recovery part. Many types of recovery processes can be selected. In ICAROB 2021, we showed a method to derive a suitable process using only one evaluation standard at a time selected from various standards.

However, in this paper, we used multiple evaluation standard at the same time to help select the recovery process. Specifically, the recovery process was selected using three of the eight evaluation standards. How to combine the evaluation standards used and how to find the optimal recovery process are future issues.

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, Using Various Evaluation Standards to Determine an Error Recovery Process in an Automation Plant, in Proc. ICAROB 2021, pp. 321-327.
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.
13.

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.

Underwater Acoustic Positioning Based on MEMS Microphone for a Lightweight Autonomous Underwater Vehicle “Kyubic”

Irmiya R. Inniyaka, Dominic B. Solpico
Department of Life Sciences and Systems Engineering, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196, Japan

Daiki Hamada, Akihiro Sugino, Rikuto Tanaka, Yuya Nishida, Kazuo Ishii
Department of Human Intelligent System, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196, Japan

E-mail:irmiya.inniyaka-350@mail.kyutech, solpico.dominic-bautista806@mail.kyutech.jp, hamada.daiki745@mail.kyutech.jp, sugino.akihiro424@mail.kyutech.jp, kai.yoshiki640@mail.kyutech.jp, y-nishida@brain.kyutech.ac.jp, ishii@brain.kyutech.ac.jp
https://www.kyutech.ac.jp

Abstract

Autonomous Underwater Vehicle positioning is important for accurate control and high-quality data collection. Conventional positioning systems are expensive. This paper describes the design and performance results of control using equations of motion and an inexpensive acoustic positioning system implemented on a lightweight AUV “KYUBIC” for Underwater Robotic competition in Okinawa 2021. The velocity used to estimate the self-position of the AUV is calculated form the equation of motion without integrating the acceleration. The acoustic positioning method is based on Super-short baseline (SSBL) principle. The system design comprises of self-made hydrophone module using MEMS microphone. The estimated distance and angle are integrated for dynamic control of AUV to locate the position of a Pinger.

Keywords: Autonomous Underwater Vehicle, Acoustic Positioning System, MEMS Mic, Underwater Competition

1. Introduction

The increasing use of Autonomous Underwater Vehicles (AUV) for underwater investigations ocean resources and installed infrastructure has made it an important tool. However, navigation remains a challenge to AUV performance because navigational accuracy is based on the accuracy of AUV position estimation underwater [1], [2]. Hybrid systems like the inertial navigation system (INS) and doppler velocity log (DVL) are expensive. Acoustic positioning systems (APS) are also expensive and are mounted on a ship and AUV for positioning. As such no available APS for low-cost AUV operation in shallow waters. This is one gap the Underwater Robotics Competition in Okinawa aims to address.

This paper describes the design of an inexpensive APS using MEMS microphone and performance results using a feedforward control on a lightweight AUV “KYUBIC” for Underwater Robotic competition in Okinawa 2021.

1.1. *Underwater Robotic Competition in Okinawa*

The Underwater Robotic Competition in Okinawa is an annual event aimed at promoting research solutions to the challenges of the underwater environment. It takes place in the actual sea. The KYUTECH Underwater Robotics team participated in the AUV Division: Intelligence/ Measurement Challenge. The purpose is to bridge the gap by designing an APS for low cost AUV operation in shallow waters. Details of the competition can be found in [3]. The following sections describes our team's strategy to meet the competition requirements.

2. Autonomous Underwater Vehicle: KYUBIC

KYUBIC is a boxed shape light weight AUV. It is a compact structure that can be operated by few people as an AUV or ROV. It has a modular architecture to facilitates for easy maintenance. A comprehensive documentation of KYUBIC development is presented in [4]. KYUBIC's control, strategy and hydrophone system are modified for the competition.

3. Hydrophone System

3.1. *System Configuration*

This section describes the configuration of the hydrophone system. The two hydrophones used in this system are connected to the Raspberry Pi 4 via an audio device (ReSpeaker 4-Mic Array for Raspberry Pi). This makes it possible to receive two devices simultaneously. The analog signals from the hydrophones are converted by the ADC inside the ReSpeaker and sent to the Raspberry Pi via I2S signals. The power supply is provided by an isolated DCDC converter to block noise from other devices.

3.2. *Acoustic Positioning Methods*

Figure 1 shows how APS estimates signal angle θ. By taking the incoming sound waves as parallel, θ is calculated from Equation (1) using the arrival time difference of the sound waves generated by Pinger, the baseline, which is the interval between hydrophones, and the speed of sound in water.

$$\theta = \sin^{-1}\frac{\boldsymbol{Sound\ velocity} \times \boldsymbol{Time\ difference}}{\boldsymbol{Baseline}} \qquad 1$$

To remove the noise of thrusters and sea, a band-pass filter is applied to extract the Pinger signal. The frequency of the Pinger is 21.164[kHz]. The sampling frequency of the hydrophone is 96[kHz]. Figure 2 shows the raw and processed data. The frequency of the bandpass filter was set to 20-22 [kHz], which is the range of the Pinger's transmission frequency.

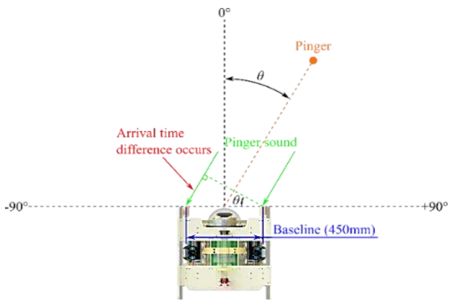


Figure 1: APS Architecture

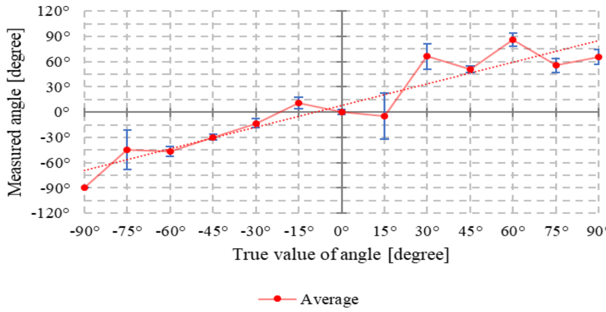


Figure 1: Result of Angle Measurement

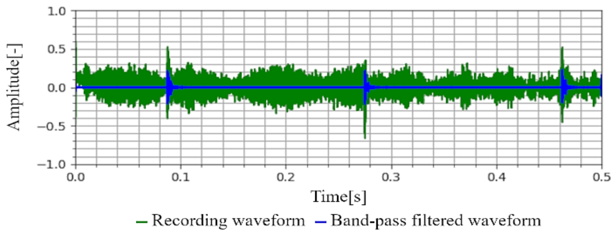


Figure 2: Band Pass Filtering

3.3. *Performance of hydrophone system*

The results of the performance evaluation of the hydrophone system are shown in Figure 3. The experiment was conducted in a pool with a diameter of 6.0[m] and a depth of 1.2[m]. The Pinger was placed in the center of the pool at a depth of 0.6[m]. The distance between the Pinger and the hydrophone array was 2.4[m]. The measurements were made at intervals of 15 degrees in the range of -90 to +90 degrees, five times each. Although the accuracy for angles other than 0° was low, the direction of the Pinger (left, front, right, etc.) could be estimated to some extent.

4. AUV Control Method

This section shows how our team controlled KYUBIC's navigation in the Intelligent/ Measurement Challenge Division of the competition. Our team aimed to reach the goal by combining position and thrust control. Until Pinger is detected. KYUBIC performs search missions by position control from target values for *Surge*, *Sway*, *Heave*, and *Yaw*.

Once Pinger signal is detected, the *thrust* is specified in the *X* and *Y* directions, while *Heave* and *Heading* control is maintained at the target value to reach the landmark. Each control was realized by feedback and feedforward control. Heave controlled the depth obtained from the Depth sensor, and Yaw controlled the heading obtained by integrating the Z-direction moment of the gyro sensor as feedback. Therefore, a feedforward control was designed using the acceleration calculated based on the equation of motion.

A description of the method for deriving the position for the feedforward control of Surge and Sway is presented. The general equation of motion of an underwater robot as expressed in the robot coordinate system is

$$\boldsymbol{M}\dot{\boldsymbol{v}}_{\boldsymbol{w}} + \boldsymbol{D}|\boldsymbol{v}_{\boldsymbol{w}}|\boldsymbol{v}_{\boldsymbol{w}} = \boldsymbol{F} \qquad 2$$

where $\boldsymbol{M}$: inertial coefficient, $\boldsymbol{D}$: fluid drag coefficient, $\boldsymbol{v}_{\boldsymbol{w}}$: water velocity, $\boldsymbol{F}$: thrust. The parameters $\boldsymbol{M}$, $\boldsymbol{D}$ in this equation were estimated by performing position control experiments. Acceleration was obtained from the equation of motion to which this estimated parameter was applied, and feedforward control was achieved.

Table 1: Parameters used for Experiment

Surge/Dive area	Surge (X)	0.5[m]
	Depth (Z)	0.2[m]
Pinger search area	Surge	0.5[m]
	Sway (Y & -Y)	1[m]
Search cycle limit	2	

5. Experimentation

5.1. *Performance of KYUBIC*

To evaluate KYUBIC's performance, experiments were conducted in a pool with diameter of 6[m] and depth of 1.2[m]. The Pinger is positioned 4.5[m] away from KYUBIC's position. The target values for the Pinger search and approach mission are as presented in Table 1. From Figure 4, the control performance for depth and heading controllers are shown. Measured value for depth has a short rise time to the target value overshoots reaching approximately 0.3[m] against the target value of 0.2[m]. Over time, the PID controller converges at the target value with smaller overshoots. This performance can be attributed inadequate tuning of the PID control. The heading control seems to keep the heading at the onset of the mission but becomes increasingly unstable when KYUBIC begins the Pinger search navigation. This is because the thrusters power navigates KYUBIC to the forward-left, forward, or forward-right direction. For better performance, adequate PID tuning of the heading controller is needed.

Figure 5 compares KYUBIC's position tracking from feedforward control to its actual position which is gotten from the DVL log. The feedforward controller navigates the *X* and *Y* target points in from start position of KYUBIC's mission. Ones the Pinger signal is detected, the thruster control program is initiated to approach the Pinger based on information from the hydrophone system.

Table 1 defines the parameters for the feedforward controller navigation.

From the Figure 5, positions [0,0] is KYUBIC's initial position. Firstly. KYUBIC's target surge value is 0.5[m] while diving to the mission depth position. However, although, the feedforward controller shows KYUBIC reaches the target position of 0.5[m], the actual position is 0.2[m] from the start position. This shows an error of 0.3[m] in the surge direction. The actual position also

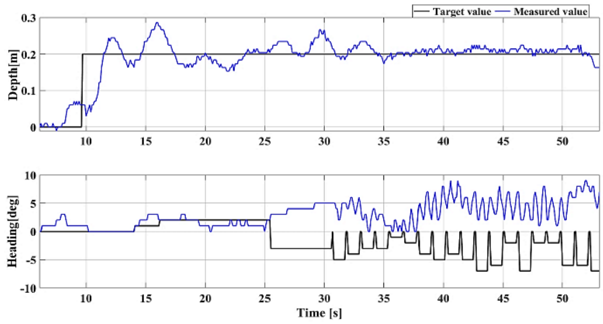


Figure 2: Controller Performance

dome drift to a maximum distance of -0.04[m] in the sway direction.

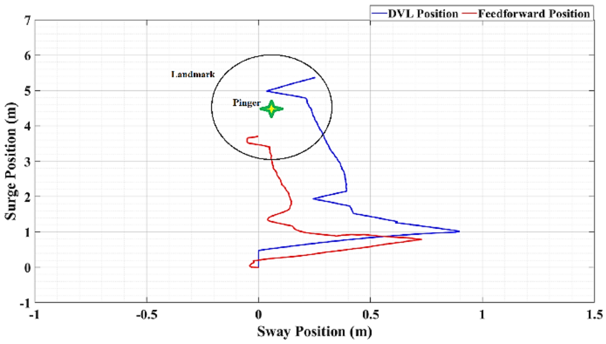


Figure 3: AUV Control and Actual Position

In the Pinger search navigation, the feedforward controller positions KYUBIC at 0.9[m] to the 1[m] target. However, it is positioned at 0.73[m] to the target position. Showing a position error of 0.27[m] without any drift error is log. This implies the feedforward controller will drift in either the surge or sway direction when there is no target value in that direction.

The Pinger is detected by the Hydrophone system is phase when KYUBIC's actual position is at 0.9[m] from the start position and the search program is terminated.

The Pinger Approach program is initiated, and the control strategy changes from *position control* to *thruster control*. In this control strategy, constant thrust in *forward-left, forward*, and *forward-right* direction based on the region of detected Pinger signal.

Figure 6 shows a relationship between KYUBIC's position and Hydrophone system information. Here, the logged information of AUV position, Hydrophone signal region and detection are presented with respect to the time of the experiment.

The previous section discusses the error in feedforward position to actual position. However, the time graph shows that KYUBIC state transition takes place at the same time in both the actual and controlled positions.

From the *Hydro_Detect* log in Figure 6, '1' means the Hydrophones can detect the Pinger signal and '0' means no detection. But, due to the limited pool size, the signal can be detected from all positions. Therefore, there is detection '1' all through the experiment.

Hydro_Region '0', '1', and '2' means the position of the Pinger is at the Left, Center, and Right for KYUBIC's position. Based on the region, controlled thrust is applied in line with the identified *Hydro_Region*. It is worth noting that due to the high detection within the pool, the search program is modified to implement Pinger search before receiving considering detected information as shown in the position arrows *A* and *B* of KYUBIC's position.

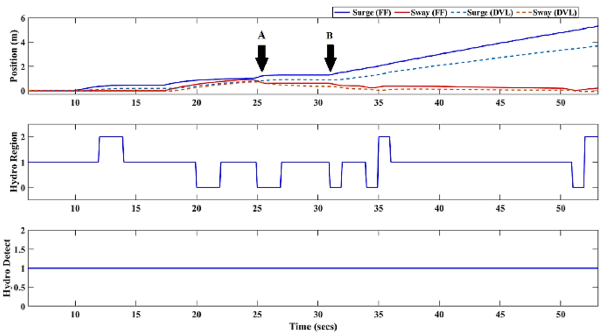


Figure 4: AUV Position vs Hydrophone System Information

In position *A*, the control strategy receives hydrophone information and switches from the *Position control* to *Thruster control*. The time for this change takes place between *A* and *B* where AUV position is seen to remain constant. At *B*, the region is Left, as such a *forward-left* thrust is applied as seen in Figure 5 from surge position of 0.9[m]. AUV navigate in the *forward-left* and then *forward* when the region is '1'. A sharp switch is seen from *forward-left* to *forward-right* when region changes from '0' to '2' at 35 [secs]. The long *forward* thrust (that drifts to the left) in Figure 5 shows the period of the region at '1'. AUV reaches Pinger position when region detects '1' and '2'while it is within the Landmark.

6. Conclusion

We developed an inexpensive acoustic positioning system for low cost AUV operation in shallow water. Using *feedforward control* and *thruster control*, we perform experiments with a KYUBIC. Results show hydrophone system can detect and successfully track the Pinger position. The feedforward controller has a 60% and 27% position error in the surge and sway directions respectively. It also drifts in either the surge or sway direction when a target value is not defined. Further improvements and tuning of the controllers as well as calculation to accurately detect Pinger angle and distance will make the APS more efficient. Sea trial are also needed to validate the performance of the hydrophone and control system because that is the basis of the developed system.

7. References

1. A. Miller, B. Miller, and G. Miller, "AUV position estimation via acoustic seabed profile measurements," AUV 2018 - 2018 IEEE/OES Auton. Underw. Veh. Work. Proc., pp. 1–5, 2018, doi: 10.1109/AUV.2018.8729708.
2. Y. Watanabe, Y. Ota, S. Ishibashi, T. Shimura, M. Sugesawa, and K. Tanaka, "An ocean experiment of inverse SSBL acoustic positioning using underwater vehicle OTOHIME," Ocean. 2016 MTS/IEEE Monterey, OCE 2016, pp. 1–5, 2016, doi: 10.1109/OCEANS.2016.7761498.
3. "7th Underwater Robot Competition in Okinawa," 2021.(Accessed Dec. 19, 2021) http://www.robo-underwater.jp/2021/rchp/JPN/index.php
4. T. Matsumura, Y. Uemura, K. Yanagise, Y. Tanaka, Y. Nishida, and K. Ishii, "Development of a handy autonomous underwater vehicle'kyubic,'" Proc. Int. Conf. Artif. Life Robot., vol. 2021, pp. 405–408, 2021, doi: 10.5954/icarob.2021.os22-4.

Authors Introduction

Mr. 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 Control, Underwater Robotics navigation and localization.

Mr. Dominic B. Solpico

He obtained his M. S. degree in Electronics Engineering at the Ateneo de Manila University, Philippines in 2015. He is currently taking his Doctoral degree course at Kyushu Institute of Technology, Japan. His research interests are in the fields of Intelligent Aquaculture, Wireless Sensor Networks (WSNs), and Field Robotics.

Mr. Daiki Hamada

He… He obtained his B.S. degree in Electrical Engineering from National Institute of Technology, Ube College. He is currently a Master's course student at the Department of Human Intelligence Systems, Graduate school of Life Science and Technology, Kyushu Institute of Technology, Japan.

Mr. Akihiro Sugino

He received his B.S. degree in Engineering in 2021 from the Faculty of Science and Engineering, Kyushu Sangyo University, Japan. 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. Tanaka Rikuto

He completed the production and electrical system technology course at Shikoku Polytechnic 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.

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.

Autonomous Underwater Vehicle with Vision-based Navigation System for Underwater Robot Competition

Kazuki Harada, Riku Fukuda, Yusuke Mizoguchi, Yusuke Yamamoto, Kota Mishima, Yoshiki Tanaka, Yuya Nishida, Kazuo Ishii

Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196, Japan

E-mail: harada.kazuki336@mail.kyutech.jp, ishii@brain.kyutech.ac.jp

Abstract

The underwater robot competition in international conference Techno-Ocean 2021 was held to advance underwater technology, in Dec. 2021. The competition consists of five leagues including AUV (Autonomous Underwater Vehicle) league, and Kyushu underwater robotics which is the Underwater Student Project team at our university that joint the AUV league using developed AUV. The wet test of the AUV league includes a Gate Pass mission which requires passing through an underwater green gate, Buoy Touch mission requiring contact with the yellow and red buoys, and Special mission requiring contact with the Pinger which has an unknown location. Our AUV navigates by image processing using underwater camera to achieve the mission. This paper explains the AUV system, the mission strategy and detail of image processing.

Keywords: Autonomous Underwater Vehicle, Robot competition, Image processing

1. Introduction

The Japan NPO Underwater Robot Network holds an underwater robotics competition every year to improve practical underwater technology and foster underwater engineers. Underwater Robot Competition Techno-Ocean Underwater Robot 2021 has a mission to perform line tracking and buoy touching. These missions simulate the automatic docking of an underwater robot to a base station on the seafloor for long-term observation, and the inspection of submarine cables in oil pipelines. This is expected to improve practical underwater technology.[1,2,3] Our laboratory has developed the AUV "KYUBIC", which can be operated by a few people, as an educational platform for participating in this competition.[4] In this paper, we report the system configuration of "KYUBIC", strategies and results of Techno-Ocean Underwater Robot Competition 2021 held in December 2021.

2. The Hovering Type AUV KYUBIC and Underwater Robot Competition

An KYUBIC is a hovering AUV developed for educational purposes. It was developed based on the concept of a versatile testbed with a hardware structure that allows for additional functions depending on the mission. The fuselage has a box-shaped structure, and a water-resistant container is stored inside the outer shell, making it highly fragile and portable.

2.1. *KYUBIC's equipment*

As mentioned earlier, KYUBIC, shown in Figure 1, was developed as an educational platform and is equipped with various sensors necessary to accomplish the mission of the underwater robotics competition. During the competition, actions such as diving and surfacing, forward and backward, etc. are performed using the ground speed obtained from the DVL (Doppler Velocity Logs) attached to the bottom of the KYUBIC, and information from the IMU sensor and depth sensor mounted on the control hull. In addition, the control hull is equipped with two camera sensors to obtain visual information about the front and bottom of KYUBIC, and image processing can be used to obtain information about obstacles. The system is also equipped with an RGB sensor to cope with lighting light fluctuations predicted in the real environment, and can measure illuminance, etc. However, it has not been implemented in software in the competition reported in this paper. In addition to the

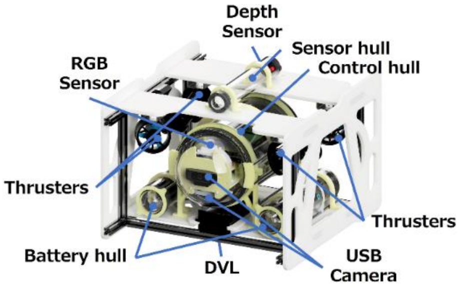


Fig. 1. Overview of KYUBIC

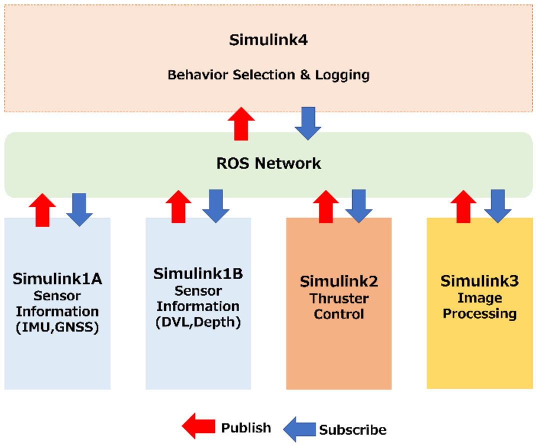


Fig. 2. system configuration

control hull, the AUV is equipped with sensor hull and two battery hulls. The sensor hull is equipped with a Wi-Fi module for wireless communication with an external PC, and a GNSS for self-positioning in competitions conducted in actual seas. The battery hull contains a control battery and a drive battery. By providing individual batteries, the risk of leakage can be distributed.

2.2. *KYUBIC's Software configuration*

The control system of KYUBIC is developed by using Mathworks' matlab/simulink and ROS. As mentioned earlier in 2.1, KYUBIC is equipped with many sensors, and it is necessary to acquire information using various sensors simultaneously while in action. In KYUBIC, a Simulink model is prepared for each sensor, and as shown in Figure. 2, a total of five Simulink (1~5) models are processed in parallel by returning ROS. There are two types of Simulink1 (Sim1), Sim1A and Sim1B. Sim1A acquires inertial data and position information from IMU and GNSS. In Sim1B, the ground speed is obtained from the DVL and the depth information from the depth sensor. In Sim2, PID control of the thrusters is performed, and in Sim3, image information acquired from the camera sensor is processed. Sim4 is a Simulink model that controls KYUBIC using the information obtained from each sensor by other Sims. Simu4 is designed using Stateflow, which is a tool for designing state transition control schemes, because Sim4 performs behavioral strategies for the KYUBIC which require complex state transitions. [4,5]

2.3. . *Competition Rules and Our Strategy*

According to the competition rules of the underwater robot competition 2021 AUV department, the robots challenge each mission set up in the indoor pool with autonomous motions. In the competition rules, the pool is divided into Area0, Area1, and Area2 as shown in Figure 3. AUVs start from (1)"Start position" and points are added by "passing through the gate" (2), "touching the buoy" (3) and “moving between areas”. The goal is reached when the AUV completes four round trips on the competition course. In this participation, we focused on the fact that each mission exists on a line, and developed strategies based on image information. Our goals were to complete missions such as "going back and forth between Area 1 and Area 2", "passing through the gate", "touching the buoy", and "surfacing at the end".

To achieve "going back and forth between areas" and "passing through the gate", we adopted line tracing as the mission strategy method. This is because, as shown in Figure 3, a buoy and a gate are located on a line at the bottom of the pool, and we thought that the robot could navigate to the front of a buoy by following the line. When the robot recognizes a buoy, it aborts line tracing, touches a buoy and transitions to the return behavior.

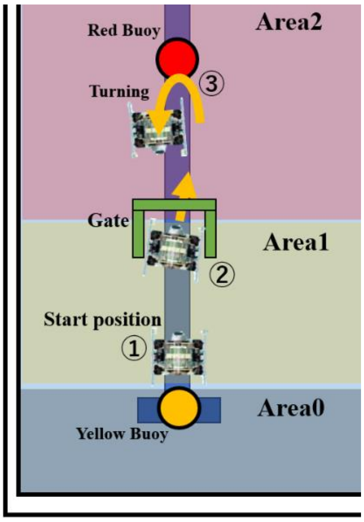


Fig. 3. Overview of competition's pool

2.4. . *Line trace method*

The robot navigates along the line using the angle information and image coordinate information of the line on the image captured by the camera for photographing underneath KYUBIC. This image processing uses the Hough transform to calculate the relative angle between the robot and the line. Figure 4 shows the flowchart to acquire angle data. In the first step, binarization is performed on the image obtained from the camera to extract the color components of the lines only, and denoising is performed by morphological transformation. Next, since the line to be traced is the boundary line between the line and the floor, edge extraction filtering is performed. For the edge image, straight line detection is performed by the Hough transform, and the angle of the line is calculated. The Yaw angle of the robot is controlled with the target angle being the sum of the line angle and the current angle of the robot. With this operation, the direction of travel of the robot becomes parallel to the line. To bring the line and KYUBIC close together, speed control is performed in the Sway direction until the difference between the line center of gravity point and the image center point becomes less than a threshold. During the line tracing, a buoy touch is performed on the front camera when a buoy is detected.

2.5. . *Buoy touch method*

As shown in Figure 5, buoy touch also performs binarization and denoising of images acquired from the

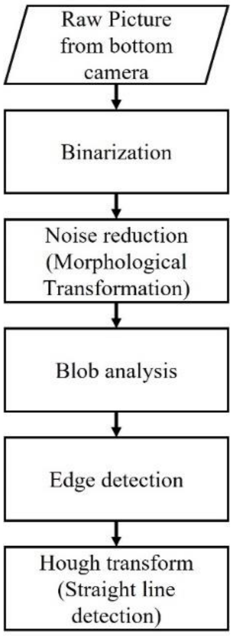


Fig. 4.
Image Processing for
Line Tracing

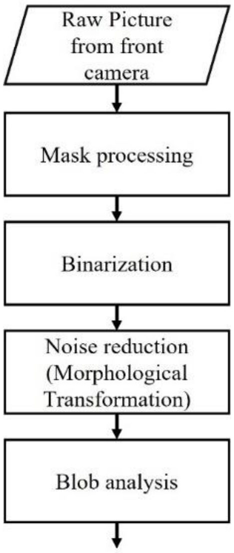


Fig. 5.
Image Processing for
Buoy touch

front camera using buoy color components. In addition to the pre-processing described above, buoy touches are masked at 10px above and 30px below the buoy and floor lines reflected on the water surface before binarization to prevent misrecognition. We label the processed binary image using connected components. Among the multiple labels, the label with the largest pixels in the region within the threshold is recognized as a buoy. It controls the speed in the Sway direction until the difference between the center of gravity of the area and the center of the image becomes less than a threshold, and then captures a buoy at the center of the KYUBIC. After that, the thrusters output thrust for a certain time, and the buoy is touched by inertial force. After the buoy touch, the robot moves backward for a certain time to avoid contact with the buoy during the round trip, and then rotates 180 degrees in the Yaw angle and returns.

3. Results of AUV Category and Consideration

The results of line tracing are shown in Figure 6. The top two graphs show the relationship between time and the difference between the region center coordinates of the detected line and the image center coordinate. The third graph shows the relationship between time and angle; between the horizontal axis of the image and the line detected by Hough transform. The fourth graph shows when the line is detected or not. A value of 1 means that the line is recognized, and a value of 0 means that the line is not recognized. Checking the Surge and Sway gap of KYUBIC, the value of the Sway gap swings. The reason for this may be that the line is not captured in the center of the image. However, Path detect was set to 1 for 140 seconds after the start of the action, indicating that the robot always detected the line and was on the line. In addition, the angle between the line and the horizontal axis of the image was recorded to be 0 degrees during that time, suggesting that the Yaw angle was kept parallel to the line on the initial position.

Figure 7 shows the motion data of KYUBIC against time. The black color indicates the target value, the blue color indicates the calculated velocity, the red color indicates the calculated thrust, and the blue and orange points in Angle indicate the Roll and Pitch. From the top, the graphs show the velocity in the Surge direction, the velocity in the Sway direction, the position in the Heave direction, the Roll angle, the Pitch angle, and the Yaw angle. The sixth graph shows whether DVL can get data or not. 1 means that DVL can get data, and 0 means that DVL is not able to get data. Focusing on the Surge velocity, Sway velocity, and DVL detection, both Surge and Sway are near the target values, but there are several periods when the DVL cannot get any data, making it unreliable. The graph shows that the thruster output is 4N during the period when the DVL is not available, because KYUBIC is set to navigate in the Surge direction with a constant thrust of 4N when the DVL is not available. The area surrounded by a green frame represents the period during which KYUBIC outputs 9N for a certain period to perform buoy touch.

Focusing on the Roll, Pitch, and Yaw angles, no significant changes were detected for Roll and Pitch and the movement of the Yaw angle was also in line with the target value. Based on these logs and the line angle information in Figure 6, we can conclude that the line trace was performed as intended during the period when DVL was available.

Figure 8 shows the results of buoy touch. From the top, each coordinate component of the detected area, the difference between the area center of gravity and the image center coordinate in the direction of the vertical axis of the image, and whether the buoy is captured at the center of the image and whether the buoy is detected. When Buoy center is 1, it means that the buoy has been detected at the center of the image, and when it is 0, it means that the buoy has not been detected at the center. When Buoy detect is 1, it means that the buoy has been detected, and when it is 0, it means that the buoy has not been detected. These graphs show that KYUBIC has recognized the buoy since about 100 seconds and has transitioned to the buoy touch.

Figure 9 shows the raw image and the binarized image during buoy touch. The markings near the center of the buoy and the two green line in the raw image are added for explanation purposes only and were not made during actual navigation. The mark near the center of the buoy corresponds to the center-of-gravity coordinate of the binarized region.

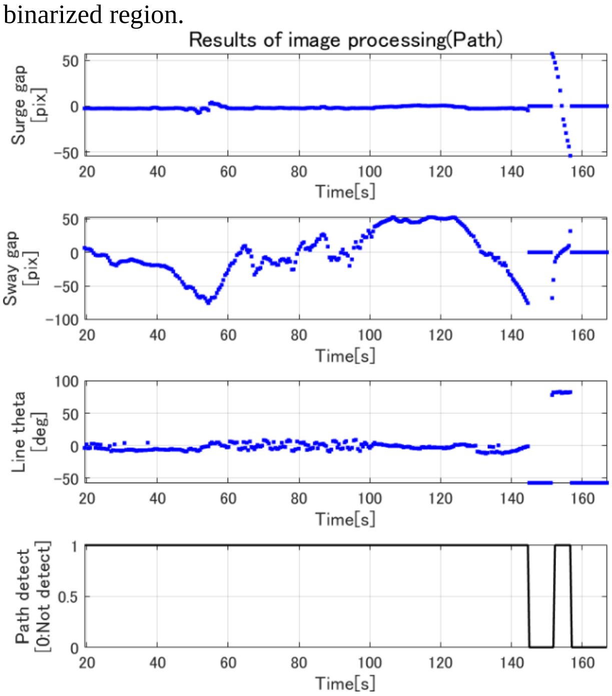


Fig. 6. Result of line trace

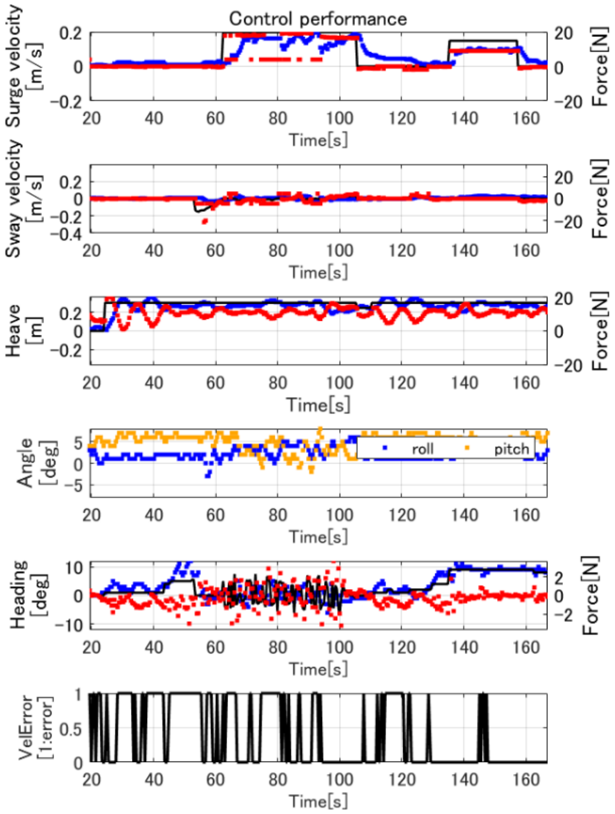


Fig. 7. Result of KYUBIC Control

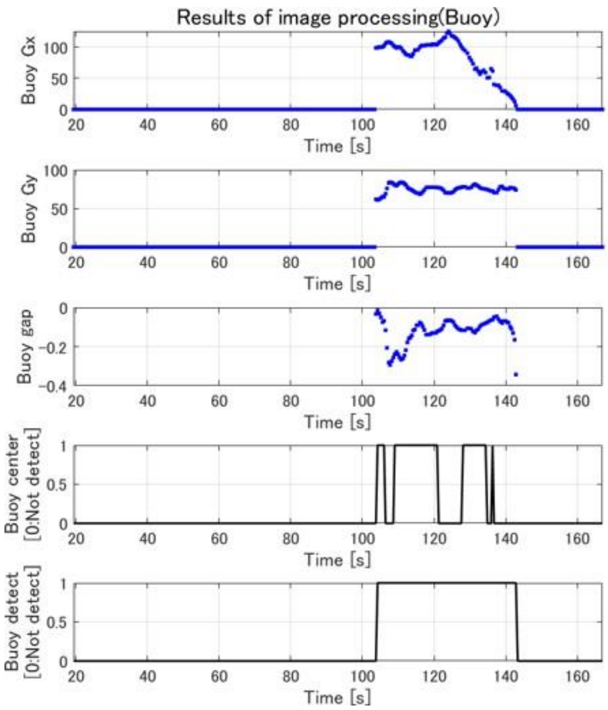


Fig. 8. Result of Buoy touch

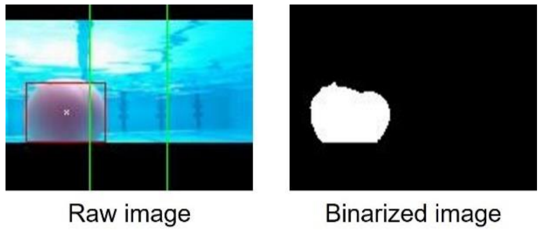


Fig. 9. Image processing results for buoy

The green line is a criterion used by KYUBIC to determine whether a buoy is centered. If the center of gravity of the detected area exists inside the two lines, the buoy is considered as centered. Figure 9 shows the image when the KYUBIC started to move forward toward the buoy. At the time of the competition, KYUBIC had once captured the buoy in the center of the screen before Figure 9 and started moving forward to the buoy, but it was unable to capture the buoy in the center of the screen in Figure 9. As a result, KYUBIC failed to touch the buoy and passed by it. The cause of these problems can be attributed to the tactics, once they started to move forward to the buoy, they did not take the tactic of canceling the buoy touch unless they lost sight of the buoy.

4. Conclusion

We implemented image processing and other processes to achieve the task of the Techno-Ocean2021 Underwater Robot Competition. As mentioned in the previous section, while we had detected the buoys, we had not been able to envision a strategy for when they were not captured in the center. As described in Chapter 2, the RGB sensor was not implemented, making the system vulnerable to changes in illumination, and in Figure 9, the camera was unable to respond to changes in illumination, resulting in blown-out areas. In the future, we will review our strategy and develop an image processing system that is robust to illumination variations.

References

1. Yuya Nishida, et al., "Robot competition for underwater technology researchers and students", Journal of Advances in Artificial Life Robotics, Vol.1, No.1, pp.11-15, 2020.
2. Toshihiro Maki, et al., "Docking Method for Hovering-Type AUVs Based on Acoustic and Optical Landmarks", Journal of Robotics and Mechatronics, Vol.30, No.1, pp.55-64, 2018.

3. Yuichi Shirasaki, "Operation of Autonomous Underwater Vehicles", NAVIGATION, Vol.158, pp.66-75, 2003.
4. Toshimune Matsumura, et al., "Development of a Handy Autonomous Underwater Vehicle "Kyubic"", Proceedings of International Conference on Artificial Life & Robotics (ICAROB2021), pp. 405-408, 2021.
5. Takashi Sonoda, et al., "System Development of AUV's Sampling Device Controller Employing MATLAB/Simulink Toolboxes", Proceedings of International Conference on Artificial Life & Robotics (ICAROB2019), pp.517-520, 2019.

Authors Introduction

Mr. Kazuki Harada

He received his Bachelar's degree from the Department of Intelligent Mechanical Engineering, Hiroshima Institute of Tehnology, Japan in 2021.
He is currently a Master's course student in Kyushu Institute of Tehnology,Japan.

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 Pre-doctoral Program, Kyushu Institute of Technology, Japan.

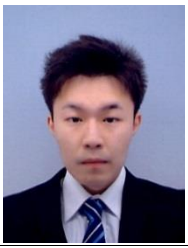
Mr.Yusuke Mizoguchi

He received his Bachelar's degree from the Undergraduate Program of Electronics, Kyoto Institute of Technology, Japan in 2021.
He is currently a Master's course student in Kyushu Institute of Tehnology, Japan.

Mr. Yusuke Yamamoto

He received his bachelor's degree from the Department of Mechanical Engineering, Faculty of Science and Engineering, Kyushu Sangyo University, Japan in 2021. He is currently Pre-doctoral Program, Kyushu Institute of Technology, Japan

Mr. Kota Mishima



He received his bachelor's degree from the Department of Mechanical engineering, Kyushu Institute of Technology, Japan in 2021. He is currently Pre-doctoral program, 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 Intelli-gence Systems, Kyushu Institute of Technology, Japan. He obtained his M. 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.

Tomato Harvesting in Greenhouse Considering the Effect of Sunlight

Kai Shioji
Dept. of Human Intelligence Systems, Kyushu institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka,808-0196, Japan
E-mail: kai.Shioji395@mail.kyutech.jp

Shinsuke Yasukawa
Dept. of Human Intelligence Systems, Kyushu institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka,808-0196, Japan
E-mail: s-yasukawa@brain.kyutech.ac.jp

Kazuo Ishii
Dept. of Human Intelligence Systems, Kyushu institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka,808-0196, Japan
E-mail: ishii@brain.kyutech.ac.jp

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. For the operation in greenhouses, the recognition system with color constancy under sunlight is necessary. In the harvesting, tomatoes should be handled gently for less damages so that the soft handling end-effector is needed. In this paper, we introduce the system configuration of the robot and the experiments conducted to solve the problem in the greenhouse.

Keywords: *Tomato-Harvesting Robot, End-effector, Color constancy, Sunlight*

1. Introduction

In agriculture, the aging and depopulation of farmers cause the shortages of farmers and manpower. The shortage of agricultural manpower has become serious problems in Japanese agriculture [1]. According to the report from the Ministry of Agriculture, Forestry and Fisheries (MAFF), the number of fulltime agricultural workers decreased by 33.6% during the past 10 years, from about 2.05 million in 2010 to about 1.36 million in 2020. The percentage of senior workers (over 65 years old) is increasing.

There are many reasons, such as cost-efficiency of the robotization, safety of the works using robots, difficulty of outdoor operations, and knowledge transfer problem from farmers to computer. As one of solutions for the problems, robot technology into the agriculture is expected to contribute to the labor-saving, improvement of production, production line automation, and also the management toward smart-agriculture.

We organize the Tomato-harvesting-robot competitions [2][3] since 2014 with the aim to foster the automation of tomato harvesting and invite young robotics researchers to agriculture. We select tomato as the target fruit as tomato is one of most important fruit vegetables. Most tomatoes are produced in the greenhouses, or large-scale farms, where the high temperature and humidity, and long harvesting force the farmers heavy works. A soft handling end-effector is needed without any damage both the surface and inside.

In addition, sunlight changes brightness and color temperature and give large influence on vision system, so color constancy is also one of key technologies [4][5].

In this paper, we introduce the tomato harvesting robot, the end-effector, and discuss the images in the greenhouse by comparing with and without the polarizing films to a camera and a lighting system.

2. Tomato harvesting robot

2.1. *Basic robot configuration and behavior*

The tomato harvesting robot is shown in Fig. 1. The robot mainly consists of mobile mechanism, a vision system, an orthogonal type manipulator with an end-effector, and a control computer (see details in [6]-[8]). The robot goes along the rails until a tomato is detected using the RGB-D camera. When a tomato is detected, the robot stops going and calculate the position of each tomato, the tomato maturity, and the harvesting order of tomatoes. Then, the manipulator approaches to the first priority target tomato and harvests it one by one with an end-effector. The obtained tomatoes move to the box through the tube attached beneath the end-effector.

2.2. *End-effector for tomato harvesting*

The harvesting is required not to give damages not only their surface but also inside. We developed a suction cutting type end-effector [8] as shown in Fig. 2. The end-effector has the cylindrical shape and consists of three modules, suction, cutting and fruit guiding modules. The suction module uses vaccume force to hold the tomato in the right position for cutting and not to give damages to the surface by preventing from grasp of the tomato. The cutting module cuts the stalk of tomato and harvests the tomato with remaining its calyx. The cutting module consists of a blade fixed to the upper edge and a U-shaped finger which is rotated by a motor to cut the stalk. Inside the module, a pair of a laser and a photoresistor is attached to check the tomato fruit is in the right position. The fruit guiding module transports the harvested tomatoes to the harvest-box and whose shape is an acrylic pipe with split lengthwise. The upper half of the pipe is fixed, and the lower half is connected to the cutting module and rotates in the pitch direction. By a torsion spring into this rotating part, the lower half pipe is closed and bridged to the hose after harvested tomato rolling

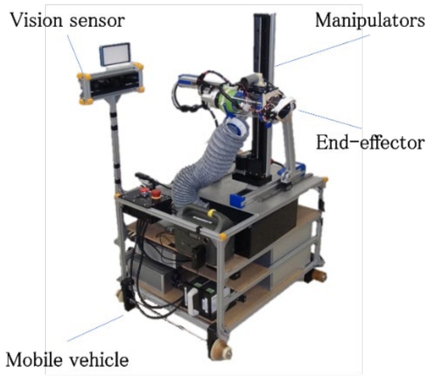


Fig. 1 The tomato harvesting robot for moving on the rail, which consists of mobile vehicle, vision system and a manipulator with an end-effector.

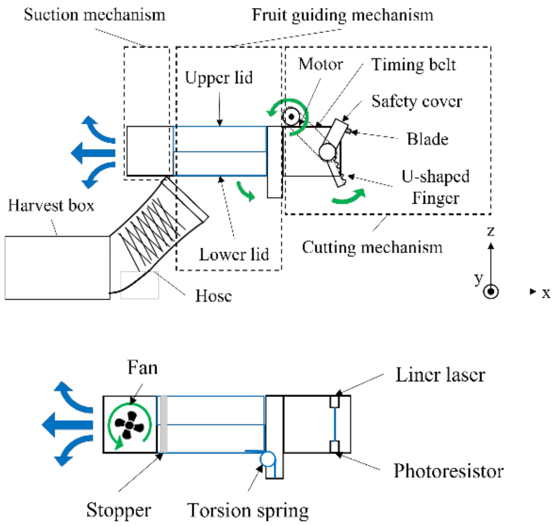


Fig. 2 The developed suction cutting type end-effector. The target tomato is absorbed using a fan, and the stem is cut.

down to the box. This hose is connected to the harvesting box, and the tomatoes are transported to the box while rolling by their own weight.

2.3. *Harvesting Process*

The flow of harvesting is shown in Fig. 3. Firstly, a target tomato is absorbed by the suction mechanism (I). When the tomato goes into the cutting mechanism, the laser beam to the photoresistor is blocked, and the photoresistor detects the object and the robot recognize the tomato is in cutting position (Ⅱ). If the grasp is successful, the U-shaped finger is rotated to cut the stalk (III). The tomato rolls over to the stopper of the fruit guiding mechanism. Then, the lower half of the pipe to open by the weight of the tomato , and the tomato goes into the hose (Ⅳ). The tomato rolls along the hose and enters to the harvest box. After the tomato dropped, the pipe is closed by the torsion spring (Ⅴ). Conventional end-effectors did not have a fruit guiding mechanism, and harvested tomatoes were grasped in the cutting mechanism and transported by manipulator to the harvest box and dropping them [8]. Even if no damage was observed on the surface of the tomatoes, still there is a possibility of the inside deformation. The proposed fruit guiding mechanism uses a flexible hose and rolls the fruit diagonally, which will reduce the impact and prevents inside damages.

2.4. *Threshold adjustment of photoresistor for tomato holding*

Whether a tomato is harvested or not is determined by the value of the photoresistor, however the photoresistor is affected by the brightness of the surrounding environment. Fig. 4 shows the results of measuring the photoresistor values indoors and outdoors. The x axis shows the time, and the y axis the sensor value at no object inside. As shown in Fig. 4(1), the photoresistor value is almost constant at around 840 in the indoor, while the value changes more than 150 between cloudy conditions (Fig. 4 (ii)) and direct sunlight (Fig. 4 (iii)) in outdoor. In this paper, we introduced the method to adjust the threshold value of the photoresistor after a constant period. Fig. 5 shows the result of setting the threshold value for the photoresistor value in Fig. 4. 10 times the value of the photoresistor is measured every 30[s] and the average is taken to get the current brightness. The threshold value is set by adding 100 to the value. The dashed line in Fig. 5 shows the threshold value, which is updated according to the value of the photoresistor.

3. Color consistency

3.1. *Polarizing films for camera and light*

In harvesting tomatoes, the proposed robot determines the maturity of tomato by the hue value of surface in the HSV color model whether the hue value is within a certain range of red, which is decided depending on the

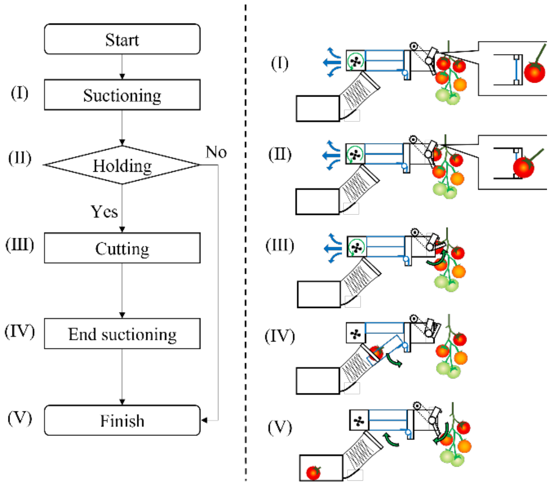


Fig.3 The process of tomato harvesting. The tomato is absorbed by air and the stem is cut when the tomato is in the right position, and roll over to the box.

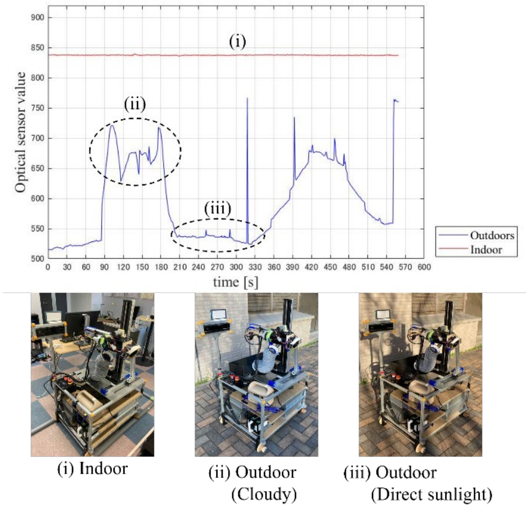


Fig. 4 The value of photoresistor to determine whether the tomato is in the right position for cutting when no object inside the cutting area exists.

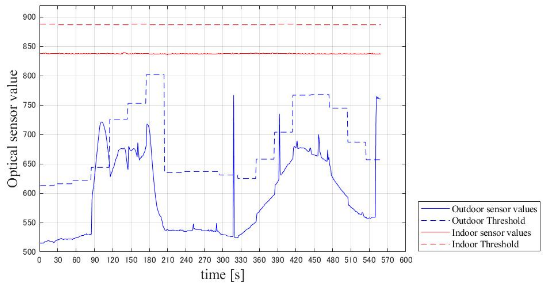


Fig. 5 The threshold is updated using previous stable sensor data to adapt the change of lighting condition.

condition of the market. In the greenhouse, due to the effects of sunlight, the color of the tomatoes appears different from their true color. If the entire image is darkened by clouds, the hue value is not calculated properly because of reduction of the color information. The vision system which can determine that “red” is red under sunlight is necessary, therefore, we conducted the evaluation experiments using a color map in our greenhouse.

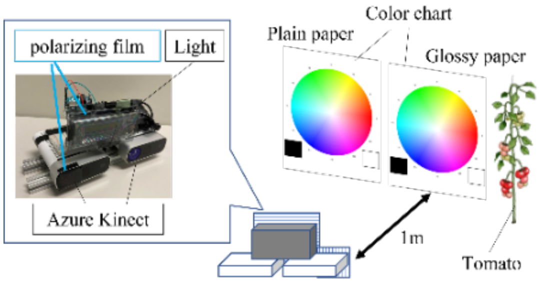
As shown in Fig. 6, a color map was set in the greenhouse and the transition of hue was recorded. As tomatoes have specular reflection on the surface with lighting on, there is a possibility that color information will be lost by white blur. To solve this problem, we used the polarizing films to attach them to one of the cameras and the light to be perpendicular to each other. Two cameras (Azure Kinect) were placed in parallel to compare the effects of polarizing film, and a light was placed above the cameras to compare the effects of light on and off with polarizing filter whose angle is perpendicular to that of camera.

3.2. *Evaluation experiments in greenhouse*

The images were taken every minute from 12:00 to 18:00. Figure 7 shows the relationship between the printed hue values and the measured hue values by comparing eight different conditions with paper type (plain or glossy), light (on or off), and polarizing film (with or without). Compared to the printed values (black line), the hue values around yellow show the differences, however, small differences in the different conditions. During the daytime, the results were similar for all time periods, and no significant changes were observed. As shown in Fig. 8, when the light was not turned on, the color information could not be observed after sunset, and it can be seen that the polarizing films reduce the specular reflection. This indicates that the lighting with the polarizing film is effective especially in dark environments such as after sunset.

4. Conclusion

In this paper, we described the problems of harvesting and recognizing tomatoes in the greenhouse. The suction cutting type end-effector can harvest tomatoes with less damaging them. Regarding color recognition, the effect of sunlight to brightness was found and the polarizing



(a) Experimental setup


(b) An example photo

Fig. 6 The experimental setup to measure the effects of sunlight. The polarizing films are attached to the one of camera and the light to be perpendicular to each other.

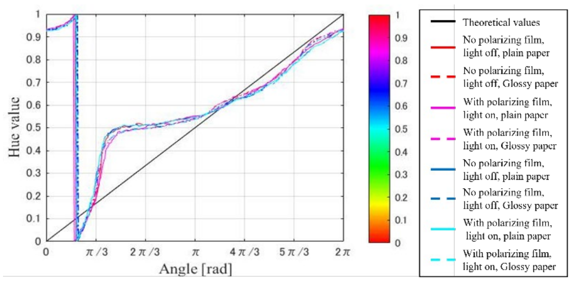


Fig. 7 Comparison of hue angle by measured values and printed values. The differences between polarizing films, lighting, paper types are compared.

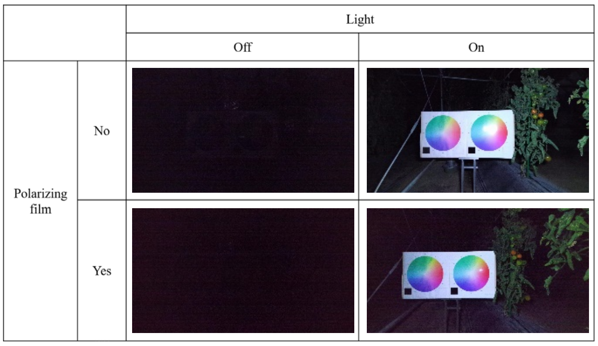


Fig. 8 The effect of polarizing films when the light is on. The reflection from surface is reduced by the films.

films will work effectively especially with the lighting devices.

of angle 90 [deg] is near the center of the aluminum foil, it is considered that it was generated by the combined wave of the ultrasonic array. The sound pressure of 0.2 [MPa] where cavitation occurs between 50 and 100 [mm] at every angle exceeded 0.2 [MPa], but in this experimental result, there is no holes or cracks.

References

1. Census of Agriculture and Forestry, Ministry of Agriculture, Forestry and Fisheries of Japan, (18th Dec., 2021).
2. T. Matsuo et. al., “Toward Smart Tomato Greenhouse: The Fourth Tomato Harvesting Robot Competition,” 6 (2), pp.138-142, 2019.
3. T. Matsuo et. al., “Toward Smart Tomato Greenhouse: The 6th Tomato-Harvesting-Robot Competition and Regulation Changes for Practical Application,” 2022, in press
4. Seishi Sekine, "Correlated Color Temperature of Daylight (1) Important Characteristics of Variation", Journal of the Illuminating Engineering Institute of Japan, Vol.79, No.2, pp.21-29, 1995 (In Japanese)
5. Keiji Uchiyama, "Color constancy and perception", The journal of the Institute of Image Information and Television Engineers, Vol.58, No.5, pp.662-668, 2004 (In Japanese)
6. Takuya Fujinaga, Shinsuke Yasukawa, Binghe Li, Kazuo Ishii, "Analysis of Characteristics of Tomato Fruits in Infrared Images Toward Automatic Tomato Harvesting System," Proc. of The 2018 International Conference on Artificial Life and Robotics, 4 pages, 2018
7. Shinsuke Yasukawa, Bingh Li, Takashi Sonoda, Kazuo Ishii, "Development of a Tomato Harvesting Robot," Proc. of the 2017 International Conference on Artificial Life and Robotics, pp.408 - 411, 2017
8. 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

Authors Introduction

Mr. Kai Shioji

He received his B.S. degree from the Kyushu Institute of Technology School of Computer Science and Systems Engineering, Japan in 2020. He is currently a Master’s course student in Kyushu Institute of Technology, Japan.

Dr. ShinsukeYasukawa

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.

Inter-University Collaboration Aimed at Integrating Different Robotic Field: Development of Underwater Robots and Soccer Robots Though these Competitions

Moeko Tominaga
Department of Integrated Systems Engineering, Nishinippon Institute of Technology,
1-11 Aratsu, Kanda-town, Miyako-gun, Fukuoka 800-0349, Japan

Jonghyun Ahn
Department of Intelligent Mechanical Engineering, Hiroshima Institute of technology,
2 Chome-1-1 Miyake, Saeki Ward, Hiroshima 731-5193, Japan

Yasunori Takemura
Department of Integrated Systems Engineering, Nishinippon Institute of Technology,
1-11 Aratsu, Kanda-town, Miyako-gun, Fukuoka 800-0349, Japan

Kazuo Ishii
Graduate School of Life and Science Systems Engineering, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka 808-0196, Japan
E-mail: tominaga@nishitech.ac.jp, j.ahn.h2@it-hiroshima.ac.jp, takemura@nishitech.ac.jp, ishii@brain.kyutech.ac.jp
www.nishitech.ac.jp, www.it-hiroshima.ac.jp, www.kyutech.ac.jp

Abstract

In robotics, the problems and solutions to be focused on may differ due to the different fields of robots to be developed. In this research, we will verify the effect of exchanging opinions and sharing knowledge by collaborating with students aiming to participate in different robot competitions between underwater robots and soccer robots. As a result of analysis using neural networks, it was found that collaborative research between universities contributes to maintaining student motivation.

Keywords: Inter-University Collaborative Education, Integrating Different Robotic Field, Robot Competition, Manufacturing Education, Under Water Robot, Soccer Robot, Neural Network

1. Introduction

Participating in robotics competitions is a very effective way to educate students about robotics[1]. This is because the results obtained by participating in the robot contest include not only technology and knowledge about manufacturing, but also collaboration with colleagues, team management, communication skills, and presentation skills. In many engineering laboratories, education is provided by participating in robot contests in each laboratory[2]. However, in many cases, participating students cannot afford to participate even if they are interested in competitions other than the robot contest they are in charge of. Not only does this narrow the horizons of students, but it also eliminates the opportunity to learn how to apply knowledge.

In this research, we will give students a chance to participate in multiple robot contests and verify the effect by promoting education in collaboration in the laboratories of three universities participating in different robot contests. The activities and results of each robot contest are shown as objective facts. In addition, as educational results for students, the results of analyzing the answers to the student questionnaires obtained at each regular meeting using the neural network are shown. The purpose of this project is to confirm the usefulness of education using multiple robot contests.

2. Educational Policy

In this project, we target Multiple types students not only 4th graders students assigned to the laboratory as shown in Table 1. All students are voluntary participants and there is no coercion to all activities, including regular meetings.

Table 1. Breakdown of students who participated in the inter-university collaborative education project.

		University			
		NIT	HIT	Kyutech	Total
Grade	Bachelor 1	1	-	-	1
	Bachelor 2	-	-	-	-
	Bachelor 3	-	3	-	3
	Bachelor 4	2	2	-	4
	Master	-	-	1	1
	Total	**4**	**5**	**1**	**10**

These students are learning basic engineering knowledge in class, but often do not know how to apply it to manufacturing. In this paper, we will create and educate relatively large-scale system robots with reference to efforts with high educational effects, but the work that each student is in charge of is optimized.

2.1. *Target robot contest and schedule*

This project was planned by faculty members from three universities, Nishinippon Institute of Technology, Hiroshima Institute of Technology, and Kyushu Institute of Technology in April, and has been implemented since May. The target robot contest is for underwater robots and soccer robots that have participated in each university so far. Table 2 shows the schedule of the entire project and the schedule of the contests.

At regular weekly meetings, the faculty members confirmed the progress of each student, pointed out presentation practice, and provided knowledge in the field of robotics. In addition, the management of each student's tasks and the scheduling up to the contest were conducted mainly by the students.

3. Activity and results

This year, two contests were held for both soccer robots and underwater robots. Due to the influence of Covid-19, the method of participating in all contests has become online.

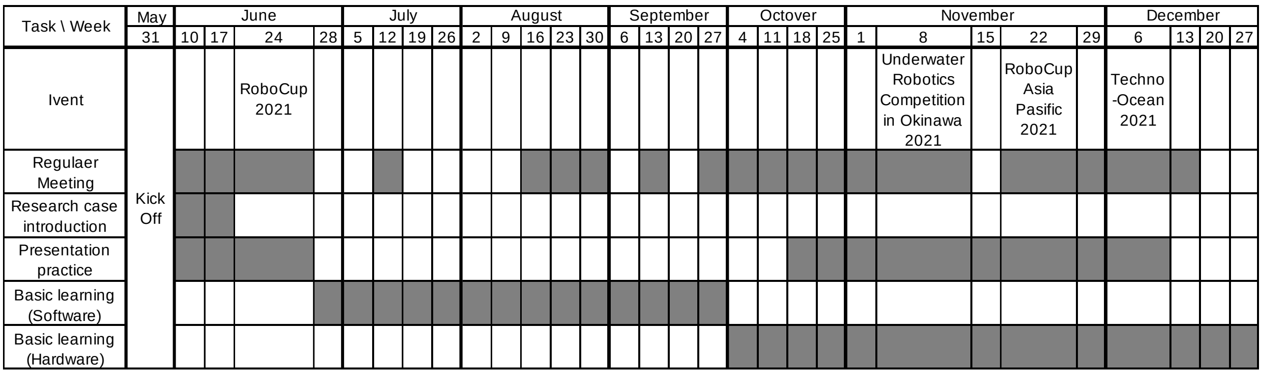
3.1. *Under water robots*

The following two were held in the underwater robot contest in 2021.

- Underwater Robotics Competition in Okinawa
- Techno-Ocean

These contests are divided into Remotely operated vehicle (ROV), Autonomous Underwater Vehicle (AUV), and Free style according to how the robot is operated. The 2 robots involved in this project participated in the freestyle category. In the freestyle category, in addition to the evaluation of the presentation,

Table 2. The planning of the inter-university collaborative education project.



the evaluation of practicality, technical ability, originality, and perfection was performed by video screening. In this project, we verified the calibration technology in sensing and the strength of the robot by the material. The robots that participated in the contest are shown in Figures 1 and 2.

J. E. N. O. S. shown in Fig. 1 is a robot whose purpose is to exterminate a large number of jellyfish[3]. The project mainly improved the strength by changing the sensing technology and parts manufacturing method and materials. As a result of the contest, the robot was not able to win the Underwater Robotics Competition in Okinawa 2021, but was able to win the Freestyle category at Techno-Ocean 2021.

M. I. R. O. C. A. shown in Fig. 2 is a robot aimed at collecting water waste, especially PET bottles. As a project, we worked on improving the Resolution and Collection Corporation. As a result of the contest, the robot finished second in the freestyle category in both Underwater Robotics Competition in Okinawa 2021 and the Techno-Ocean 2021.

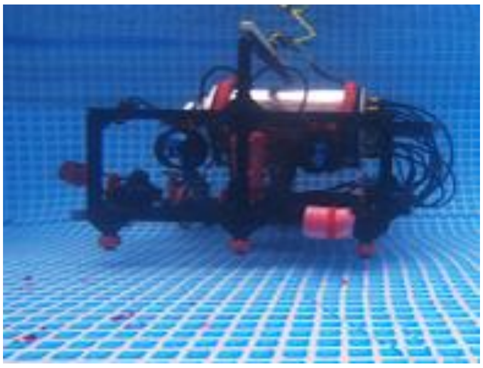


Fig. 1. Underwater robot for the purpose of exterminating jellyfish “J. E. N. O. S.”

Fig. 2. Water robot for the purpose of collecting water waste “M. I. R. O. C. A.”

3.2. *Soccer robots*

The following two were held in the soccer robot contest.

- RoboCup 2021
- RoboCup Asia Pacific 2021

RoboCup is a global landmark project, with competitions divided into various leagues. The original mission is to make a team of robots capable of winning against the human soccer World Cup champions by 2050. In this project, the students participated as "Hibikino-Musashi", a team participating in the medium-sized league. The robots that participated in the contest are shown in Figures 3.

In both contests, the presentation of the research (Scientific Challenge) and the introduction of the robot by video (Technical Challenge) were evaluated instead of the ranking by the soccer competition. Since both contests are international contests, the project worked to improve presentations in English and listen to presentations from other teams in the contest. We also shared the basics of sensing with a microcomputer and the basic knowledge of action decisions using AI[4]. The students worked on improving the omni wheel, which is the drive wheel, and the robot cover.

As a result of the contest, Hibikino-Musashi could not win the prize at the world competition RoboCup 2021, but at the RoboCup Asia Pacific 2021 where the students took advantage of their reflection, Hibikino-Musashi won both challenges.

Fig. 3. The soccer robot aiming to play soccer with humans “Musashi”

4. Evaluation

As shown in the feature map of Fig. 4, a tensor self-organizing map was used to evaluate changes in student motivation over the season.

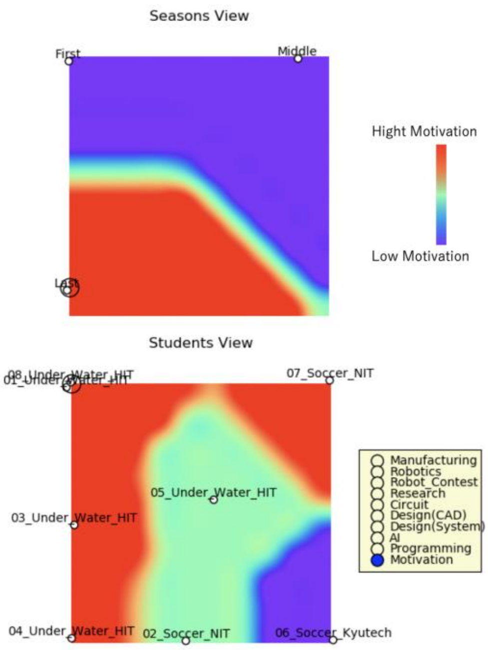


Fig. 4. Evaluation of motivation by TensorSOM

5. Conclusion

In this study, we examined whether collaboration between universities had a positive effect on student motivation, and confirmed that motivation improved as the season passed. In addition, this project was able to achieve good results in the contest as a result of cooperation between universities.

References

1. K. Jitjumnong, P. Chujai, N. Kamata, “Robot Contest for Innovative Development in Education Technology”, International Journal of Mechanical Engineering and Robotics Research, Vol. 9, No. 3, pp. 395-400, 2020.
2. J. Kawata, J. Morimoto, M. Higuchi, S. Fujisawa, “The Educational Effected of Practical Manufacturing Activities in Graduation Research”, Journal of Robotics and Mechatronics, Vol.31 No. 3, pp. 391-404, 2019.
3. J. Ahn, S. Chikushi, S. Yasukawa, T. Sonoda, “Development of Autonomous Underwater Vehicle for Small-size Jellyfish Extermination and its Evaluation”, Journal of Japan Society for Design Engineering, pp. 1-10, 2021.
4. M. Tominaga, Y. Takemura, K. Ishii, "Behavior Selection System for Human-Robot Cooperation Using Tensor SOM", Journal of Robotics, Networking and Artificial Life, Vol .7(2), pp.81-85, 2020.
5. T. Iwasaki, T. Furukawa, “Tensor SOM and tensor GTM: Nonlinear tensor analysis by topographic mappings”, Neural Networks, Volume 77, pp. 107-125, 2016.

Authors Introduction

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. 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.

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.

Prof. Kazuo Ishii

He received Ph.D. 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 Intelligences Systems, and the Assistant of 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.

Exercise on Environmental Monitoring and Control of Green house by IoT Devices toward Smart Agriculture

Yuya Nishida[*1]
Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196, Japan

Ryugo Mchizuki[*2]
Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196, Japan

Shinsuke Yasukawa[*3]
Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196, Japan

Kazuo Ishii[*4]
Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196
*E-mail: y-nishida[*1], s-yasukawa[*3], ishii[*4]@brain.kyutech.ac.jp, mochizuki[*2]@lsse.kyutech.ac.jp*
https://www.kyutech.ac.jp

Abstract

As crops in greenhouses are widely distributed, IoT devices placed near the crops should be stand-alone and modular, and data from the devices are collected over the networks. Smart agriculture requires knowledge of a wide range of fields including electricity, information, and image processing. We have designed an AI and IoT technology exercise on environmental monitoring and control of a greenhouse where we have been preparing for grow up of tomatoes and other vegetables.

Keywords: Smart agriculture, green house, mechatronics exercise, MATLAB/Simulink exerciese

1. Introduction

In the future estimation of 2050, the food demand increases 70% and the production decreases 15% caused by global warming, and farmer population be 1/5 of current workers. We need an agricultural system of twice production with the same farmland area and 5 times effective operations until 2050 [1]. For the sustainable society, smart agriculture including robot technology, AI, IoT is one of the solutions for food issues. As crops in greenhouses are widely distributed, IoT devices placed near the crops should be stand-alone and modular, and data from the devices are collected over the networks. Camera is often used for monitoring of growth status of the crops, and only resulted information by image processing should be transmitted [2-3]. Smart agriculture requires knowledge of a wide range of fields including electricity, information, and image processing [4-5].

Authors have designed an AI and IoT technology exercise on environmental monitoring and control of a greenhouse where we have been preparing for grow up of tomatoes and other vegetables. The exercise that was

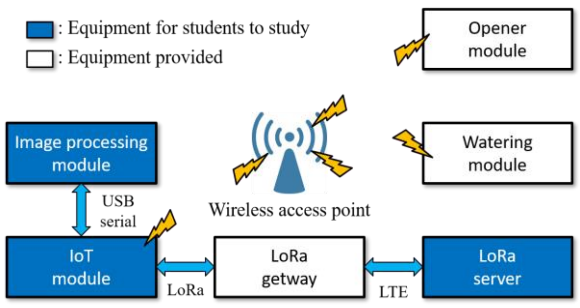


Fig.1 System architecture for acoustic communication

a part of the joint graduate school consists of mechatronics exercise using micro-computer, MATLAB/Simulink exercise for image processing, and system construction exercise for smart agriculture. This paper explains detail of the exercise, and shows system constructed by the students and results of rubric.

2. Exercise outline

2.1. *The exercise in the joint graduate school*

Graduate schools of engineering of Kyushu Institute of Technology, the University of Kitakyushu, and Waseda University established joint graduate school car electronics course supported by strategic university collaboration support project of ministry of education, culture, sports, science and technology in 2008. Then, the graduate schools established intelligent car robotics course based on the car electronics course, with Kyushu Institute of Technology as the representative school. Car electronics and robotics intelligences is taught in the course, supported by program for promoting inter-university collaborative education. To develop highly professional human resources that become a leader in the next generation of intelligent robotics and AI technology, joint graduate school intelligent car robotics & AI that AI sub course for acquiring the basic technology of AI was added has been started since 2019. Authors offered Exercise on Environmental Monitoring and Control of Green house by IoT Devices as comprehensive practicum subjects in the joint graduate school in 2021. The exercise has mechatronics exercise, MATLAB/Simulink exercise and system construction exercise toward smart agriculture, using the system shown in Fig.1. Students learn how to use each module and how to communicate with each module.

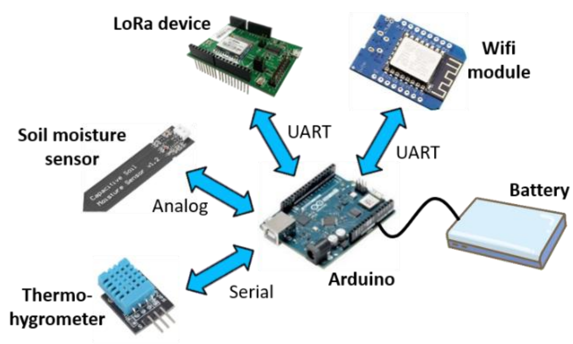


Fig. 2 IoT module using an Arduino

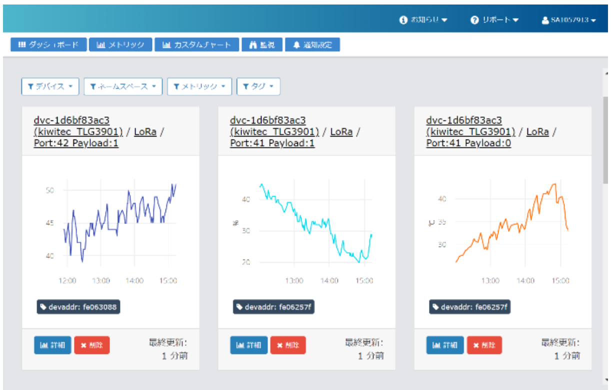


Fig.3 Graphs on the LoRa server

2.2. *Mechatronics exercise*

The mechatronics exercise is separated three sections, and IoT module using an Arduino as shown Fig.2 is used for the exercise. In first section, students make simple peripheral circuits using Arduino and program for that circuits, and learn Arduino functions of digital I/O and A/D conversion. The first section is easy for students who have learned Arduino, but some students are new to Arduino, so they learn about basic electrical circuits and how to use Arduino. DC motor driver using n-channel enhancement type FET and motor drive by PWM signal from Arduino are taught in second section. The driver has a tachometer for measurement angular velocity of the motor, and low pass filter is located between signal pin of the tachometer and A/D pin of Arduino. Students implement PID controller using PWM in the driver with Arduino to learn feedback control for the motor. In third section, LoRaWAN communication that is the part of LPWA (**L**ow **P**ower **W**ide **A**rea) is taught using IIJ

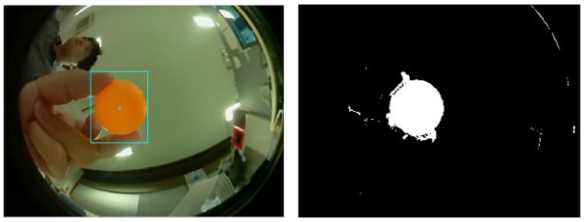
(Internet initiative Japan) LoRa device and LoRa gateway. The device transmits data received from Arduino to the gateway up to 10km away by LoRa of class C that transmits at any time unlike other classes, and the gateway uploads the data by LTE communication to LoRa server. Arduino program for communication to the device and LoRa script on the gateway for data upload are provided by teachers. Students make Arduino program that data measured by soli moisture sensor and thermo-hygrometer are sent to the device, and set the server to display a graph of the data as shown in Fig.3.

2.3. *MATLAB/Simulink exercise*

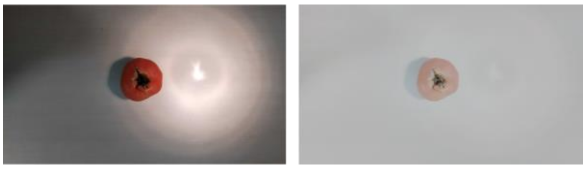
The MATLAB/Simulink exercise is separated two sections, and image processing is learned by using Raspberry Pi with a dedicated camera. By installing support package for Raspberry Pi on the MATLAB/Simulink, it is possible to develop and debug the program for Raspberry Pi on the MATLAB/Simulink. In first section, basic image processing is taught to learn method required to monitor the condition of crops. Students learn color object detection, noise filter and edge extraction, and implement the ball detection program in Raspberry Pi as shown Fig.4. Image enhancement which is the part of pre-processing method is taught in second section. The appearance of craps in photo-image changes greatly with time and weather because the intensity of sun light depends on its position and the percentage of clouds. Image enhancement is often used before the object detection because simple image processing may not detect the crops in outdoor. Students learn single scale retinex filter how human perceive color and light is modeled in the section. The retinex filter that is high-pass filter removes illumination component from the image as shown in Fig.5. Crop detection becomes robust against sun light by incorporating pre-processing by retinex filter.

2.4. *System construction exercise*

Students who have finished mechatronics and MATLAB/Simulink exercises are divided into groups to work system construction in system construction exercise. Each group develop IoT system for agriculture by utilizing the technology learned in the exercises and explain its development purpose, concept, functions, performance to teachers and audience. Students can use



(a) Detection results (b) Binarized image
Fig.4 Orange ball tracking by image processing



(a) Original image (b) Enhanced image
Fig.5 image enhancement using retinex filter

Fig.6 Students who measure soil moisture

watering and opener modules shown as Fig.1, and Arduino can control moisture of the ridge by using watering module and house temperature by using opener module. Teachers who hear group's presentation evaluate by usefulness, perfection, operability, quality of presentation and quality of demonstration.

3. Results of the exercise

Authors gave the exercise that was taken seven students in the first and second year of the master's program on September 2021. Mechatronics and MATLAB/Simulink exercise were conducted in a hybrid of face-to-face and online by using Zoom. Almost students had no experience using Arduino and MATLAB/Simulink, and making electronic circuits. Because actual equipment is used for the exercise, teachers was difficult to completely support for the students who participated remotely. However, all students have completed all the exercises.

Fig.7 Students who work near outdoor culture

Seven students who have completed the two exercises were divided into team A and team B for the system construction exercise. At this time, the opener module could not be used because it failed. Team A worked on the development of a system used in the green house. Figure 6 shows students who measures soi moisture of ridges in the house. Team A developed soil moisture monitoring system for cucumbers using a camera and sensors. The system can upload the data measured by image processing and soil moisture sensor every few minutes, and users can monitor its data on the LaRa Sever. Team B worked on the development of a system for outdoor culture as shown in Fig.7, and developed the system measures degree of ripeness of green pepper. The system judges the degree of ripeness based on the color information in HSV color model, and transmits its results on the LaRa server. Although the systems of two teams were not perfect, students understood and utilized the techniques that they learn in the exercise to develop the system. Rubric results before and after our exercise as shown in Fig.8 shows that almost students acquired knowledge of micro-computers, image processing and MATLAB/Simulink by our exercise.

4. Conclusions

We have designed an AI and IoT technology exercise on environmental monitoring and control of a greenhouse where we have been preparing for grow up of tomatoes and other vegetables. The exercise consists of mechatronics, MATLAB/Simulink and system construction exercises, technology required for smart agriculture can be learned in the exercise. Almost students who took the exercise acquired technology related to micro-computers, image processing technology and control by MATLAB/Simulink, but a few

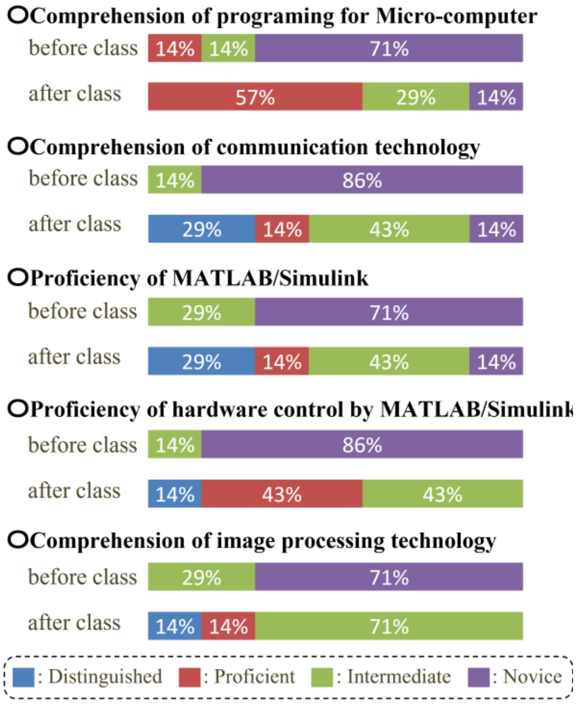


Fig.8 Rubric results before and after our exercise

students couldn't acquire them. In the next school year, authors plan to teach from a little more basic technology.

References

1. Lesline Lipper, et al., Climate-smart agriculture for food security, Nature Climate Change, Vol.4, pp.1068-1072, 2014
2. Takuya Fujinaga, et al., Image Mosaicing Using Multi-Modal Images for Generation of Tomato Growth State Map, Journal of Robotics and Mechatronics, Vol.30, No.2, pp.187-197, 2018
3. Takayuki Matsuo, et al., Tomato-harvesting Robot Competition: Aims and Developed Robot of 6th Competitions, Proc. of International Conference on Artificial Life and Robotics, 2021
4. Kazuo Ishii, et al., Tomato-harvesting-robot competition towards smart agriculture, Proc. of International Conference on Artificial Life and Robotics, Hiroshima, 2021
5. Kazuo Ishii, et al., A greenhouse project toward smart agriculture, Proc. of International Conference on Artificial Life and Robotics, Hiroshima, 2021

Authors Introduction

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. 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.

Dr. Shinsuke Yasukawa

He is an Associate Professor at Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology. His research interests include information processing in biological systems and their applications in robotics.

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

Evaluation of roller arrangement of sphere by omnidirectional integral value

Kenji Kimura
Engineering course and International Baccalaureate course Fukuoka Daiichi High School
22-1 Tamagawamati, Minamiku, 815-0037, Fukuoka, Japan

Yusuke Abematsu
Kagoshima Gyokuryu High School
20-57 Ikenouecho, 892-0806, Kagosima, Japan

Hiroyasu Hirai
Graduate School of Life Science and engineering, Kyusyu Institute of Technology
2-4 Hibikino,Wakamatsu-ku, Kitakyushu 808-0196, Fukuoka, 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:kimuken.19772058@gmail.com, abematu0081008@gmail.com, multipletitan@gmail.com, ishii@brain.kyutech.ac.jp

Abstract

Roller arrangement problem is a sphere conveyance problem of driving rollers. In this research, the roller arrangement problem, viewed as an evaluation function, is thought of as mean of roller kinetic energy with respect to the sphere direction. Furthermore, theoretically, we derive the function, and find the contact point such that the evaluated value is minimal.

Keywords: Omnidirectional movement, Angular velocity vector of the sphere, Kinetic energy of the sphere

1. Introduction

A sphere, one of the basic shapes of robots, is used not only as a multi-fingered 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, and there are various arrangements, depending on the application of the movement mechanism.

Figure 1 shows the roller contact type for the number of actuators(N_w) per sphere.

Examples of mechanisms driven by two rollers include a power transmission mechanism by Wada et al. [1] (see **Figure 1**(a)), a mobile device using Miyamoto's ball [2] (see **Figure 1**(b)), and. The abovementioned mechanisms can be used for the roller of a wheelchair. In the roller arrangement, a roller fixed in advance is performed on the equator parallel to the horizontal plane, and the sphere can be rotated in two degrees of freedom by generating an angular velocity vector on the plane by Kimura [3].

The ball holding mechanism [4] (see **Figure 1**(c)) is intended to transport the ball, and the roller is placed in the upper hemisphere to hold the ball by friction.

Here, roller arrangement problem is considered a sphere conveyance problem by driving rollers.

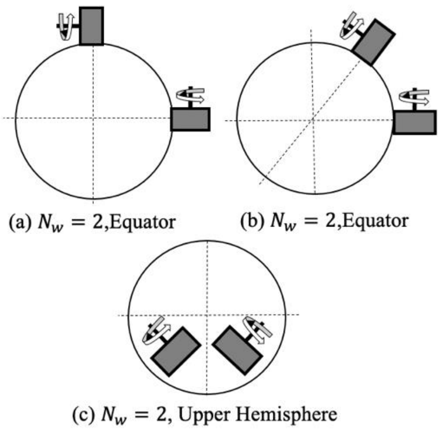


Figure 1 Type of roller arrangement for sphere mobile robot

In this research, in the case of omnidirectional movement, we define an evaluation function as mean of roller's kinetic energy with respect to sphere direction angle, and we also derive the exact formula. Furthermore, theoretically, we find the contact point such that the evaluated value (mean of roller's kinetic energy) is minimal. Additionally, we perform simulation and present energy distribution of several contact points on a sphere.

2. Derivation of theoretical evaluation function

In this chapter, we calculate the omnidirectional energy integral of the driving rollers.

As shown in **Figure 2**, 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 (i = 1 or 2) is in point contact with the sphere at a position vector $\boldsymbol{P_i}$ ($\boldsymbol{P_1} \neq \boldsymbol{P_2}$). $\boldsymbol{\omega}$ denotes the angular velocity vector of the sphere. Because of $\boldsymbol{\eta_1}, \boldsymbol{\eta_2} \in \mathbf{span}\{\boldsymbol{P_1}, \boldsymbol{P_2}\}$ (omnidirectional condition), $\boldsymbol{\omega}$ is on span$\{\boldsymbol{P_1}, \boldsymbol{P_2}\}$. sphere direction φ ($0° \leq \varphi < 360°$) is the angle from x-axis and ρ is the angle from xy-plane to $\boldsymbol{\omega}$. Now, given the sphere mobile velocity $\boldsymbol{V}$ (the center velocity of sphere).

2.1 Kinetic energy of the roller

Consider two rollers (right cylinder) with radius R, mass M, moment of inertia I, and roller's angular velocity ω_i. The total kinetic energy of the rollers is given by Eq. (1).

$$E = I(\omega_1^2 + \omega_2^2) = \frac{M}{2}(\|\boldsymbol{\omega} \times \boldsymbol{P}_1\|^2 + \|\boldsymbol{\omega} \times \boldsymbol{P}_2\|^2) \tag{1}$$

2.2 Mean of kinetic energy of rollers

To evaluate the value for roller arrangement, we define the follows expressions. Eq. (2) presents the mean of kinetic energy by integrating the total kinetic energy of the rollers with respect to the direction φ ($0° \leq \varphi \leq 360°$).

$$E_M = \frac{1}{2\pi}\int_0^{2\pi} E\, d\varphi \tag{2}$$

(i) Case of arbitrary arrangement

Quoting Equation (12) of Paper [5] (Kimura) as follows:

$$\begin{aligned}\|\boldsymbol{\omega} \times \boldsymbol{P}_1\|^2 + \|\boldsymbol{\omega} \times \boldsymbol{P}_2\|^2 &= (\|\boldsymbol{e_3} \times \boldsymbol{P}_1\|^2 + \|\boldsymbol{e_3} \times \boldsymbol{P}_2\|^2)\omega_z^2 \\ &+ 2(\langle \acute{\boldsymbol{\omega}} \times \boldsymbol{P}_1, \boldsymbol{e_3} \times \boldsymbol{P}_1\rangle + \langle \acute{\boldsymbol{\omega}} \times \boldsymbol{P}_2, \boldsymbol{e_3} \times \boldsymbol{P}_2\rangle)\omega_z \\ &+ \|\acute{\boldsymbol{\omega}} \times \boldsymbol{P}_1\|^2 + \|\acute{\boldsymbol{\omega}} \times \boldsymbol{P}_2\|^2\end{aligned} \tag{3}$$

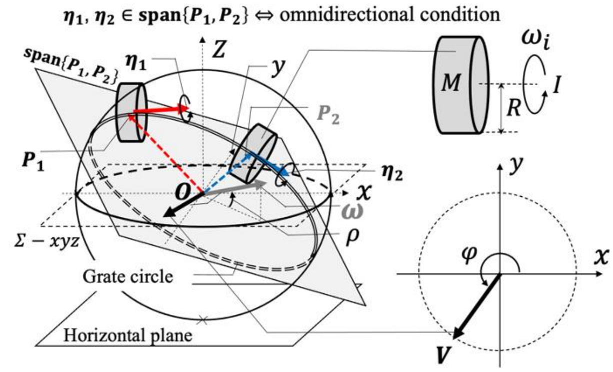


Figure 2 The sphere rotational motion by driving rollers at $\boldsymbol{P_i}$.

where

$$\boldsymbol{P_i} = r[\cos\theta_{i,1} \cos\theta_{i,2}, \sin\theta_{i,1} \cos\theta_{i,2}, \sin\theta_{i,2}]^T \tag{4}$$

$$\boldsymbol{e_3} = [0, 0, 1]^T,\ \acute{\boldsymbol{\omega}} = [\omega_x, \omega_y, 0]^T,\ \omega_z = \|\boldsymbol{V}\| \tan\rho / r$$

Using $\boldsymbol{P}_i = [P_{i,x}, P_{i,y}, P_{i,z}]^T$, $\boldsymbol{e_3} \times \boldsymbol{P}_i$, $\acute{\boldsymbol{\omega}} \times \boldsymbol{P_1}$ are represented as follow.

$$\boldsymbol{e_3} \times \boldsymbol{P}_i = [-P_{i,y}, P_{i,x}, 0]^T \tag{5}$$

$$\acute{\boldsymbol{\omega}} \times \boldsymbol{P}_i = [\omega_y P_{i,z}, \omega_x P_{i,z}, \omega_x P_{i,y} - \omega_y P_{i,x}]^T \tag{6}$$

Using Eqs. (5), $\|\boldsymbol{e_3} \times \boldsymbol{P}_i\|^2$ is calculated in teams of $\boldsymbol{P}_i = [P_{i,x}, P_{i,y}, P_{i,z}]^T$.

$$\|\boldsymbol{e_3} \times \boldsymbol{P}_1\|^2 + \|\boldsymbol{e_3} \times \boldsymbol{P}_2\|^2 = P_{1,x}^2 + P_{1,y}^2 + P_{2,x}^2 + P_{2,y}^2 = 2r^2 - P_{1,z}^2 - P_{2,z}^2 \tag{7}$$

Using Eqs. (5) and Eqs. (6),

$$\begin{aligned}&\langle \acute{\boldsymbol{\omega}} \times \boldsymbol{P}_1, \boldsymbol{e_3} \times \boldsymbol{P}_1\rangle + \langle \acute{\boldsymbol{\omega}} \times \boldsymbol{P}_2, \boldsymbol{e_3} \times \boldsymbol{P}_2\rangle \\ &= -\frac{\|V\|}{r}\{(P_{1,x}P_{1,y} + P_{2,x}P_{2,z})\sin\varphi \\ &\qquad + (P_{1,y}P_{1,z} + P_{2,y}P_{2,z})\cos\varphi\}\end{aligned} \tag{8}$$

Using Eqs. (6),

$$\begin{aligned}\|\acute{\boldsymbol{\omega}} \times \boldsymbol{P}_1\|^2 + \|\acute{\boldsymbol{\omega}} \times \boldsymbol{P}_2\|^2 &= \frac{\|V\|^2}{r^2}\{P_{1,z}^2 + P_{2,z}^2 \\ &+ (P_{1,y}^2 + P_{2,y}^2)\sin^2\varphi + (P_{1,x}^2 + P_{2,x}^2)\cos^2\varphi \\ &+ 2(P_{1,x}P_{1,y} + P_{2,x}P_{2,y})\sin\varphi\cos\varphi\}\end{aligned} \tag{9}$$

Thus. By substituting Eqs. (7), Eqs. (8) and Eqs. (9) for Eqs. (3), It can be represented in teams of $P_{i,x}$, $P_{i,y}$, $P_{i,z}$.

Here, using expressions,

$$\frac{1}{2\pi}\int_0^{2\pi}\sin^2\varphi\, d\varphi = \int_0^{2\pi}\cos^2\varphi\, d\varphi = \frac{1}{2} \tag{10}$$
$$\frac{1}{2\pi}\int_0^{2\pi}\sin\varphi\cos\varphi\, d\varphi = 0$$

Integral of Eqs. (8) and Eqs. (9) by $\varphi(0° \le \varphi \le 360°)$ will be calculated as follow.

$$\int_0^{2\pi}(\langle\acute{\boldsymbol{\omega}}\times\boldsymbol{P}_1, \boldsymbol{e_3}\times\boldsymbol{P}_1\rangle + \langle\acute{\boldsymbol{\omega}}\times\boldsymbol{P}_2, \boldsymbol{e_3}\times\boldsymbol{P}_2\rangle)\omega_z d\varphi \tag{11}$$
$$= \pi\frac{\|V\|^2}{r^2}\frac{1}{P_{1,x}P_{2,y}-P_{1,y}P_{2,x}}\{ (P_{1,x}P_{1,z}+P_{2,x}P_{2,z})(P_{1,y}P_{2,z}-P_{2,y}P_{1,z}) + (P_{1,y}P_{1,z}+P_{2,y}P_{2,z})(P_{1,z}P_{2,x}-P_{1,x}P_{2,z})\}$$

$$\int_0^{2\pi}\|\acute{\boldsymbol{\omega}}\times\boldsymbol{P}_1\|^2 + \|\acute{\boldsymbol{\omega}}\times\boldsymbol{P}_2\|^2\, d\varphi \tag{12}$$
$$= \pi\frac{\|V\|^2}{r^2}(2P_{1,z}^2 + 2P_{2,z}^2 + P_{1,x}^2 + P_{1,y}^2 + P_{2,x}^2 + P_{2,y}^2)$$
$$= \pi\frac{\|V\|^2}{r^2}(2r^2 + P_{1,z}^2 + P_{2,z}^2)$$

By substituting Eqs. (3),(7), (11) and (12) into Eq. (1), E_M can be represented as

$$\frac{4r^2}{M\|V\|^2}E_M = \frac{2r^2 - P_{1,z}^2 - P_{2,z}^2}{(P_{1,x}P_{2,y}-P_{1,y}P_{2,x})^2}\{(P_{1,y}P_{2,z}-P_{2,y}P_{1,z})^2 + (P_{1,z}P_{2,x}-P_{1,x}P_{2,z})^2\} + \frac{2}{P_{1,x}P_{2,y}-P_{1,y}P_{2,x}}\{ (P_{1,x}P_{1,z}+P_{2,x}P_{2,z})(P_{1,y}P_{2,z}-P_{2,y}P_{1,z}) + (P_{1,y}P_{1,z} + P_{2,y}P_{2,z})(P_{1,z}P_{2,x}-P_{1,x}P_{2,z})\} + 2r^2 + P_{1,z}^2 + P_{2,z}^2 \tag{13}$$

By theoretical calculation, we get the following properties.

[Property 1]: Optimality of the evaluated value

If $(\theta_{1,2}, \theta_{2,2}) = (0,0)$ ($\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ are on the equator), E_M takes the minimal value $M\|V\|^2/2$.

Method of proof is follow. Using inequality

$$(x^2+y^2)(2r^2-x^2-y^2) \ge (px+\alpha y)^2 + (qx+\beta y)^2 \tag{14}$$

, AM-GM inequality and Cauchy-Schwarz inequality, [**Property1**] is proved.

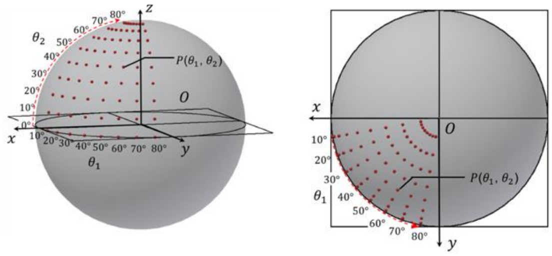


Figure 3 The distribution of contact points on the upper hemisphere. (a) Isometric view. (b) Right overhead view.

Table 1 The distribution of energy function $E_M(\theta_1, \theta_2)$ in the upper hemisphere.

80°	32.19	29.40	25.12	19.87	14.29	9.04	4.76	1.97
70°	8.32	7.67	6.66	5.43	4.12	2.89	1.88	1.23
60°	3.91	3.65	3.25	2.76	2.24	1.75	1.35	1.09
50°	2.38	2.25	2.07	1.83	1.59	1.36	1.17	1.04
40°	1.68	1.62	1.53	1.41	1.29	1.18	1.08	1.02
30°	1.32	1.29	1.25	1.20	1.14	1.08	1.04	1.01
20°	1.13	1.12	1.10	1.08	1.05	1.03	1.02	1.00
10°	1.03	1.03	1.02	1.02	1.01	1.01	1.00	1.00
0°	1	1	1	1	1	1	1	1
θ_2, θ_1	**10°**	**20°**	**30°**	**40°**	**50°**	**60°**	**70°**	**80°**

(ii) Case of symmetry arrangement

Especially, in case of symmetry arrangement ($P_{1,x} = -P_{2,x}, P_{1,y} = P_{2,y}, P_{1,z} = P_{2,z}$), using $(\theta_1, \theta_2) = (\theta_{1,1}, \theta_{1,2})$, $(\theta_{2,1}, \theta_{2,2})$. Eq. (13) is represented as follow.

$$E_M(\theta_1, \theta_2) = \frac{M\|V\|^2}{4r^2}\{2r^2 - 2P_{1,z}^2 + \frac{2P_{1,z}^2(r^2 - P_{1,z}^2)}{P_{1,y}^2}\}$$
$$= \frac{M\|V\|^2}{2}\frac{(1-\cos^2\theta_1\cos^2\theta_2)}{\sin^2\theta_1} \tag{15}$$
$$(0° < \theta_1 < 90°, 0° \le \theta_2 < 90°)$$

By theoretical calculation, we prove the following fact.

[Property 2]: Monotonicity of $E_M(\theta_1, \theta_2)$.

(i) When θ_1 increases, $E_M(\theta_1, \theta_2)$ also decrease.

(ii) When θ_2 increases, $E_M(\theta_1, \theta_2)$ also increase.

3. simulation of Evaluation value on sphere

This section presents the simulation results $E_M(\theta_1, \theta_2)$ (Eq.(15)), with $0° < \theta_1 < 90°$, $0° \le \theta_2 < 90°$, $\|\boldsymbol{V}\| = 1$ [m/s], $M = 2$, and $r = 1$.

Figure 3 shows the contact points on the upper hemisphere. **Table 1** shows the distribution of $E_M(\theta_1, \theta_2)$

at the contact points on the upper hemisphere in steps of θ_1 $(0° < \theta_1 < 90°)$ and θ_2 $(0° \leq \theta_2 < 90°)$.

As shown in **Table 1**, the value increases from the lower left of the table to the right and upward correspondingly [**Property2**]. $E_M(\theta_1, \theta_2)$ diverges infinitely as it approaches $(\theta_1, \theta_2) = (90°, 0°)$. In particular, when $\theta_2 = 0$, $E_M(\theta_1, \theta_2)$ is constant regardless of the contact position.

As shown in [1] and [2], when two constraint rollers are placed on the equator, the evaluation value is constant regardless of the angle of the two position vectors (see [**Property 1**]).

In the ball holding mechanism (evaluation of the placement of the world team) [4], the roller arrangement is on the upper hemisphere for ball transportation, but it is less-energy efficient than on the equator. Since the ball is not fixed by a pole caster, it is required to be placed on the upper hemisphere.

4. Conclusion

In this research, we defined an evaluation function as mean of roller's kinetic energy with respect to sphere direction angle and derived the exact formula. Furthermore, theoretically, we proved that points on equator are minimal.

Future issues include consideration of motion related to variable mechanisms with offset.

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, 2012The JapanSocietyofMechanicalEngineers.Vol.78,No.790, pp.2162-2170, 2012.

[3] K. Kimura, S. Chikushi , et al., Motion Analysis of a Sphere Driven by Dollers, Vol.38, No.5,pp.485-495, 2020.

[4] S. Chikushi, M. Kuwada, et al., Development of Next-Generation Soccer Robot"Musashi150"for RoboCup MSL,*30th Fussy System Symposium*, pp. 624-627,2014

[5] K. Kimura, K. Ishii, The Spherical Robot Transfer Problem With Minimal Total Kinetic Energy, The 2021 international Conference on Artificial Alife and Robotics (ICAROB2021), pp.226-270,2021.

Authors Introduction

Dr. Kenji Kimura

He is a mathematical teacher (International Baccalaureate) and engineering course chief in Fukuoka Daiichi High School, where he has been since 2014. He received the ME (mathematics) from Kyusyu University in 2002 and received his Ph.D. degree in engineering from Kyushu Institute of Technology in 2020.

Mr. Yusuke Abematsu

He is a mathematical teacher in Kagoshima Gyokuryu High School, where he has been since 2018. He received the ME (mathematics) from Kyusyu University in 2002.

He is positively working on establishing an educational method that connects closely mathematics and engineering.

Mr. Hiroyasu Hirai

He received his B.E., M.E., in Computer Science from Nippon Bunri University, Japan, in 2013, 2016, respectively. He is a 2nd year student in the doctoral program of the Kyushu Institute of Technology. He is an engineer of Garuda incorporated in 2016. His research interest includes autonomous robotics and machine learning(deep learning) .

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.

Robot Technology, and its Development Trend

Takao Ito
Graduate School of Engineering, Hiroshima University
1-4-1 Kagamiyama, Higashi Hiroshima, 739-7527, Japan
E-mail: itotakao@hiroshima-u.ac.jp

Abstract

Robot technology has been changed dramatically with massive development of internet environment. The author reviewed a plethora of literature and investigated advanced robot technologies. Today, most of typical robot technologies are used in single-cause-oriented products, such as robot vacuum cleaner and Asimo, a humanoid robot invented by Honda. These advanced products played important role in our modern society. For further development, a networking robot system with advanced technologies of internet and artificial intelligence is required to copy with the uncertainty in the future. Different technology should be combined and linked together for multiple-goal-oriented approach in the networking robot system. For evaluating the validity of our new system, a centrality index is introduced in this research.

Keywords: networking robot system, swarm, optimization, fragility.

1. Introduction

A plethora of studies on robot technology from the viewpoint of materials engineering, information processing technology such as pattern recognition and image processing, artificial intelligence, network, and biotechnology have been published in the past decades. However, studies focusing on the review of current robot technology to clarify its development trend are still sparse. This paper proposes a new networking robot system and a critical index to measure the structure of the system after reviewing the relevant literatures. This paper contributes to the development of robot technology by: (1) Identifying one of the future development directions, and (2) Proposing a new index to measure system structure to ascertain the effective system for future development.

This manuscript is structured as follows: Section 2 introduces the background of this research with literature review. In section 3, the paper explicates networking robot system, and calculates its structure using a new index of fragility. Based on our discussion, the study limitations are identified and directions for future research are proffered in the final section.

2. Background

A robot is the product of the robotics field, where programmable machines are built that can assist humans or mimic human actions while robotics is the intersection of science, engineering and technology that produces machines, called robots, that substitute for and/or replicate human actions [1]. Robot technology has been recognized as playing an important role not only in industrial society, but also in our daily life. Recently, many studies focusing on robot have been published.

A tracking system, also known as a locating system, one of the important technologies in robotics, is used for the observing of objects on the move and supplying a timely ordered sequence of location data for further processing. A tracking control of a two-wheeled mobile robot in both kinematic and dynamic models has been invested [2]. Chen and Jia suggested to use differential flatness and PD-spectral theory for controller design and proved the effectiveness of their proposed method by simulation results. Another problem of sensor selection for maneuvering target tracking in the cluttered environment also has been studied. By modeling the target dynamics as jump Markov linear systems, Li and Jia developed a decentralized tracking algorithm by applying the extended Kalman filter and the probabilistic data association technique [3]. Furthermore, Gholami et al., presented an inverse kinematic controller using neural networks for trajectory controlling of a delta robot in real-time. They found that the error in trajectory tracking is bounded, and the negative effect of joint backlash in trajectory tracking is reduced in the presence of external disturbance using the control scheme developed by their team [4]. In addition, a fast terminal sliding mode controller is designed for an active suspension gravity compensation system with unknown bounds of uncertainties and disturbances recently. Duan et al., proposed a new control scheme and verified that it can reduce the chattering effectively and render the high-precision tracking performance [5].

A small mobile robot developed by Alife Robotics Corporation Ltd., under demonstration is shown as in Figure 1.

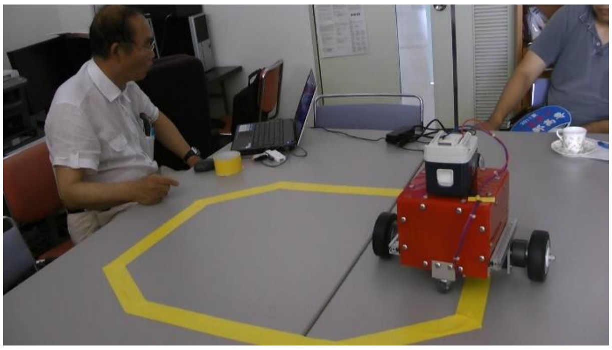


Figure 1. A small mobile robot developed by Alife Robotics Corporation Ltd.

The quadruped robots are considered as the effective tools for space exploration, military application, industrial use, and many more in different practical situations. Kitani et al., proposed the asymmetric amplification of the output waveforms of central pattern generators for excessive vibrations of quadruped robots in the roll direction when turning by controlling their hip yaw joint. They confirmed that the proposal method can suppress 43.7% vibration of the robot body in the roll direction and 7.4% vibration in the pitch direction compared with the conventional method [6]. A detailed survey concentrates on various design and development approaches for the quadrupedal robot, and environment perception techniques have been discussed [7].

To develop automatic harvesters is another important task. Matsuno et al. introduced new research of tomato harvesting robot and found smart tomato greenhouse aiming at promoting the automated tomato harvesting will reduce working time of harvesting after detailed analysis of the results mainly from tomato harvesting robot competition [8].

As illustrated in Figure 2, a new apple picking robot has been introduced in Australia farm [9].

Figure 2. A new apple picking robot.

Recently, molecule Robotics have been realized by using bio molecules such as DNA and proteins. DNA and proteins are well structured with a kind of "intelligence" for adapting to environment change. Suzuki and Taniguchi found that DNA molecule can sense concentration of single strand input sequence or quasi-input and chooses higher concentration one [10].

Table 1. Ten problems existing in robot technology.

	Ten problems
1	New materials and manufacturing schemes
2	Bionic robot and biological hybrid robot
3	Power and energy
4	Cluster robot
5	Navigation and exploration
6	The artificial intelligence of the robot
7	Brain computer interface
8	Social interaction
9	Medical robot
10	The ethics and safety of robot

3. Networking robot System

For further development of robot technology, at least ten problems exist. They are shown in Table 1 [11].

3.1. *Swarm system*

As one of the future development direction, swarm robotics become a hot issue today. Swarm robotics is an approach to the coordination of multiple robots as a system which consist of large numbers of mostly simple physical robots [12]. The difference between swarm robots and individual robots is that a swarm can commonly decompose its given missions to their subtasks. Compared with individual robots, a swarm is more robust to partial swarm failure and is more flexible about different missions [13]. One of the potential applications for swarm robotics is in search and rescue missions [14]. The key point of swarm system is the miniaturization and cost because a large number of robots are required in the system.

Undoubtedly swarm system should be one of the important directions for its further development in next decades. In real society, not only miniaturization, but also different robots having different functions will be combined together for different goals and purposes. Based on the survey mentioned above, and the real issues in our modern society, a networking robot system is proposed in this paper. The networking robot system is a novel approach to the combination and coordination of many different autonomous robots with different functions using internet and sense technology to obtain a certain task.

To realize optimization of the networking robot system, structure analysis is required.

3.2. *Networking robot system and Fragility*

As indexes of network structure, a large number of indexes have been developed for efficient and effective structure analysis. One of the most effective indexes is centrality. Calculation of centrality depends on its definition. At least more than 400 definitions have been developed [15]. For calculating the centrality of networking robot system, degree and fragility will be introduced as below.

Degree, as one of the basic indices of centrality, is considered as the basic index in network analysis. It can be calculated as follows [16].

$$C_D(p_k) = \sum_{i=1}^{n} a(p_i, p_k). \qquad (1)$$

where

i ≠ k;

a (pi, pk) = 1 if and only if pi and pk are connected by a line;

a (pi, pk) = 0 otherwise.

In an asymmetric network, two indexes of out-degree and in-degree should be calculated.

Fragility is a physical term, which characterizes how rapidly the dynamics of a material slow down as it is cooled toward the material transition [17]. Accordingly, fragility is defined as the ratio of the entire degree of and the entire degree after moving a specific node. It will be illustrated as follows.

$$\mathrm{F(p)}= \frac{C_D(\overline{p})-C_D}{C_D}. \qquad (2)$$

where

C_D: Entire degree of a given network;

$C_D(\bar{p})$: Entire degree after removing node p.

The equation of the entire network is defined as below [18].

$$C_D = \frac{\sum_{i=1}^{n}[C_D(p^*)-C_D(p_i)]}{n^2-3n+2}. \qquad (3)$$

where

$$C_D(p^*) = maxC_D(p_i). \qquad (4)$$

Data were drawn from two typical network system: networking robot system A and B. System A is an ordinary system and system B is a famous organization for its success compared with system A. Data on these two systems in 2007 is illustrated as Figure 3.

Figure 3. The structure of network.

Table 2. Results of fragility-sales regression model.

Sales	Networking Robot System A	Networking Robot System B
Fragility		
Partial regression coefficient	-1996633	-29733934.04
Standard coefficient	-0.1262	-0.7798
t value	-1.0945	-12.0763
Probability	0.2773	0
Correlation coefficient	-0.1262	-0.7798
Partial correlation coefficient	-0.1262	-0.7798
Coefficient of determination (R^2)	0.01593	0.60807
Multiple correlation coefficient	0.12621	0.77979
F value	1.19787	145.83724
Degree of freedom	1, 74	1, 94
AIC	2508.45	3044.02
DW ratio	1.1641	0.9052
Data number	76	96

The results of fragility-sales regression model are shown in Table 2.

Compared with system A, the probability of fragility is significant, and coefficients of determination are higher. The model is effective for good performance organization. Furthermore, both correlation coefficients of the two models are negative. Thus, the assumption of higher fragility having an inverse association with sales is confirmed. Higher fragility, lower performance holds.

4. Conclusion

In this paper, a new index of network system fragility is proposed. To prove its validity, regression model using fragility and sales has been tested. There are different perspectives such as economy, efficiency, safety, and liability to evaluate a system. To draw firmer conclusions, future studies should be examined using multi-year data. Furthermore, additional indexes for these conclusions are required.

References

1. What is robotics? Built In, retrieved November 14, 2021.
2. Liming Chen, Yingmin Jia, Variable-poled Tracking Control of a Two-wheeled Mobile Robot Using Differential Flatness, Journal of Robotics, networking and Artificial Life, Volume 1, Issue 1, 12-16, June 2014. retrieved November 14, 2021.
3. Wenling Li, Yingmin Jia, Global sensor selection for maneuvering target tracking in clutter, Journal of Robotics, networking and Artificial Life, Volume 3, Issue 2, 128-131, September 2016. retrieved November 14, 2021.
4. Akram Gholami, Taymaz Homayouni, Reza Ehsani, and Jian-Qiao Sun, Inverse Kinematic Control of a Delta Robot Using Neural Networks in Real-Time, Robotics 2021, 10(4), 115. retrieved November 14, 2021.
5. Meng Duan, Jiao Jia, Takao Ito, Fast terminal sliding mode control based on speed and disturbance estimation for an active suspension gravity compensation system, Volume 155, Mechanism and Machine Theory, January 2021, retrieved November 14, 2021.
6. Makoto Kitani, Ryo Asami, Noritaka Sato, Yoshifumi Morita, Tomofumi Fujiwara, Takahiro Endo, Fumitoshi Matsuno, Suppression of Roll Oscillation in Turning of Quadruped Robot by Asymmetric Amplification of Central Pattern Generator Output Waveform, Journal of Robotics, networking and Artificial Life, Volume 6, Issue 2, 79-83, September 2019, retrieved November 14, 2021.
7. Priyaranjan Biswal, Prases K. Mohanty, Development of quadruped walking robots: A review, Ain Shams Engineering Journal, Volume 12, Issue 2, 2017-2031, June 2021, retrieved November 14, 2021.
8. Takayuki Matsuo, Takashi Sonoda, Yasunori Takemura, Masanori Sato, Kazuo Ishii, Toward Smart Tomato Greenhouse: The Fourth Tomato Harvesting Robot Competition, Journal of Robotics, networking and Artificial Life, Volume 6, Issue 2, 138-142, September 2019, retrieved November 11, 2021.
9. The newest robots 2021, Incredible and technologically advanced robots, retrieved November 11, 2021.
10. Yasuhiro Suzuki, Rie Taniguchi, Toward Artificial Intelligence by using DNA molecules, Journal of Robotics, networking and Artificial Life, Volume 5, Issue 2, 128-130, September 2018, retrieved November 14, 2021.
11. Development of robot technology: Ten problems existing, retrieved November 12, 2021.
12. Junyan Hu, Parijat Bhowmick, Alexander Lanzon, Two-layer distributed formation-containment control strategy for linear swarm systems: Algorithm and experiments, International Journal of Robust and Nonlinear Control. 30 (16): 6433–6453, retrieved November 12, 2021.
13. Eugene Kagan, Nir Shvalb, Irad Ben-Gal, Autonomous Mobile Robots and Multi-Robot Systems: Motion-Planning, Communication, and Swarming, John Wiley & Sons, 2019, retrieved November 12, 2021.

14. Khalil Al-Rahman Youssefi, Modjtaba Rouhani, Swarm intelligence based robotic search in unknown maze-like environments, Expert Systems with Applications, Volume 178, 15 September 2021, 114907, retrieved November 14, 2021.
15. Alphabetically/Chronologically sorted list, retrieved November 14, 2021.
16. Nieminen J. (1974). On centrality in a graph, Scandinavian Journal of Psychology 15, 322-336, 1974, Retrieved December 15, 2021.
17. Ediger M. D., Angell, C. A., Nagel S. R., (1996). Supercooled liquids and glasses, Journal of Physical Chemistry 100 (31): 13200–13212, August 1996, retrieved November 14, 2021.
18. Linton C. Freeman (1978/79). Centrality in Social Networks Conceptual Clarification, Social Networks 1, 215-239, retrieved November 14, 2021.

Authors Introduction

Dr. Takao Ito

Dr. Takao Ito 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 Design Tool of Hula Costumes

Taketo Kamasaka, Kodai Miyamoto
Graduate School of Engineering, University of Miyazaki,
Miyazaki-City, Miyazaki, Japan

Makoto Sakamoto*, Satoshi Ikeda, Amane Takei, Kenji Aoki
Faculty of Engineering, University of Miyazaki,
Miyazaki-City, Miyazaki, Japan

Tsutomu Ito
National Institute of Technology, Ube College
Ube-City, Yamaguchi, Japan

Takao Ito
Graduate School of Engineering, Hiroshima University
Higashi-Hiroshima, Hiroshima, Japan

E-mail: hm16011@student.miyazaki-u.ac.jp, hm16043@student.miyazaki-u.ac.jp, fruits2000jp@yahoo.co.jp
**Corresponding Author*

Abstract

Miyazaki Prefecture has a large hula population, probably due to its similarity to Hawaii in mythology and climate. On the other hand, many hula costumes are handmade, and it costs tens of thousands of yen to produce an original design. So we are developing a 3D CAD (computer-aided design) system for hula costumes, based on the idea that we can reduce the number of failures by checking the behavior of the fabric when danced in the designed costume before production.

Keywords: Cloth-Simulation, Hula Costume, Computer Graphics, Costume Design

1. Introduction

In recent years, three-dimensional computer graphics (3DCG) technology has been applied in various fields; there are several apparel design tools using 3DCG, but there is no precedent specific to the hula.[1, 2] Therefore, as a preliminary step in developing a design tool for hula costumes, we developed a tool that allows users to select and simulate the design and color of a skirt. In addition to the design and color, this tool has the ability to deform the human body model based on the body measurements. We are also planning to develop a function to change the length of the skirt. With these features, you will be able to find the right skirt size for you.

2. Research Background

In this study, we improved the 2D hula costume design tool shown in Fig. 2, which outputs the execution results as shown in Fig. 1, to output the execution results in real time with 3D animation as shown in Fig. 3. By changing from 2D to 3D animation, it is easier to visualize the finished skirt.

Fig. 1. 2D design tool execution results.

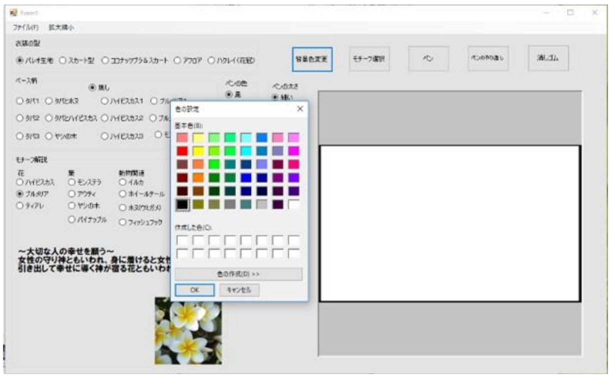


Fig. 2. 2D design tool.

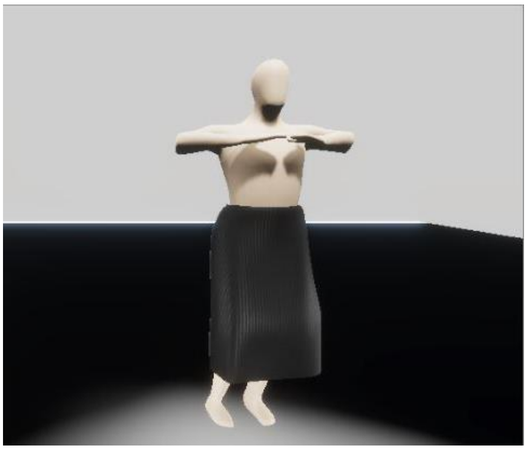


Fig. 3. 3D design tool execution results.

3. Simulation Environment

This study was simulated in the environment shown in Table 1.

Table 1. Simulation Environment

OS	Windows 10 Pro
Memory	16GB
Processor	Intel® Core ™ i7-7700 3.60GHz
GPU	Quadro M4000
Programming language	C#
Development Software	Unity
CG modeling software	Blender

4. Processing details of each function

This section describes the processing of the functions we have developed for deforming the human body model based on body measurements and selecting the color and design of the skirt.

4.1. *Transformation of human body models based on body measurements*

The human body model will be deformed based on the measurements of the actual body. In this case, we will measure the length of the four parts of the body and the circumference of the seven parts. Then, by inputting the measurements into the input form shown in Fig. 4, the human body model is deformed into the same body shape as the person who took the measurements.

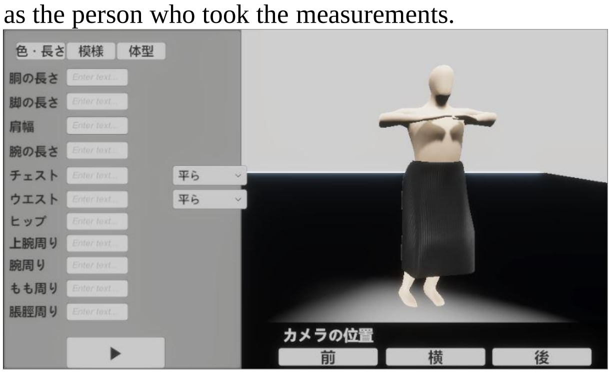


Fig. 4. Input form for measurement values.

4.1.1. Transformation of the length of each part

This time, assuming the height of the human model before deformation to be 160 cm, the following procedure was used to determine the magnification/reduction ratio of the length of the torso, legs, shoulder width, and arms.

(i) The vertical length of each bone of the human body model was calculated using the coordinates of the image in Fig. 5. The results are shown in Table 2 below. The head, neck, chest, back, hips, buttocks, thighs, shins, and feet were measured for vertical length. Shoulders, upper arms, forearms, and hands were measured in horizontal length.

Table 2. Length of each bone in coordinates.

Part	Length
Head	79
Neck	40
Chest	66
Back	54
Waist	43
Hip	49
Thigh	146
Shin	150
Foot	37
Shoulder	69
Upper Arm	74
Forearm	98
Hand	59

(ii) The vertical lengths of the head, neck, chest, back, hips, thighs, shins, and feet were added together and all divided by the head length of 79. The result of that calculation is about 8.40.

(iii) Since we assumed that the model before deformation is 160 cm, we divide the result of step 2 by 160. The result of this calculation is about 19.04.

(iv) Then, the following formula is used to calculate the length of each bone on the process of 160cm height, L_{Born}.

$$L_{Born} = 19.04 \times L_i \div 79 \tag{1}$$

Let L_i be the length of each bone in the coordinate.

(v) Finally, dividing the actual measurements of each part of the body by the calculation results of step 4 determines the magnification/reduction ratio of the parts of the human body model.

The measured body, leg, and arm lengths are the sum of the back/waist, thigh/shin, and upper arm/forearm lengths, respectively.

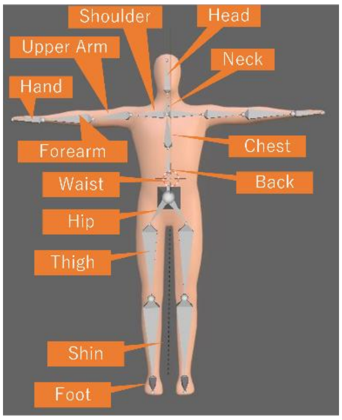


Fig. 5. 3D model of a human body (front view).

4.1.2. Transformation of the length of the perimeter of each part

Assuming the height of the human model before deformation to be 160 cm, the following procedure was used to calculate the magnification/reduction ratio for the chest, waist, buttocks, upper arms, forearms, thighs, and shins.

(i) The long and short diameters of the chest, waist, hips, upper arms, forearms, thighs, and shins of the human model were calculated using the coordinates of the images in Figures 5, 6, and 7. The results are shown in Table 3 below.

Table 3. Long and short diameters of each part.

Part	Long Diameter	Short Diameter
Chest	148	96
Waist	112	67
Hip	125	76
Upper Arm	37	33
Forearm	29	26
Thigh	67	57
Shin	47	41

(ii) Consider the cross section of each part of the body as an ellipse. The approximate formula for the length L of the perimeter of the ellipse is shown below.[3]

$$L = 2\sqrt{4(a-b)^2 + \pi^2 ab} \tag{2}$$

Here, a is the short diameter and b is the long diameter.

(iii) If the ratio A of short and long diameters is fixed, the following equation can be used to calculate the short and long diameters from the perimeter length. When $b = Aa$,

$$a = \sqrt{\frac{L^2}{4\{4-(8+\pi^2)A + 4A^2\}}} \tag{3}$$

(iv) From the short and long diameters a and b in Step 3, the magnification factors s_a for the short diameter and s_b for the long diameter are calculated by the following formula.

$$s_a = \frac{2a}{19.036 \times L1_i \div 79} \tag{4}$$

$$s_b = \frac{2Aa}{19.036 \times L2_i \div 79} \tag{5}$$

Here, $L1_i$ is the long diameter of each part, and $L2_i$ is the short diameter of each part.

Since the ratio of the long and short diameters of the chest and waist can vary greatly from person to person, in this study, for the chest and waist, there were three items to choose from: "flat," "slightly round," and "round" in terms of the shape seen from the side. The ratios of the long and short diameters when each item was selected are shown in Table 4 below.

Table 4. Ratio of the long and short diameters of each part when

each item is selected.

Item	Ratio (Long Diameter: Short Diameter)
Flat (Chest)	154.2 : 100
Slightly round (Chest)	150 : 100
Round (Chest)	145 : 100
Flat (Waist)	100 : 59.8
Slightly round (Waist)	100 : 70
Round (Waist)	100 : 80

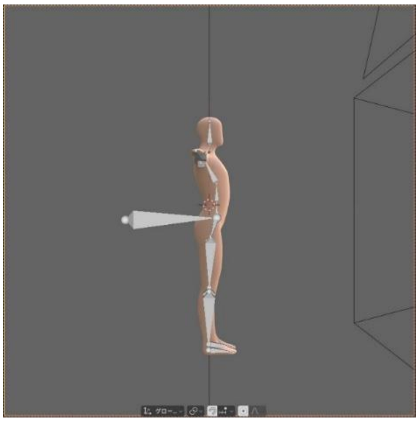


Fig. 6: 3D model of a human body (side view).

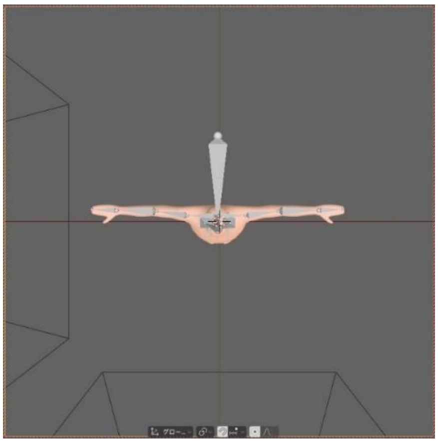


Fig. 7: 3D model of a human body (top view).

4.2. *Change the color of the skirt*

The color of the skirt is changed by moving each RGB slider in Fig. 8.

Fig. 8. Change the color of the skirt.

The RGBA color model was used to represent the colors. In this model, colors are represented by determining the R (red), G (green), B (blue), and A (transparency) parameters from 0% to 100%.[4] In this system, the parameter is set to 0 for 0%, and 1 for 100%. A is fixed at 1 to avoid transparency. The initial state is (R,G,B,A)=(0,0,0,1), and the color can be changed by moving each slider. The slider is 0 on the left and 1 on the right. The processing procedure is shown below.

(i) Set the color parameters to (R,G,B,A) = (0,0,0,1) and the color of the skirt to black.
(ii) If the skirt has a texture set, set the texture parameter to null.
(iii) Get the RGB parameters of the current skirt color.
(iv) Get the value of each slider.
(v) Assign the value of each slider to each RGB parameter.

By repeating steps 3 through 5, the color of the skirt can be changed in real time.

4.3. *Change the pattern of the skirt*

The pattern of the skirt can be changed by clicking on the button for each pattern in Fig. 9.

Fig. 9. Change the pattern of the skirt.

The pattern is changed by setting the skirt color to the white parameter (R,G,B,A)=(1,1,1,1) and mapping the chosen texture to the skirt when the button is pressed.

Currently, we are using Unity's "Triplanar", which projects images onto the skirt, so the pattern stays in a certain position when the human model moves.

5. Execution result

The following shows the results of the run when the human model was deformed and the color and pattern of the skirt were changed.

5.1. *Results of deforming a human body model*

Fig. 10 shows a variant of the human body model based on the average of measurements of Japanese males aged 20 to 24. This time, the torso length and arm length were calculated approximately, and the average values of the entire Japanese male population were input for hip and upper arm circumference.[5, 6, 7]
The result, as you can see in Fig. 10, is that the legs have rotated inward. This is thought to be due to the enlarged hip bones.

Fig. 10. Model of an average Japanese male body shape.

The following Fig. 11 to 13 show the deformation of the human body model when the shape of the waist from the side is changed to "flat", "slightly round", and "round" on the model of the average of the measurements of Japanese men aged 20 to 24. It can be seen that there is a difference in the roundness of the shape of the stomach between the "flat" and "rounded" waists.

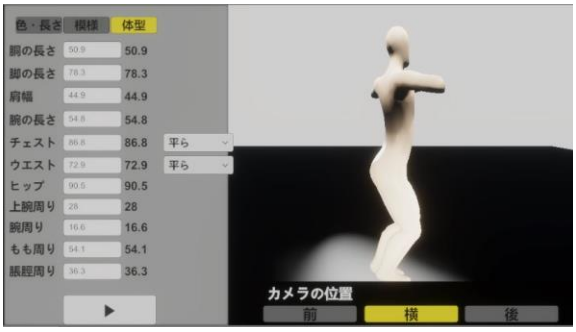


Fig. 11. Change the shape of the belly (flat).

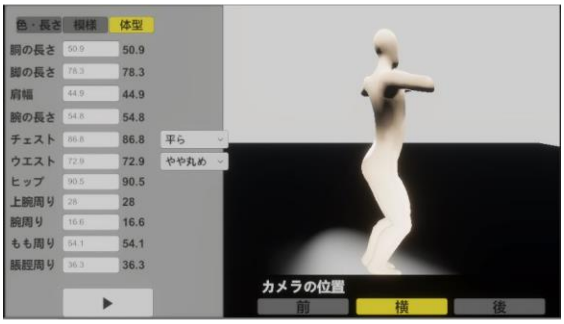


Fig. 12. Change the shape of the belly (slightly round).

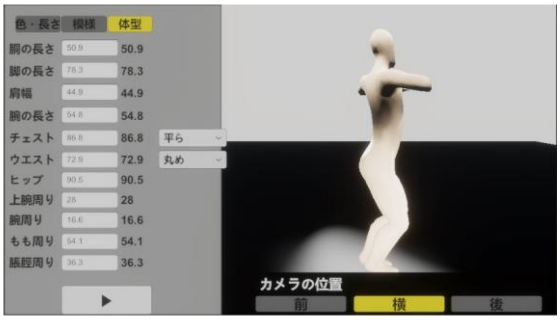


Fig. 13. Change the shape of the belly (round).

Fig. 14 shows a variant of the human body model based on the average of measurements of Japanese women aged 20 to 24. This time, the torso length and arm length were calculated approximately, and the average of all Japanese women was input for the upper arm circumference.[5, 8, 9]
The result, as can be seen in Fig. 14, is an unnatural balance between the head and the torso, since the head size remains the same as before the deformation.

Fig. 14. Model of an average Japanese female body shape.

The following Fig. 15 to 17 show the deformation of the human body model when the shape of the chest from the side is changed to "flat", "slightly rounded" or "rounded" on the model of the average of the measurements of Japanese women aged 20 to 24. It can be seen that there is a difference in roundness in the shape of the chest between "flat" and "rounded" chests.

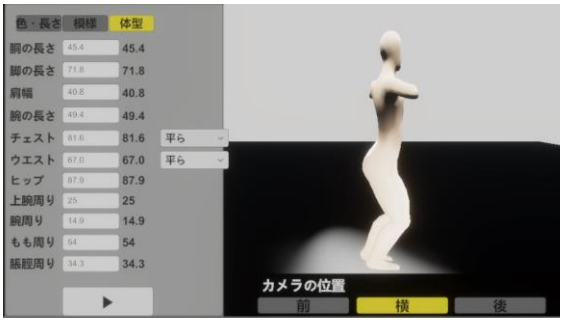


Fig. 15. Change the shape of the chest (flat).

Fig. 16. Change the shape of the chest (slightly round).

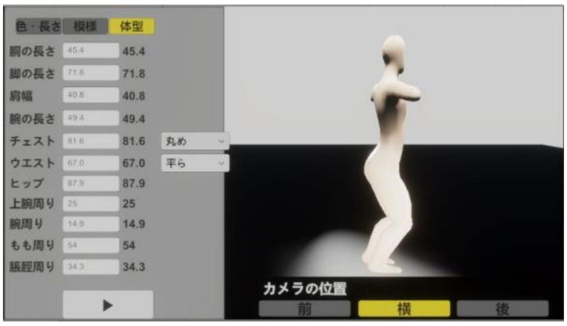


Fig. 17. Change the shape of the chest (round).

5.2. *Result of changing the color of the skirt*

When all the color sliders are moved to the right, the skirt color will be white, as shown in Fig. 18. When only the green slider is moved to the right, the color of the skirt will be green, as shown in Fig. 19.

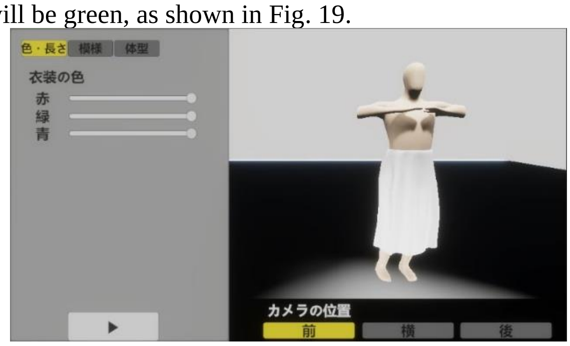


Fig. 18. Change the color of the skirt (white).

Fig. 19. Change the color of the skirt (green).

5.3. *Result of skirt pattern change*

The results of the run when the skirt pattern is set to Hibiscus are shown in Fig. 20 and Fig. 21. When the human body model is not moving as shown in Fig. 20, the texture mapping to the skirt appears to be successful. However, when the animation of the human body model is played back as shown in Fig. 21, the texture is projected at a fixed position, indicating that it does not follow the movement of the skirt.

Fig. 20. Change the pattern of the skirt (before playing the animation).

Fig. 21. Change the pattern of the skirt (during animation playback).

6. Consideration

The following is a discussion of the results of the transformation of the human body model, the change of the color of the skirt, and the change of the skirt pattern.
In the first function, the deformation of the human body model, the expansion of the hips caused the legs to rotate inward. This could be solved by rotating the hip bones in accordance with the hip expansion. In addition, when the model was transformed into a model of an average Japanese female body shape, the balance of the body shape became unnatural due to the fact that the head size did not change. This could be solved by changing the size of the head to match the expansion of the torso length, etc. In the future, we will input the average human body dimensions of a four-year-old, which is the age when most people start hula dancing, and verify how the human body model deforms.
In the second function, changing the color of the skirt, the color looks a little bright, probably due to the lighting in the virtual space. In the future, we would like to adjust the lighting so that it expresses more correct colors.
In the third function, changing the pattern of the skirt, we found that the texture was projected at a fixed position and did not follow the movement of the skirt. This suggests that the method of mapping the image onto the skirt needs to be reviewed.

7. Conclusion

This time, we have developed a tool that can deform the human body model based on measurements and change the color and pattern of the skirt. As for the deformation of the human body model, we would like to do four things in the future: rotate the hip bones to match the expansion of the hip, change the size of the head to match the body shape, deform the human body model based on the average human body size of four-year-olds who tend to start hula dancing, and display an explanation of where to measure the length.
For the color change, we would like to adjust the lighting to represent the correct color. We would also like to display the percentage of the slider that has been moved when it is moved.
In changing the pattern, I would like to revise the texture mapping method so that the image can be mapped while following the skirt correctly.
And in the future, I would like to add the ability to change the length of the skirt, display the length, and generate a pattern, to develop a practical tool.

Acknowledgements

I would also like to express my sincere gratitude to Associate Professor Masato Sakamoto, who gave me a great deal of advice on the direction of my research when I was stuck in my research.

References

1. KIKI FASHION [Online]. https://f.kiki.ooo/82695432679
2. Rose Sinclair, "Textiles and Fashion: Materials, Design and Technology," Woodhead Publishing, 2024.
3. "Takakazu Seki's approximate formula for finding the elliptic circumference." [Online]. http://www.tcp-ip.or.jp/~n01/math/analysis/seki/seki.pdf
4. ZD Net Japan, "RGBA" [Online]. https://japan.zdnet.com/glossary/exp/RGBA/?s=4
5. Paroday, "Average human body dimensions (hand length, wrist circumference, foot length, foot width, etc.)" [Oneline]. https://paro2day.blog.fc2.com/blog-entry-399.html
6. BELCY, "Average waist size for men and women｜Calculating standard size and what is the ideal three sizes?" [Online]. https://belcy.jp/48403
7. BELCY, "What is the average or ideal thickness of a man's arms? Where to measure, how to measure, and muscle training for thicker upper arms." [Oneline]. https://belcy.jp/72734
8. Activel!, "What is the average thickness of a woman's arms? Find out the ideal thickness and the thickness of models!" [Online]. https://activel.jp/bodymake/xAIqx
9. Domani, "What is the golden ratio of three sizes? Learn how to measure correctly and find the right size for you!" [Online]. https://domani.shogakukan.co.jp/332248

Authors Introduction

Mr. Taketo Kamasaka

He is a master student at Department of Computer Science and System Engineering, University of Miyazaki. His current research interests are computer graphics, entertainment computing, and so on.

Mr. Kodai Miyamoto

He is a master student at Department of Computer Science and System Engineering, University of Miyazaki. His current research topic is a fundamental study on educational support using VR technology

Prof. Makoto Sakamoto

He received the Ph.D. degree in computer science and systems engineering from Yamaguchi University in 1999. He is presently a professor in the Faculty of Engineering, University of Miyazaki. His first interests lay in hydrodynamics and time series analysis, especially the directional wave spectrum. He is a theoretical computer scientist, and 2. 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 forth. He has published many research papers in that area. His articles have appeared in journals such as Information Sciences, Pattern Recognition and Artificial Intelligence, WSEAS, AROB, ICAROB, SJI, IEEE and IEICE (Japan), and proceedings of conferences such as Parallel Image Processing and Analysis, SCI, WSEAS, AROB and ICAROB.

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. 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 the Use of XR Technology to Support Science Education

Kodai Miyamoto, Taketo Kamasaka
Graduate School of Engineering, University of Miyazaki,
Miyazaki-City, Miyazaki, Japan

Makoto Sakamoto*, Masahiro Yokomichi, Satoshi Ikeda, Amane Takei
Faculty of Engineering, University of Miyazaki,
Miyazaki-City, Miyazaki, Japan

Tsutomu Ito
National Institute of Technology, Ube College
Ube-City, Yamaguchi, Japan

Takao Ito
Graduate School of Engineering, Hiroshima University
Higashi-Hiroshima, Hiroshima, Japan

E-mail: hm16043@student.miyazaki-u.ac.jp, hm16011@student.miyazaki-u.ac.jp, fruits2000jp@yahoo.co.jp
**Corresponding Author*

Abstract

According to the results of a survey on science teaching in 2016, the percentage of students who answered that they like science is lower than other subjects. However, more than 80% of students said that they like experiments and observations. In addition, a 2019 survey on smartphone penetration showed that about 90% of students are familiar with the technology. Also, XR technology has made remarkable progress in recent years. Based on the above, I conducted this research because I thought that creating a simulation application using XR technology with smartphones would change the way we think about science classes. In this paper, we have developed a simulation application for science experiments. The subjects were asked to experience the created application and answer a questionnaire. As a result, the average score was 4 out of 5, which was not a bad result. At the same time, however, we found a problem. The problem was that since this was a simulation application, the user experience was not very good. So we wanted to make it a little easier to use. While improving the problems, I would like to create applications for other fields as well.

Keywords: Education, science, chemistry, physics, experiment / observation, virtual reality, simulation application.

1. Introduction

In the materials related to science surveyed by the Science Working Group in 2016, it was found that the percentage of high school students who answered "I like studying science" and "It is important to study science" is lower than other subjects.[1] However, the percentage of elementary and junior high school students who answered "I like experiments and observations" exceeds 80%.[2]

However, not enough science experiments are being conducted. The main reasons for not conducting science experiments are lack of time, equipment, and space for preparation and cleanup.[3]

In addition, a survey of smartphone penetration in 2019 showed that about 90% of students have one.[4] Most schools have PC classrooms. In addition, each classroom is equipped with a computer for teachers. Under these circumstances, it is believed that the educational effects can be further enhanced in various classes.

In light of the above, the purpose of this research is to contribute to education by creating a science simulation application using VR technology to solve the current situation where science experiments are not sufficiently conducted.

Fig. 1. Example of an English conversation learning application that supports VR[5]

2. Physics experiment application

In this study, we attempted to simulate the physics of "projectile motion" and "falling body motion".

In the initial state, the selection screen shown in the figure below is displayed, and by selecting a button object, you can move to the "Projectile Motion" and "Falling Body Motion" screens. After transitioning to each screen, the screen transition to the selection screen is made by selecting "Return to Selection Screen".

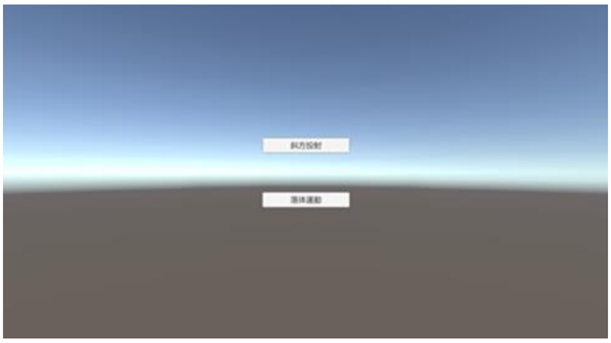


Fig. 2. The start screen I created

Since these experiments require large experimental tools and it is sometimes difficult to obtain accurate values, we thought it would be suitable to conduct them in a virtual environment.

2.1. *Development environment*

This survey was conducted in the environment shown in Table 1.

Operating system	Windows10
Programming language	C#
Software	Unity 2019.2.15f1

Table. 1. Development environment

2.2. *Projectile motion*

Projectile motion moves from the initial position. After deciding the angle and speed and pressing Start Button, the object will be fired and the distance will be displayed. If it is difficult to check the current situation, you can zoom in and out on the ball with the zoom below.

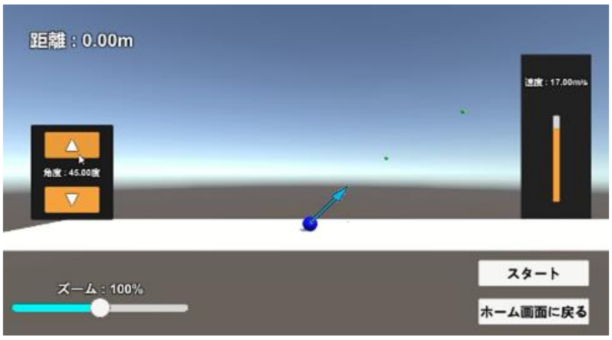


Fig. 3. Created projectile motion application

2.3. *Falling exercise*

I made a slope that controls the falling speed by rolling the ball while changing the angle of the slope. Determine the angle of tilt and press Start to start spinning the sphere. The tilt angle, ball speed, and ball position are displayed.

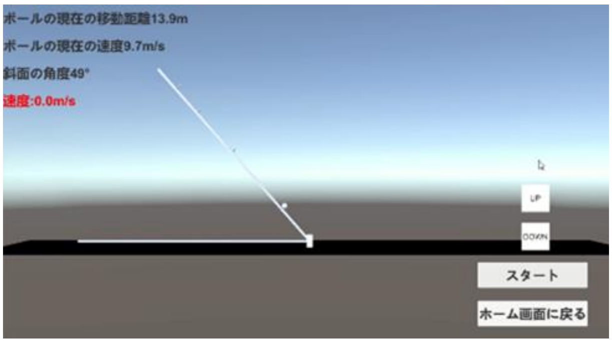


Fig. 4. Created falling exercise application

3. Chemistry experiment application

In this study, we tried to create a prototype chemical experiment simulation of the "flame color reaction" and the "silver mirror reaction".

In the initial state, the selection screen shown below is displayed, and by selecting the button object, you can move to the "Flame Color Reaction" and "Silver Mirror Reaction" screens. After moving to each screen, selecting "Back" will move the screen to the previous screen.

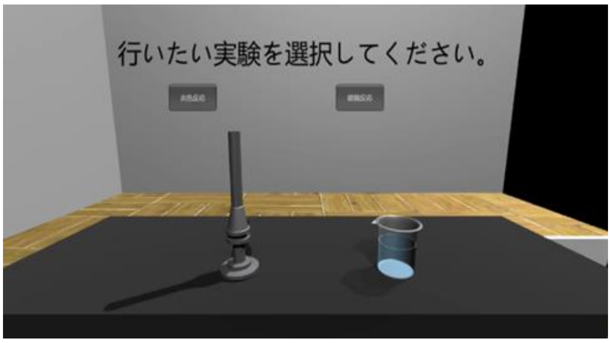


Fig. 5. The start screen I created

These experiments require a lot of solutions and experimental tools, which can be difficult to prepare and clean up, and the experiments take a long time, so we thought this would be a good solution to the lack of class time.

3.1. *Development environment*

This survey was conducted in the environment shown in Table 2.

Operating system	Windows10
Programming language	C#
Software	Unity 2019.2.15f1
	Blender 2.81

3.2. *Flame color reaction*

First, select an aqueous solution. There are seven types of aqueous solutions. From left to right: lithium chloride solution, sodium chloride solution, potassium chloride solution, calcium chloride solution, strontium chloride solution, barium chloride solution, and copper (II) chloride solution.

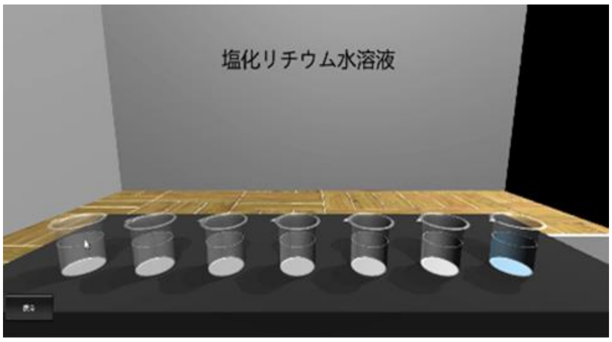


Fig. 6. Created flame color reaction application

When you select an aqueous solution, a flame color reaction occurs. Lithium chloride aqueous solution is red, sodium chloride aqueous solution is yellow, potassium chloride aqueous solution is purple, calcium chloride aqueous solution is orange, strontium chloride aqueous solution is beni, barium chloride aqueous solution is yellowish green, and copper(II) chloride aqueous solution is green.

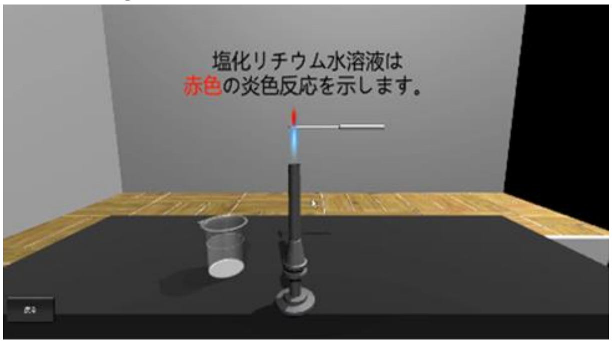


Fig. 7. Created flame color reaction application

3.3. *Silver mirror reaction*

The process of silver precipitation is shown with the chemical reaction equation.

Fig. 8. The first stage of the application of the silver mirror reaction that I made

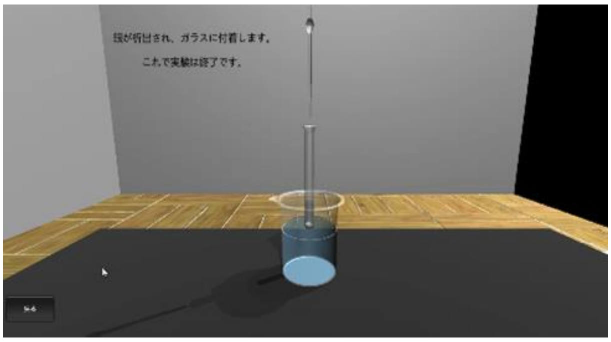


Fig. 9. The final stage of the application of the silver mirror reaction made

4. Evaluation experiment

An evaluation experiment was conducted to verify whether the developed science experiment simulation application was useful. We conducted a questionnaire to 10 people, including students in the laboratory, and evaluated the usefulness and usability of the system. The contents of the questionnaire are the following four points.

Evaluation item 1: Was it easy to operate?
Evaluation item 2: Was the result of the experiment easy to understand?
Evaluation item 3: Was it easy to imagine a physical phenomenon?
Evaluation item 4: Opinions and points to be improved (free description)
Evaluation items 1 to 3 will be evaluated on a scale of 1 to 5 points, and evaluation item 4 will be freely described.

4.1. *Experimental result*

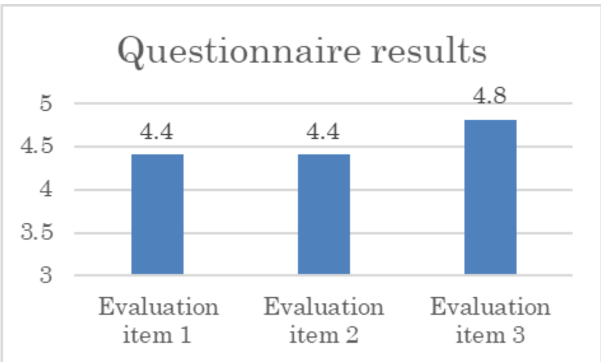


Fig. 10. Questionnaire average score

The average score of 4.4 was obtained for the evaluation items 1: Was it easy to operate? and 2: Was the result of the experiment easy to understand? from this result, we can say that the application is easy for anyone to use.

For the third question, " Was it easy to imagine a physical phenomenon?", the average score was 4.8. From this result, we can say that it is easy to visualize the phenomena in science.

In response to evaluation item 4, " Opinions and points to be improved (free description)", the following comments were given as good points: "It was good that the sound showed that the gravitational acceleration is constant in the vertical direction," and "It was easy to understand the trajectory of the sphere.

On the other hand, the following comments were made as "improvements": "It would be easier to understand the difference between the new object and the previous one if the distance of the previous object is displayed," "It would be easier to operate if the speed and angle can be input by the experimenter," "It would be easier to understand if the trajectory of the ball is connected with a line," and "It would be easier to understand how far the object actually flew if there is a memory in the background.

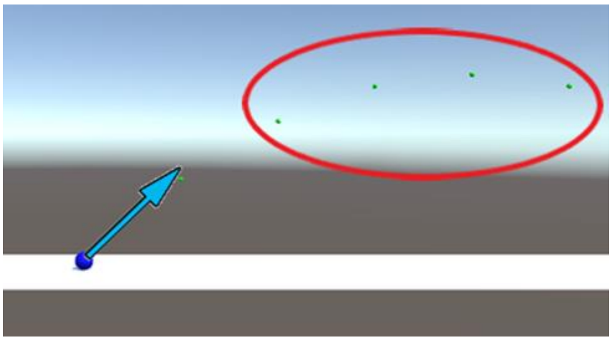


Fig. 11. Examples of questionnaire improvement points

5. Consideration

We asked five people to help us evaluate the application through a questionnaire, and the results show that the application was useful. The results show that the application is useful, but there are a few things that could be improved.

First of all, the distance measurement was done when the object was completely stopped. However, the problem of oblique projection in high school physics often involves finding the distance to the part of the object that lands on the ground, so I think we need to devise a way to make a physics experiment simulation application that is closer to the kind of material that students actually use on a regular basis.

Regarding evaluation item 4, there were many opinions about the expansion of functions. Extension to VR was also mentioned as an extension of the function. With regard to expansion to VR, effects such as immersive feeling and easy image of physical phenomena can be expected. However, since VR-specific equipment is required, I think we must devise ways to make it easy for anyone to use.

6. Future tasks

If we can create a simulation application that is similar to the teaching materials that students usually handle, we can support education not only in science but also in various subjects. Also, before creating an application, you need to think about what kind of application can support student education.

This time, it could not be expanded to VR. If it is possible to expand to VR, it will be possible to realize a more immersive feeling by actually expressing the sound and science room during the experiment, so this is a future issue.

7. Summary

As mentioned at the beginning, VR technology has grown remarkably in recent years. In this research, we tried to create an application that can perform science experiments using virtual reality, but as mentioned in future tasks, special equipment is required to reproduce VR. It has become a trial of only CG images. However, from the results of the evaluation experiment, it was found that a useful application was created. Ultimately, I think it is to create a system that will help support education.

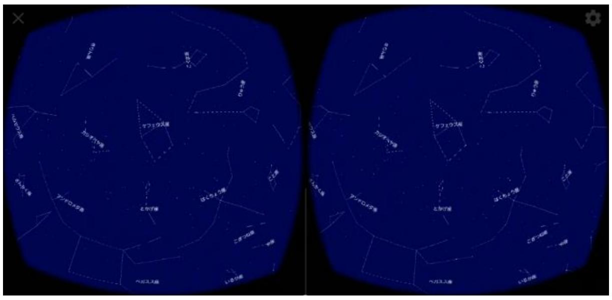


Fig. 12. Extension to VR

References

1. Curriculum Subcommittee Science Working Group "Materials on Science", February 5, 2016. https://www.mext.go.jp/b_menu/shingi/chukyo/chukyo3/060/siryo/__icsFiles/afieldfile/2016/02/19/1367079_7.pdf
2. National Academic Achievement / Learning Situation Survey, 2015. https://www.nier.go.jp/15chousakekkahoukoku/hilights.pdf
3. "Results of the 2012 Survey on the Status of Science Education in Junior High Schools (preliminary report). " https://www.jst.go.jp/cpse/risushien/secondary/cpse_report_016.pdf
4. "[Latest version] The smartphone penetration rate in 2019 is released to the public by age group, region, and age group! It's a smartphone-only era! I predicted how marketing should change in the future." Marketing Research Camp. https://marketing-rc.com/article/20160731.html
5. "AEON provides VR English conversation app ..." Hospitality "simulated experience improves ability". https://resemom.jp/article/2017/07/18/39278.html

Authors Introduction

Mr. Kodai Miyamoto

He is a master student at Department of Computer Science and System Engineering, University of Miyazaki. His current research topic is a fundamental study on educational support using VR technology.

Mr. Taketo Kamasaka

He is a master student at Department of Computer Science and System Engineering, University of Miyazaki. His current research interests are computer graphics, entertainment computing, and so on.

Prof. Makoto Sakamoto

He received the Ph.D. degree in computer science and systems engineering from Yamaguchi University in 1999. He is presently a professor in the Faculty of Engineering, University of Miyazaki. His first interests lay in hydrodynamics and time series analysis, especially the directional wave spectrum. He is a theoretical computer scientist, and 2. 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 forth. He has published many research papers in that area. His articles have appeared in journals such as Information Sciences, Pattern Recognition and Artificial Intelligence, WSEAS, AROB, ICAROB, SJI, IEEE and IEICE (Japan), and proceedings of conferences such as Parallel Image Processing and Analysis, SCI, WSEAS, AROB and ICAROB.

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. 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. 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.

Analysis of 5x5 board Quoridor

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

Takeo 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

Abstract

Quoridor is a perfect information game played by two or four players. The purpose of this study is to classify this game as a first move win, a second move win, or a draw. However, the analysis of the standard version of this game, 9x9 board, requires a huge amount of calculation, so in this paper, we show one of a perfect plays of a reduced version with a 5x5 board and one fence each.

Keywords: Perfect play, Retrograde Analysis, Quoridor, combinatorial theory

1. Introduction

Quoridor is a strategy game designed by Mirko Marchesi and released by Gigamic Games* in 1997. In the year it was released, Quoridor was selected for the Mensa Mind Game award†.

Quoridor is played on a 9x9 board with two or four players (sometimes unofficially three). Quoridor for two-player starts from the initial position in Fig.1 with 20 fences per player. Do either "move one of own pawn" or "install one fence" on one's turn, advance the turn in order, the player who first reaches any space on the opposite side of the game board wins.

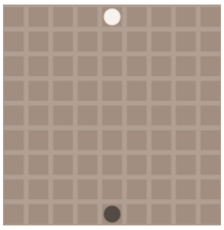


Fig 1. Initial position of Quoridor for two-player.

Quoridor for two-player is categorized into two-player zero-sum finite deterministic games of perfect information[1,2]. Games in this class are possible to look ahead in theory, thus if both players choose constantly the best move, these are classified into win, loss or draw game for the first player (the sequence obtained in this way is called perfect play). This game is still young among board games and has not been studied much, so it is not clear whether it is a win, a loss or a draw game for the first move player.

Therefore, this paper uses retrograde analysis to show one of a perfect plays of a reduced version with a 5x5 board and one fence each.

2. Quoridor

In this paper, we deal with a miniature board Quoridor. 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.

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*

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.

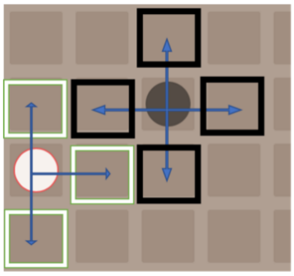


Fig 2. How to move pawn.

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 two sets of two squa res. 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*

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.

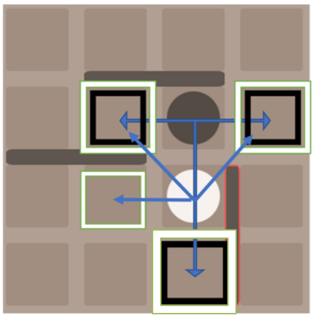


Fig 3. When two pawns are next to each other or when the path is blocked by a wall.

2.6. *End of the game*

The first player who reaches one of the five squares opposite his base line is the winner.

3. Retrograde Analysis

In this study, we conducted an experiment using receding 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 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".

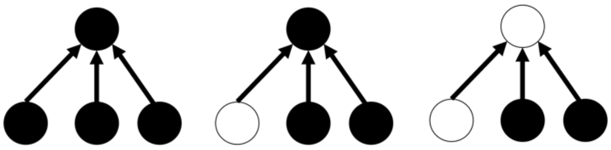


Fig 4. Retrograde Analysis

In Fig. 4, the circle represents a position of the game. The black circle and the white circle mean the black victory phase and the white victory phase, respectively. The lower circles represent the condition that can be transitioned from the upper circle.

In retrograde analysis of game, the result of the position is determined in the reverse order from the final position where the victory or defeat is decided to the initial position. For example, if the next move is only a black victory, that position is a black victory, see Fig.4.

4. Results

Some of the results of the regression analysis are shown in Fig. 5 which shows a game tree at a depth of 2 from the initial condition. "W" indicates that the victory of the first move is confirmed, and "L" indicates that the victory of the second move. The "P" part represents the movement of the pawn. The "F" part represents the placement of the fence. Pawn movements are divided into vertical movements and horizontal movements. Since there are many Fs, some are omitted.

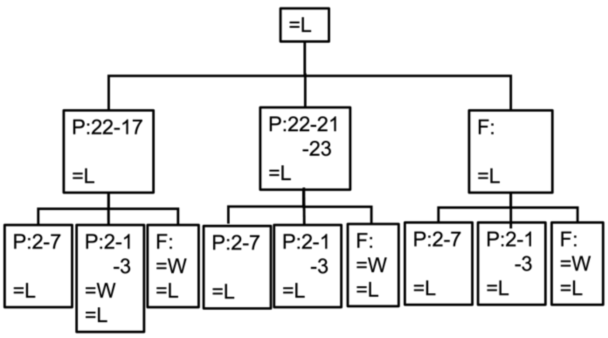


Fig 5. Some of the results of the regression analysis

For the position of the pawn in Fig. 5, the board number is specified as shown on the left in Fig. 6. Similarly, regarding the position of the fence, the position number is given to the fence as shown on the right in Fig. 6.

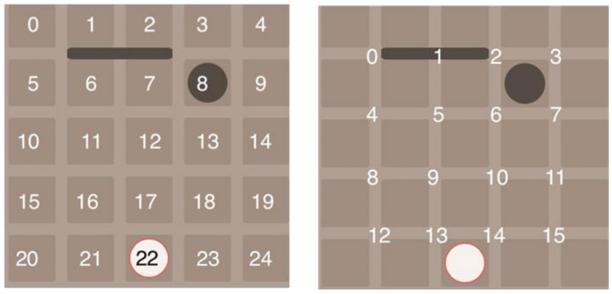


Fig. 6. Pawn and fence position number

Fig. 7 shows A perfect plays of Quoridor with a 5x5 board and one fence each. If the first move and the second move make the best move to each other, the first move must jump over or make a detour, so the first move loses by one move.

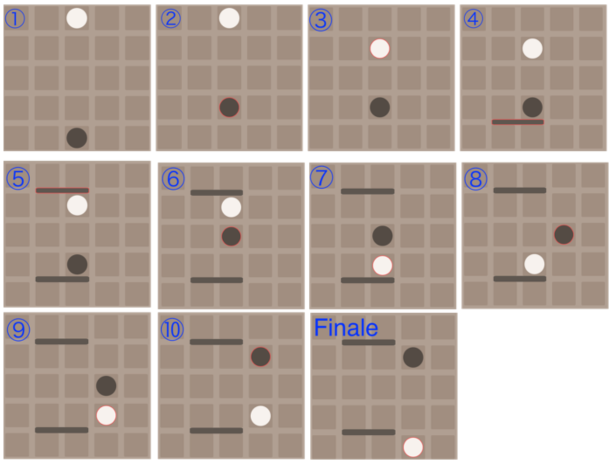


Fig. 7 A perfect plays of Quoridor with a 5x5 board and one fence each.

The number of positions obtained in this experiment was 298807 (There are $3.9905*10^{42}$ in the regular version[6].), excluding 768 positions not allowed by the rules. Furthermore, considering the turn of each position, the number is doubled to 597614. Only 20 of them were undecided.

5. Conclusion

After the experiment, I tried to play the 3x3 and 5x5 board with one fence each, and found that the game was won by the rear player in both cases. Since the only perfect play in this case is when the fence is placed horizontally in the center of the board where it can block the opponent's path, we can expect the same kind of moves in the 7x7 and 9x9 board as in the 5x5 board, resulting in a backward win. On the other hand, in 6x5 and 8x5, where the length of the rows is even, the game is expected to be won by the first move. I would like to confirm these two points.

References

1. Teturo Tanaka An Analysis of a Board Game “Doubutsu Shogi” Journal of Information Processing, Vol. 48, No. 11, pp. 3470–3476(2007).
2. J. Romein and H. Bal: Solving the Game of Awari using Parallel Retrograde Analysis, IEEE Computer, Vol. 36, No. 10, pp. 26 – 33(2003).
3. https://www.ultraboardgames.com/quoridor/game-rules.php
4. Schaeffer, J., N. Burch, Y. Bjornsson, A. Kishimoto, M. Mueller, R. Lake, P. Lu, and S. Sutphen, Checkers is solved, Science, 317, 1518–1522, 2007.
5. H. E. Bal and L. V. Allis. Parallel Retrograde Analysis on a Distributed System. In Supercomputing ’95, San Diego, CA, December 1995.
6. Victor Massague Respall, Joseph Alexander Brown and Hamna Aslam, "Monte carlo tree search for quoridor", *19th International Conference on Intelligent Games and Simulation GAME-ON 2018*, 2018.

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

He received the Ph.D. degree in computer science and systems engineering from Yamaguchi University in 1999. He is presently a professor in the Faculty of Engineering, University of Miyazaki. His first interests lay in hydrodynamics and time series analysis, especially the directional wave spectrum. He is a theoretical computer scientist, and 2. 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 forth. He has published many research papers in that area. His articles have appeared in journals such as Information Sciences, Pattern Recognition and Artificial Intelligence, WSEAS, AROB, ICAROB, SJI, IEEE and IEICE (Japan), and proceedings of conferences such as Parallel Image Processing and Analysis, SCI, WSEAS, AROB and ICAROB.

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.

A perfect play in 4×12 board of Othello

Tomoyasu Toshimori
Graduate School of Engineering, Miyazaki University, Japan, hm17033@student.miyazaki-u.ac.jp

Makoto Sakamoto
Department of Computer Science and System Engineering, Miyazaki University, Japan, saka2000@cc.miyazaki-u.ac.jp

Takao Ito
Graduate School of Engineering, Hiroshima University, Japan, itotakao@hiroshima-u.ac.jp

Satoshi Ikeda
Department of Computer Science and System Engineering, Miyazaki University, Japan, bisu@cs.miyazaki-u.ac.jp

Abstract

In 2015, Hiroki Takeshita started to analyze a reduced-form Othello board with a 4x6 board. Later, he analyzed the 4×8 and 4×10 boards and discovered perfect play. The analysis of 4×12 board Othello was sub-problemed by Takeshita and further sub-problemed by Toya Shotaro in 2019. In this paper, we present the solution to the remaining subproblem of 4 × 12 board Othello and the discovery of perfect play.

Keywords: combinatorial theory, combinatorial optimization, perfect play, miniature Othello.

1. Introduction

In 1993, the mathematician Joel Feinstein discovered that in a reduced form 6×6 Othello board, white wins when both players do their best, and announced this in the newsletter of the British Othello Association[1]. It was published in the newsletter of the Othello Society of Great Britain. The number of search phases reached about 40 billion, and the search period was described as two weeks. In 2015, in Hiroki Takeshita's senior thesis, he performed the same analysis of the 6×6 board Othello as Feinstein did and confirmed that Feinstein's result was correct. Since then, Takeshita has been working on the analysis of reduced board Othello and announced that he has succeeded in the complete analysis of 4×4, 4×6, 4×8, and 4×10 boards[2, 3].

The search space for the next 4×12 board was expected to be about 1000 times larger than the 4×10 board. Since it takes time to analyze such a huge number of phases, we tried to reduce the search space by referring to the perfect play of 4×10 board. The result of the search from the 7th moves onward, when the procedure up to the 6th move of the 4×12 board is the same as that of the 4×10 board, is that Black wins[4]. The result of the search after the 7th move was confirmed that Black won. In 2019, Shotaro Toya solved one of the subproblems.

In this study, we will analyze the subproblemized procedure and prove that Black wins the 4×12 board Othello. Since 13 out of 14 subproblems have already been analyzed by Takeshita, we will analyze the remaining subproblem. This subproblem has a larger search space than the other subproblems and takes more time to analyze, so it is further subdivided by Toya.

2. Othello*

The rules of Othello are as follows[5]. The two opponents are divided into two groups, one playing the white stones and the other the black stones, and place the stones as shown in Fig. 1. The player who chooses the black stone is the first to move. The opponent must place a stone in an empty square between the stones of his own color and the stones of his opponent's color on his turn. It can be placed vertically, horizontally, or diagonally. The stone of your opponent's color becomes a stone of your own color. After placing a stone, it is your opponent's turn to play. If there is no square that can hold a stone of the opponent's color, it is the opponent's turn without placing a stone (pass). The game ends when the board is completely filled with stones or when both opponents are unable to place any stones. At the end of the game, the side with more stones of its own color on the board is the winner.

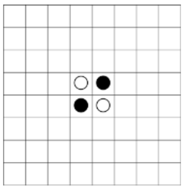


Fig. 1. Othello's setup.

Othello is classified as a two-player zero-sum finite definite perfect information game. As such, it can result in either a must-win first move, a must-win second move, or a draw[6].

3. Techniques

The program used for the analysis in this study is based on the negative-max-alpha-beta method with depth-first search. The negative-max method is an algorithm for finding the best order that exists in a game tree, with the same computational process for black and white. The best order in this research is the sequence of steps when you assume that you play the best move, and your opponent plays the worst move for you. The negative-max method requires that all nodes be explored. It can be used in conjunction with the alpha-beta method to reduce the computation time by not searching for moves that will never be selected.

4. Experiments

Hiroki Takeshita predicted that the perfect play of 4×12 board from the first move to the sixth move is identical to that of 4×8 and 4×10 boards. Fig. 2. shows the procedure of the perfect play with the 6th move fixed, and Fig. 3. shows the procedure from the first move to the 6th move.

The result of the perfect play with the 6th move fixed was "Black 42, White 0". If we can show the same best result before the sixth move, we can prove that the procedure including the fixed six moves is perfect play. There are 95 problems that are necessary to prove that the game is

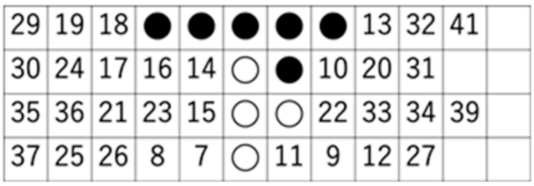


Fig. 2. A perfect play fixed from move 1 to move 6.

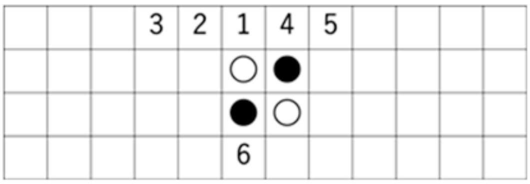


Fig. 3. A position fixed from move 1 to move 6.

played perfectly up to the 6th move as well. the 14 subproblems are shown in Table 1., and the board with the corresponding coordinates is shown in Fig. 4. Table 1. shows the 14 subproblems, and Fig. 4. shows the board with the corresponding coordinates. The blacked-out areas in Table 1. are problems that have already been analyzed.

*Othello is a registered trademark.

Table 1. Executing sub-problems to prove our perfect play.

Depth 6	Depth 4	Depth 2
f1e1d1g1h1e3	**f1e1d1e3f4g1**	**f1e3f4e1d1g1**
f1e1d1g1h1h2	**f1e1d1f4e2g1**	**f1e3f4e1d1h2**
	f1e1d1f4e2h1	**f1e3f4g1h1h2**
	f1e1d1h2h3f4	**f1g1h1e1d1e3**
	f1e1d1h2h3h4	**f1g1h1e1d1f4**
		f1g1h1e1d1h2
		f1g1h1e3g4h3

Of the 14 problems, those in Table 1. that are not blacked out are unsolved problems. Fig. 5. shows the board of an unsolved problem, where the X mark indicates a square where a black stone can be placed.

a1	b1	c1	d1	e1	f1	g1	h1	i1	j1	k1	l1
a2	b2	c2	d2	e2	○	●	h2	i2	j2	k2	l2
a3	b3	c3	d3	e3	●	○	h3	i3	j3	k3	l3
a4	b4	c4	d4	e4	f4	g4	h4	i4	j4	k4	l4

Fig. 4. Coordinates on 4×12 board.

5. Conclusion

As a result of analyzing the unsolved problem shown in Table 1., we found "Black 47 White 0". Since all problems in Table 1. were found to be perfect play, Takeshita's perfect play shown in Fig. 2. and Fig. 3. was proved. Fig. 7. shows the procedure of the perfect play discovered in this study.

					●	●	●				
				X	●	●		X			
				○	○	○	○				
			X	X	X	●	X	X			

Fig. 5. The board of the unsolved problem of 4×12 board Othello

In the 4×12 board Othello, it turned out that black wins as a result of both sides continuing to play their best moves. This is the same result as for the 4×6, 4×8, and 4×10 boards, so it is expected that black will also win in the 4×14 board.

●	●	●	●	●	●	●	●	●	●	●	●
●	●	●	●	●	●	●	●	●	●	●	●
●	●	●	●	●	●	●	●	●	●	●	●
●	●	●	●	●	●	●	●	●	●	●	

Fig. 6. Endgame board for perfect play of unsolved problems.

31	32	22	27	8	1	2	3	10	33	42	43
28	23	24	19	13	○	●	17	34	26	38	39
30	16	15	12	4	●	○	6	25	35	40	41
29	21	20	9	7	14	5	11	18	36	37	

Fig. 7. Procedure of perfect play for unsolved problems.

References

1. British Othello Federation: Forty Billion Nodes Under The Tree - The Newsletter of the British Othello Federation, pp. 6–8 (1993).
2. Yuki Takeshita, Satoshi Ikeda, Makoto Sakamoto and Takao Ito: "Perfect Analysis in miniature Othello," Proceedings of the 2015 International Conference on Artificial life and Robotics, pp.39 (2015).
3. Yuki Takeshita, Makoto Sakamoto, Takao Ito and Satoshi Ikeda: "Perfect Play in Miniature Othello," Proceedings of the Ninth International Conference on Genetic and Evolutionary Computing, pp.281-290 (2015).
4. Yuki Takeshita, Makoto Sakamoto, Takao Ito, Tsutomu Ito and Satoshi Ikeda: "Reduction of the search space to find perfect play of 6×6 board Othello," Proceedings of the 2017 International Conference on Artificial life and Robotics, pp.675-678 (2017)
5. Japan Othello Association: Nihon osero renmei kyougi ru-ru, [online]www.othello.gr.jp/r_info/rule.html (Retrieved 24 November. 2021)
6. Jonathan Schaeffer, Neil Burch, Yngvi Björnsson, Akihiro Kishimoto, Martin Müller, Robert Lake, Paul Lu and Steve Sutphen: "Checkers Is Solved," Science, Vol. 317, pp. 1518-1522 (2007).

Authors Introduction

Mr. Tomoyasu Toshimori

He is a student of the Department of Computer Science and Systems Engineering, Faculty of Engineering, University of Miyazaki.

Prof. Makoto Sakamoto

He received the Ph.D. degree in computer science and systems engineering from Yamaguchi University in 1999. He is presently a professor in the Faculty of Engineering, University of Miyazaki. His first interests lay in hydrodynamics and time series analysis, especially the directional wave spectrum. He is a theoretical computer scientist, and 2. 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 forth. He has published many research papers in that area. His articles have appeared in journals such as Information Sciences, Pattern Recognition and Artificial Intelligence, WSEAS, AROB, ICAROB, SJI, IEEE and IEICE (Japan), and proceedings of conferences such as Parallel Image Processing and Analysis, SCI, WSEAS, AROB and ICAROB.

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.

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

Amane Takei
Faculty of Engineering, University of Miyazaki,1-1, Gakuen Kibanadai-Nishi
*Miyazaki, 889-2192, Japan**

Kento Ohnaka
Faculty 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

E-mail:takei@cc.miyazaki-u.ac.jp, hl16004@student.miyazaki-u.ac.jp, saka2000@cc.miyazaki-u.ac.jp
http://www.miyazaki-u.ac.jp/

Abstract

In this presentation, a parallel microwave 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. Then, a numerical result by a microwave oven model is shown. The simplified Berenger's PML is introduced which these eight corners are given the average value of all PML's layers.

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 **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:

$$\iiint_{\Omega} (1/\mu)\mathrm{rot}\mathbf{E}_\mathrm{h} \cdot rot\mathbf{E}_\mathrm{h}{}^{*}dv - \omega^2 \iiint_{\Omega} \varepsilon\mathbf{E}_\mathrm{h} \cdot \mathbf{E}_\mathrm{h}{}^{*}dv = j\omega \iiint_{\Omega} \mathbf{J}_h \cdot \mathbf{E}_h{}^{*}dv. \quad (1)$$

The equation contains complex numbers and becomes a complex symmetric matrix. In the present study, the electric field **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).

$$\begin{bmatrix} K_{II}^{(1)} & 0 & 0 & K_{IB}^{(1)}R_B^{(1)\mathrm{T}} \\ 0 & \ddots & 0 & \vdots \\ & & K_{II}^{(N)} & K_{IB}^{(N)}R_B^{(N)\mathrm{T}} \\ 0 & 0 & & \\ R_B^{(1)}K_{IB}^{(1)\mathrm{T}} & \cdots & R_B^{(N)}K_{IB}^{(N)\mathrm{T}} & \sum_{i=1}^{\mathrm{N}} R_B^{(i)}K_{BB}^{(i)}R_B^{(i)\mathrm{T}} \end{bmatrix} \begin{bmatrix} u_I^{(1)} \\ \vdots \\ u_I^{(N)} \\ u_B \end{bmatrix} = \begin{bmatrix} f_I^{(1)} \\ \vdots \\ f_I^{(N)} \\ f_B \end{bmatrix} \quad (2)$$

$$K_{II}^{(i)}u_I^{(i)} = f_I^{(i)} - K_{IB}^{(i)}u_B^{(i)} \quad (i = 1, \ldots, N) \quad (3)$$

$$\left\{\sum_{i=1}^{\mathrm{N}} R_B^{(i)}\left\{K_{BB}^{(i)} - K_{IB}^{(i)\mathrm{T}}\left(K_{II}^{(i)}\right)^{-1}K_{IB}^{(i)}\right\}R_B^{(i)T}\right\}u_B = \sum_{i=1}^{\mathrm{N}} R_B^{(i)}\left\{f_B^{(i)} - K_{IB}^{(i)T}\left(K_{II}^{(i)}\right)^{-1}f_I^{(i)}\right\} \quad (4)$$

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 equation (5) as follows:

$$Su_B = g, \quad S = \sum_{i=1}^{N} R_B^{(i)}S^{(i)}R_B^{(i)T}, \quad S^{(i)} = K_{BB}^{(i)} - K_{IB}^{(i)T}\left(K_{II}^{(i)}\right)^{-1}K_{IB}^{(i)}. \quad (5)$$

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 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. Fig. 1 shows a schematic diagram of Berenger's PML absorbing boundary.

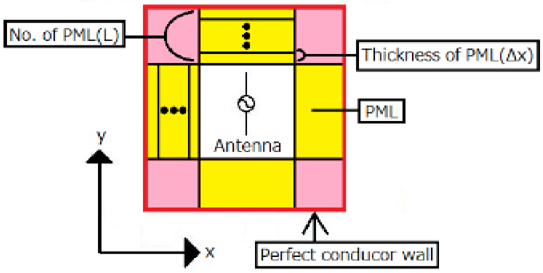


Fig. 1. PML absorbing boundary

In this paper the distribution of the electric conductivity for PML is expressed as follows:

$$\sigma = \sigma_{max}\left[\frac{\left(L - \hat{L}(x)\right)\Delta x}{L\Delta x}\right]^M \quad (6)$$

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:

$$|R(\phi)| \cong exp\left[-\frac{2\sigma_{max}L\Delta x}{(M+1)\varepsilon_0 c}\cos\phi\right] \quad (7)$$

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:

$$\sigma_{max} = -\frac{(M+1)\varepsilon_0 c}{2L\Delta x}\ln|R(0)| \quad (8)$$

In the present study, we construct a PML using (6) through (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.

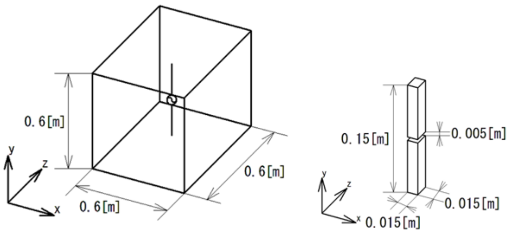
3.2. *Numerical results*

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:

$$I(y) = I_0\cos\left(\frac{2\pi}{\lambda}y\right) \quad : -l \le y \le l \quad (9)$$

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. Fig. 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. Fig. 7 shows a visualization diagram of the electric field obtained by analysis.

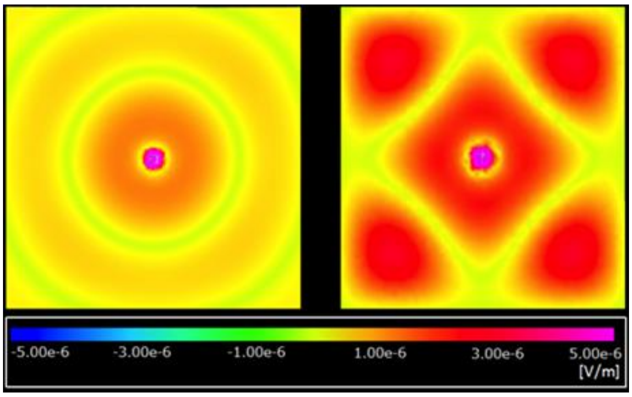


Fig. 3. **V**isualization of the analysis result (electric field) (Left: PML(9), Right: PML(0) (perfect conductor wall))

In Fig. 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 Fig. 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:

$$E_\theta = j60I\frac{e^{-jkr}}{r}\cdot\frac{\cos\left(\frac{\pi}{2}\cos\theta\right)}{\sin\theta} \qquad (10)$$

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. Fig. 4 shows a plot of the numerical analysis solution e_θ and the theoretical solution E_θ in increments of 1 [deg].

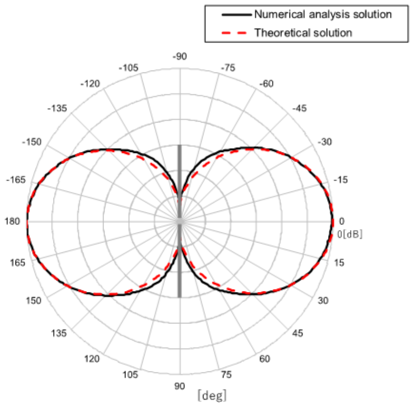


Fig. 4. Numerical and theoretical solutions in the E plane

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.

$$\arcsin\left(\frac{2\,l^2}{r\lambda}\right) + 90 \le \theta \le 90 - \arcsin\left(\frac{2\,l^2}{r\lambda}\right) \qquad (11)$$

$$E_{err} = \frac{\sum_{i=\theta_{Min}}^{\theta_{Max}} \frac{|e_i - E_i|}{E_i}}{\theta_{Max} - \theta_{Min} + 1} \times 100 \quad [\%] \qquad (12)$$

In the calculations shown in Fig. 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. Conclusion

In the present paper, we propose a simplified method that omits the directionality to Berenger's PML hyperbolic problems like a high-frequency electromagnetic field analysis and gives the average value of the electric conductivity of each layer at the corner of the model. Performance evaluation reveale that sufficient absorption performance can be obtained. In the accuracy verification by directivity evaluation of the dipole antenna, when the maximum element side length is set to 1/20 of the wavelength and the PML to be given is set to 9 layers, the error rate of the numerical solution and the theoretical solution is about 1.70 [%]. It is found that a highly accurate solution can be obtained. In addition, when the tolerance range of the far-field error rate that is considered to be sufficiently practical is set to less than 5 [%], an eight-layer PML is found to be optimal. In addition, the usefulness of the proposed method for a frequency band of 1.2 GHz or higher, which is used in microphones and mobile phones, is demonstrated.

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. TEAM Workshop, Tucson, AZ, USA, 1998.
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.
3. Fujikura LTD. Homepage

Authors Introduction

Prof. Amane Takei

He is working as Associate Professor for Department of Electrical and systems Engineering, University of Miyazaki, Japan. His research interest includes high performance computing for computational electromagnetism, iterative methods for the solution of sparse linear systems, domain decomposition methods for large-scale problems. Prof. Takei is a member of IEEE, an expert advisor of 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).

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

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.

Geometry Structure Oriented Nonlinear Internal Model Based Manifold Consensus

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, 30 Xueyuan 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
E-mails:songhpu@126.com, weicunzhang@263.net,daifz@tust.edu.cn, xiaohm@huel.edu.cn, smfei@seu.edu.cn

Abstract

This note comes with manifold consensus based on nonlinear internal model. To be special, scheme demonstrated here is not necessary to inject the nonlinear internal model with additional extraneous augmented system. And this amazing result is made possible empowered by geometry structure, to be specific, Riemannian metric is employed to modeling the internal model of the nonlinear manifold. In case of completeness, the consensus of a first order linear agent and another one second order oscillator is provided to verify the suggested program.

Keywords: geometry structure, manifold following, Riemannian metric, oscillator agent

1. Introduction

The consensus of the single simple agent and the complex dynamics of the nonlinear agent is full of meaning[1], the promising application of this research can be seen in different areas, especially in robotic fields[2-5]. Examples such as task cooperation, teams' agreement and such et al. The advantage of this heterogeneous

† Corresponding author

agents' deployment not only saves money but also sharpens the problem.

However, it is so tough to realize consensus between nonlinear agent and linear agent, like second oscillator agent in nonlinearity and one first order linear agent in linearity. Traditional method is to inject nonlinear internal model by introducing extraneous augmented system. It has several disadvantages, the first one is that it is not easy to found the exact nonlinear internal model, especially when some uncertainties exist, and another adverse point is that the traditional method needs the augmented dimension, the augmented system increases the dimensionality of the system, and this complicates the system further.

This note takes advantage of the geometry characteristics of the nonlinear agent, especially its Riemannian metric specialty to compact the nonlinear internal model. The compacted nonlinear internal model and the linear internal model is introduced in the control function at first, then followed by the high gain linear feedback controller. The jointed action of them makes consensus possible.

The note will be arranged as follows. At start, problem formulation will be proposed, then after, the main results will be presented in the second part. followed will be some conclusions.

2. Geometry Structure Oriented Nonlinear Internal Model

This part will begin with the introduction of the geometry structure oriented nonlinear internal model.

2.1. *Problem formulation*

Assume the following agents' model is already.

$$\begin{cases} \dot{X} = f(X) + h(X)U_X, X \in R^n, \\ \dot{Y} = AY + BU_Y, Y \in R^m, \\ n > m, \end{cases} \tag{1}$$

In Eq.1, we have two kinds of agents, one is the nonlinear agent, its state demonstrated as X, and another one is linear agent, its state is demonstrated as Y here. Our aim is to design controller to realize the consensus between X and Y.

Comment 1: It is not necessary to realize the consensus of all the sub variables of the linear agent to the nonlinear agent all the times, for this requirement is too strong to be practical. The accepted result is that some of the linear agent sub variables and some of the nonlinear agent sub variables come to consensus.

2.2. *The geometry structure*

Riemannian metric can be used to describe the nonlinear degree of the dynamics of the nonlinear agents.

Let *M* be a Riemannian manifold, *TM* is the tangent vector among *M*, the metric on *TM* is often denoted by $g = (g_p)_{p\in M}$ [6-7]. Under local coordinates, Riemannian metric is often noted as:

$$g = \sum\nolimits_{ij}^{`} g_{ij} dx_i \otimes dx_j$$

or simply

$$g = \sum\nolimits_{ij}^{`} g_{ij} dx_i dx_j \,,$$

where

$$g_{ij}(p) = \left\langle (\frac{\partial}{\partial x_i})_p, (\frac{\partial}{\partial x_p})_p \right\rangle_p \tag{2}$$

In straight and flat space, the Riemannian metric is reduced to Euclidean metric.

Comment 2: Take one circle in Euclidean space as example, its radius length is *r*, the value of *r* can demonstrate its nonlinear degree, the bigger the value of *r*, the smaller the nonlinear degree of the circle, the Riemannian metric can be characterized as its arc length.

2.3. *Nonlinear internal model*

Riemannian metric as the representation of the nonlinear degree of the dynamics of the nonlinear agents, and it can be used as the nonlinear internal model to construct the controller. Take Eq.1 as example, the controller designed under nonlinear internal model empowered by Riemannian metric can be factorized as three steps:

At the first step, the nonlinear internal model and linear internal model of the nonlinear agent and the linear agent is prepared, and at the second step the linear feedback law is added, and the third step is to pour them two together to the agent systems properly.

To be formalized, the controller for Eq.1 can be described as follows:

$$\begin{cases} U_X = [h(X)]^{-1} \left\lfloor g_{ij}(Y) - g_{ij}(X) + K(Y - X) \right\rfloor, \\ U_Y = B^{-1} \left[g_{ij}(X) - g_{ij}(Y) + K(X - Y) \right], \end{cases} \tag{3}$$

where that capital K is the coefficient matrix of the linear feedback controller, it reflects the feedback gain.

Comment 3: Here the nonlinear internal model is embedded in geometry metric, and the geometric metric is related to its local coordinate, that is geometric is localized quantity.

Comment 4: The regulation law of Equ.3 is composed of two different parts, one is from the nonlinear internal model representation, and another one is from the linear feedback controller.

Comment 5: The nonlinear internal model is responsible for driving the controlled system into the flat and straight space, and the linear feedback controller is in charge of driving the error of the consensus into the satisfied region.

3. Case Study

In this section, the case study of geometry metric oriented nonlinear internal model based consensus will be realized for an oscillator agent of the second order and a first order linear agent.

3.1. *The mathematical model*

The mathematical model of the targeted system is described as follows:

$$\begin{cases} \dot{X} = f(X) + h(X)U_X, X \in R^n, \\ \dot{Y} = AY + BU_Y, Y \in R^m, \\ f(X) = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}, \\ h(X) = 1, \\ X \in R^2, \\ A = 0, \\ B = 1, \\ Y \in R^1, \end{cases} \quad (4)$$

For oscillator agent without controller coming in, its locus is a circle and the radius of the circle is determined by the initial value of the agent. And its geometry metric is listed as follows:

$$\begin{cases} g_{ij}(X) = \frac{1}{2}\left|x_i x_j\right|, \\ g_{ij}(Y) = \mathrm{y}_i y_j, \end{cases} \quad (5)$$

Under this assumption, we can write down the regulation law of the controller to Equ.4 as follows:

$$\begin{cases} U_X = U_{Xa} + U_{Xb}, \\ U_{Xa} = y^2 - \frac{1}{2}\left|x_1 x_2\right|, \\ U_{Xb} = k(y - x_1), \\ U_Y = U_{Ya} + U_{Yb}, \\ U_{Ya} = \frac{1}{2}\left|x_1 x_2\right| - y^2, \\ U_{Yb} = k(x_1 - y), \end{cases} \quad (6)$$

Comment 6: Notice that abiding to the match principle, regulation action to oscillator is added on its first sub variable, for we want to realize the consensus among the first sub variable of the oscillator to the first order linear agent.

3.2. *Simulation results*

To oscillator agent and linear agent described in Eq.4, we select feedback gain *k=400*, and the initial values of the oscillator as -8, 2, and the initial value of the linear agent as 0, the simulation result is described in Fig.1 as follows:

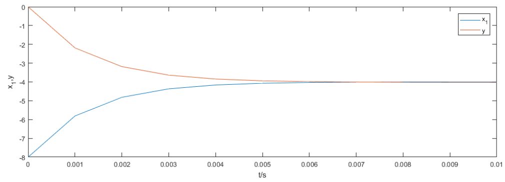


Fig.1 Case study example results

It is not so difficult for us to get that the consensus of the linear agent and the oscillator agent is made possible here.

3.3. *Further comments*

Inspiration initialized here can be generalized to the system of chaos, synchronization as well as robotics cooperation, it is just start of the suggested scheme, promising results are expected in the near future.

System complexity, could not be tractable if we are kept apart away from the concrete systems. Studies like performance improvement of vehicle driving[8-9], can be the main carrier. Time varying and distributive delays, can also be the negative factors. Adaptive strategy in Jia[10] gave a good example.

4. Conclusion

Nonlinear internal model is badly reputed in construction for its complexity, this note provides an alternative way to construct nonlinear internal model to realized consensus among oscillator agent and the linear agent, some promising results will be expected of this research in the near future.

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). Special thanks for Professor Yingmin Jia of Beihang University and Professor Masanori Sugisaka for their encouragement.

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, Z. Fu, F. Wang, Socialized multi-agent system rendezvous via networks of networks, *Journal of Robotics, Networking and Artificial Life*, **2** (2016) 136-139.
3. 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
4. Y. Song, Consensus of agents with mixed linear discrete dynamics, *International Journal of Control, Automation and Systems*, **14** (2016) 1139-1143.
5. 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
6. S. Chen, L. Zhao, W. Zhang, P. Shi, Consensus on compact Riemannian manifolds, Information Sciences, 268(2014) 220-230
7. A. Sarlette, R. Sepulchre, Consensus optimization on manifolds, *SIAM Journal on Control and Optimization*, **48**(2009) 56-76.
8. 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.
9. 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
10. 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.

Authors Introduction

Prof. Yunzhong Song
He received his PhD degree in Zhejiang University, China in 2006. He is currently a full professor in Henan Polytechnic University, China.

Prof. Weicun Zhang
He received his PhD degree in 1993 from Tsinghua University in China. He is currently a full professor in University of Science and Technology Beijing, China.

Dr. Fengzhi Dai
He received the Dr. Eng. from the Division of Materials Science and Production Engineering, Oita University, Japan, in 2004. His major field of study is robotics.

Prof. Huimin Xiao
He received his PhD degree in Automaic 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.

Prof. Shumin Fei
He received his PhD degree in Automaic Control Theory and Its Applications in 1995 from Beihang University in China. He is currently a full professor in Southeastern University, China.

Consensus Control of Linear Discrete-time Multi-agent Systems with Limited Communication Data Rate

Jintao Hu, Yingmin Jia*, Yaxin Li
The Seventh Research Division and the Center for Information and Control, School of Automation Science and Electrical Engineering, Beihang University (BUAA), 37 Xueyuan Road, Haidian District, Beijing, 100191, China
E-mail: hujintao@buaa.edu.cn, ymjia@buaa.edu.cn
www.buaa.edu.cn

Abstract

This paper investigates the consensus problem of linear discrete-time multi-agent systems with limited communication data rate and the cooperative-antagonistic interactions. A consensus control protocol is designed based upon a dynamic encoding-decoding mechanism. By means of the proposed control protocol, it is guaranteed that the agents can attain bipartite consensus if the signed undirected graph is connected and structurally balanced/unbalanced. Moreover, the clear form of the convergence rate is given.

Keywords: Multi-agent systems, signed undirected graph, bipartite consensus, structural balance, encoding and decoding

1. Introduction

Consensus is one of the hottest topics in the field of multi-agent systems (MASs) due to its wide application in lighting. Many important and interesting problems about consensus have been studied in recent years.

In many relevant works, it's assumed that the agents can achieve exact information of the states of the neighbors by local communication. However, in practice, the communication ability of the network is limited. Thus, the encoding and decoding scheme is introduced to deal with this problem[1-3].

In addition, most of the works in the literature assume that the relationship between agents is cooperative. However, cooperation often coexists with antagonism. By utilizing the property of structural balance/unbalance of communication networks, the bipartite consensus is introduced where all gents attain agreement concerning a value, which is same for all in modulus but not in sign[4-5].

Motivated by the above discussion, we consider a consensus control protocol based on a dynamic encoding-decoding mechanism. The agents can achieve bipartite consensus under the control scheme. What's more, the clear form of the convergence rate is given.

2. Preliminaries and Problem Formulation

2.1. *Signed Graph*

Let $\mathcal{G}(\mathcal{V},\mathcal{E},\mathcal{A})$ be a signed undirected graph, where $\mathcal{V}=\{v_1,v_2,\cdots,v_N\}$ is a set of N agents with i denoting the ith agent, $\mathcal{E}\subseteq\mathcal{V}\times\mathcal{V}$ is the edge set of agents, and $\mathcal{A}=[a_{ij}]\in\mathbb{R}^{N\times N}$ with the signed elements is the weighted adjacency matrix of $\mathcal{G}$. The set of agents who

*Corresponding author.

can communicate with agent i is denoted by $\mathcal{N}_i = \{j \in \mathcal{V} \mid (i,j) \in \mathcal{E}\}$. The Laplacian matrix of $\mathcal{G}$ is represented by $\mathcal{L} = \mathcal{C} - \mathcal{A}$, where $\mathcal{C} = diag(\sum_{j \in \mathcal{N}_1} |a_{1j}|, \cdots, \sum_{j \in \mathcal{N}_N} |a_{Nj}|)$. A signed graph $\mathcal{G}$ is structurally balanced if the nodes in $\mathcal{V}$ are divided into two parts $\{\mathcal{V}_1, \mathcal{V}_2\}$, where $\mathcal{V}_1 \bigcup \mathcal{V}_2 = \mathcal{V}$ and $\mathcal{V}_1 \bigcap \mathcal{V}_2 = \varnothing$, bring $a_{ij} \ge 0$ $\forall i, j \in \mathcal{V}_p \left(p \in \{1,2\}\right)$, $a_{ij} \le 0$ $\forall v_i \in \mathcal{V}_p$, $v_j \in \mathcal{V}_q$, $p \ne q$ $\left(p, q \in \{1,2\}\right)$, otherwise, the signed graph $\mathcal{G}$ is structurally unbalanced.

2.2. *Problem Formulation*

In this paper, we consider the bipartite consensus control for a multi-agent system with the discrete-time dynamics

$$x_i(k+1) = Ax_i(k) + Bu_i(k) \tag{1}$$

Where $x_i(k) \in \mathbb{R}^n$ and $u_i(k) \in \mathbb{R}$, respectively, represent the state and input of the agent i at time k. $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times 1}$ denote the state and input matrix.

A uniform quantizer is utilized to quantize the state of agent. The quantizer $q(x)$ is defined as follows:

$$q(x) = \begin{cases} 0, & -1/2 \le x \le 1/2 \\ b, & (2b-1)/2 < x \le (2b+1)/2, \\ & b = 1, \cdots, L-1 \\ L, & x > (2L-1)/2 \\ -q(-x), & x < -1/2 \end{cases} \tag{2}$$

Obviously, the level of the quantizer is $2L+1$. When $|x| \le L + 1/2$, the quantization error of the uniform quantizer satisfies $|x - q(x)| \le 1/2$. Assume that the information transmitted between agents only can be binary number, but not exact value. Thus, the corresponding encoder and decoder are designed. The encoder of agent i is proposed as follows:

$$\begin{cases} \hat{\phi}_i(k) = l(k-1)Z_i(k) + A\hat{\phi}_i(k-1), \hat{\phi}_i(0) = 0 \\ Z_i(k) = Q(\dfrac{x_i(k) - A\hat{\phi}_i(k-1)}{l(k-1)}) \end{cases} \tag{3}$$

Where $\hat{\phi}_i(k)$ and $Z_i(k)$ are the internal variable and output of the encoder, respectively. Apart from this, $Q(\cdot) = [q(\cdot), \cdots, q(\cdot)]^T$ is the product quantizer. $l(k) = l_0 r^k$ is the scaling function, where $l_0 \in \mathbb{R}$ and $r \in (0,1)$, which prevents the quantizer from saturation. The decoder of agent j which receive the information from agent i is designed as follows:

$$\begin{cases} \hat{\varphi}_{ji}(0) = 0 \\ \hat{\varphi}_{ji}(k) = l(k-1)Z_i(k) + A\hat{\varphi}_{ji}(k-1) \end{cases} \tag{4}$$

Where $\hat{\varphi}_{ji}(k)$ which is connected with agent i is the internal variable of the decoder of agent j.

The objective of this paper is to propose a distributed control scheme based on dynamic encoding and decoding, such that the MASs achieve bipartite consensus when communicate graph is structurally balanced/unbalanced.

For this purpose, we consider the following protocol

$$u_i(k) = K \sum_{j \in \mathcal{N}_i} |a_{ij}| (\hat{\phi}_i(k) - \text{sgn}(a_{ij}) \hat{\varphi}_{ij}(k)) \tag{5}$$

Where $K \in \mathbb{R}^{n \times 1}$ is the control gain and $\text{sgn}(\cdot)$ represents the signum function.

In order to facilitate the subsequent analysis, we give the following definition and assumption.

Definition 1. *The consensus of Multi-agent systems can be achieved, if there exists the control law such that*

$$\lim_{k \to \infty} \| x_i(k) - \text{sgn}(a_{ij}) x_j(k) \| = 0$$

For any initial state $x_i(0)$, $\forall i \in [1, N]$.

Assumption 1. $\|x_i(0)\|_\infty \le C_x, \forall i \in [1, N]$ *and* C_x *can be achieved by all agents.*

3. Main Results

3.1. *Structurally balanced graph*

To get the main result, we need divide the agents into two groups $d_i \in \{1, -1\}$ $i = 1, \cdots, N$, where the agents are in the same group when they are cooperating, i.e., $a_{ij} > 0$.

The relevant lemmas are given as follows:

Lemma 1. *For a connected signed and structurally balanced graph, there always exists a diagonal matrix* $D = diag(d_1, \cdots, d_N)$ *such that all elements of* $D\mathcal{A}D$ *is nonnegative, and* $0 = \lambda_1(\mathcal{L}) < \lambda_2(\mathcal{L}) \le \cdots \le \lambda_N(\mathcal{L})$.

Lemma 2. *For any* $P \in \mathbb{R}^{n \times n}$ *and* $\varepsilon > 0$, *we can obtain that* $\|P^k\| \le M\eta^k, \forall k \ge 0$, *where* $M = \sqrt{n}(1 + 2/\varepsilon)^{n-1}$ *and* $\eta = \rho(P) + \varepsilon\|P\|$.

Theorem 1. *Consider the multi-agent system (1) with the structurally balanced and connected signed undirected communication network. Suppose that* (A, B) *is controllable,* $\prod_j |\lambda_j(A)| < (1 + \lambda_2/\lambda_N)/(1 - \lambda_2/\lambda_N)$ *and Assumption 1 holds. Select appropriate control gain* K *such that* $\rho(J(K)) < 1$, *where* $J(K) = diag(A, A + \lambda_2 BK, \cdots, A + \lambda_N BK)$. *For any* $r \in (\rho(J(K)), 1)$ *and* $\sigma \in (0, (r - \rho(J(K)))/\|J(K)\|)$, *let*

$$M(K, r) = \frac{\|A\|_\infty + 2d^* \|BK\|_\infty}{2r} + \frac{n^{3/2} \lambda_N^2 M \sqrt{N} \|BK\|_\infty^2}{r^2(1-\eta)}$$

and for any $L > M(K, r) - 1/2$, *let*

$$l_0 > \max\{\frac{\|A\|_\infty C_x}{L+1/2}, \frac{4r^2 C_x(1-\eta)}{\sqrt{n}\lambda_N \|BK\|_\infty}\}$$

Then under the proposed control scheme (5) given by (2), (3) and (4), the closed-loop system satisfies the bipartite average consensus

$$\lim_{k\to\infty} x_i(k) = d_i \sum_{j=1}^N d_j x_j(0),\ i = 1,\cdots,N$$

Proof. Let the signed average state error be $\delta_i(k) \triangleq x_i(k) - d_i x_{ave}(k)$, where $x_{ave}(k) \triangleq 1/N \sum_{i=1}^N d_i x_i(k)$ is the signed average state of the MASs. It yields that

$$\lim_{k\to\infty} \|\delta_i(k)\| \le 1/N \sum_{j=1}^N \lim_{k\to\infty} \|x_i(k) - \operatorname{sgn}(a_{ij}) x_j(k)\| \quad (6)$$

Simultaneously, $\lim_{k\to\infty} \|\delta_i(k)\| = 0$ also can implies the consensus of the MAS. Thus, the bipartite consensus of structurally balanced network is equivalent to $\lim_{k\to\infty} \|\delta_i(k)\| = 0$, for all $i \in [1, N]$.

According to the encoder and decoder, it can be achieved that $\hat{\varphi}_{ij}(k) \equiv \hat{\phi}_j(k)$, $j \in \mathcal{N}_i$, $\forall i \in [1, N]$ easily. Let the estimation error be $e_i(k) = x_i(k) - \hat{\phi}_i(k)$. Inserting the controller into the linear discrete system, we have

$$x_i(k+1) = Ax_i(k) + BK \sum_{j=1}^N |a_{ij}| (x_i(k) - \operatorname{sgn}(a_{ij}) x_j(k))$$
$$-BK \sum_{j=1}^N |a_{ij}| (e_i(k) - \operatorname{sgn}(a_{ij}) e_j(k)) \quad (7)$$

Choose $X(k) = [x_1^T(k), \cdots, x_N^T(k)]^T$, $E(k) = [e_1^T(k), \cdots, e_N^T(k)]^T$ and $\Delta(k) = [\delta_1^T(k), \cdots, \delta_N^T(k)]^T$. Then

$$X(k+1) = (I_N \otimes A + \mathcal{L} \otimes BK) X(K) - (\mathcal{L} \otimes BK) E(K) \quad (8)$$
$$\Delta(K+1) = (I_N \otimes A + \mathcal{L} \otimes BK) \Delta(K) - (\mathcal{L} \otimes BK) E(K) \quad (9)$$

Denote $\Phi(k) = [\phi_1^T(k), \cdots, \phi_N^T(k)]^T$. Then

$$X(k+1) - (I_N \otimes A)\Phi(k)$$
$$= (I_N \otimes A - \mathcal{L} \otimes BK) E(k) + (\mathcal{L} \otimes BK) \Delta(k) \quad (10)$$

This together with (3) leads to

$$E(k+1) = (I_N \otimes A - \mathcal{L} \otimes BK) E(k) + (\mathcal{L} \otimes BK) \Delta(k)$$
$$-l(k) Q(((I_N \otimes A - \mathcal{L} \otimes BK) E(k) + (\mathcal{L} \otimes BK) \Delta(k)) / l(k)) \quad (11)$$

Let $R(k) = \Delta(k)/l(k)$ and $S(k) = E(k)/l(k)$ .Next, we will prove that the quantizer is never saturated.

From Assumption 1, we can see that

$$\|(I_N \otimes A - \mathcal{L} \otimes BK) S(0) + (\mathcal{L} \otimes BK) R(0)\|_\infty$$
$$\le \|A\|_\infty C_x / l_0 < L + 1/2 \quad (12)$$

Thus, the quantizer is unsaturated at time $k = 0$. In order to demonstrate the feasibility of the quantizer, mathematical induction is utilized. Assume that the quantizer is unsaturated for time $[0, k]$, i.e., $\sup_{1\le i\le k+1} \|S(i)\|_\infty \le 1/2r$.

Choose $\xi_i \in \mathbb{R}^N$ such that $\xi_i^T \mathcal{L} = \lambda_i \xi_i^T$. The unitary matrix $\Xi = [D1_N/\sqrt{N}, \xi_2, \cdots, \xi_N]$ is introduced such that $\Xi^T \mathcal{L} \Xi = diag(0, \lambda_2, \cdots, \lambda_N)$. It is easy to derive that

$$(\Xi \otimes I_n)^T (I_N \otimes A + \mathcal{L} \otimes BK)(\Xi \otimes I_n) = J(K) \quad (13)$$

Let $\tilde{R}(k) = (\Xi \otimes I_n)^T R(k)$ and $\tilde{S}(k) = (\Xi \otimes I_n)^T (\mathcal{L} \otimes BK) S(k)$, then

$$\tilde{R}(k+1) = r^{-1} J(K) \tilde{R}(k) - r^{-1} \tilde{S}(k) \quad (14)$$

Separate $\tilde{R}(k)$ into two parts $[\tilde{R}_1^T(k), \tilde{R}_2^T(k)]^T$, where $\tilde{R}_1(k) = \mathbf{0}_n$ represents the first n elements of $\tilde{R}(k)$. Similarly, $\tilde{S}(k)$ is in two parts $[\tilde{S}_1^T(k), \tilde{S}_2^T(k)]^T$. Then

$$\tilde{R}_2(k+1) = \left(\frac{J(K)}{r}\right)^{k+1} \tilde{R}_2(0) - \frac{1}{r} \sum_{i=0}^k \left(\frac{J(K)}{r}\right)^{k-i} \tilde{S}_2(i) \quad (15)$$

Combine Lemma 2 and the condition given in Theorem 1, we can get that $\|(J(K)/r)^k\| \le M\eta^k$, where $M = \sqrt{n(N-1)}(1+2/\varepsilon)^{n(N-1)-1}$ and $\eta = \rho(J(K)/r) + \varepsilon \|J(K)/r\| < 1$. This together with (15) leads to

$$\|\tilde{R}_2(k+1)\| \le M\eta^{k+1} \|\tilde{R}_2(0)\| + (1/r) M \sum_{i=0}^k \eta^{k-i} \|\tilde{S}_2(i)\| \quad (16)$$

And given $\|\tilde{R}_2(0)\| = \|R(0)\| \le 2\sqrt{Nn} C_x / l_0$, we have

$$(1/r) M \sum_{i=0}^k \eta^{k-i} \|\tilde{S}_2(i)\|$$
$$\le n\sqrt{N} \lambda_N M \|BK\|_\infty (1 - \eta^{k+1}) / (2r^2(1-\eta)) \quad (17)$$

Since $\eta \in (0,1)$, (16) can be rewrote as

$$\|\tilde{R}_2(k+1)\| < \max\{\frac{2M\sqrt{Nn} C_x}{l_0}, \frac{n\sqrt{N}\lambda_N M \|BK\|_\infty}{2r^2(1-\eta)}\} \quad (18)$$

Noting that $l_0 > 4r^2 C_x (1-\eta) / (\sqrt{n} \lambda_N \|BK\|_\infty)$, we have

$$\|(I_N \otimes A - \mathcal{L} \otimes BK) S(k+1) + (\mathcal{L} \otimes BK) R(k+1)\|_\infty$$
$$< L + 1/2 \quad (19)$$

Thus, the $2L+1$ level uniform quantizer will not be saturated for time $k \ge 0$.

Considering $\|R(0)\|_\infty \le 2C_x / l_0$, then

$$\sup_{k\ge 0} \|R(k)\|_\infty < \max\{\frac{2C_x}{l_0}, \frac{n\sqrt{N}\lambda_N M \|BK\|_\infty}{2r^2(1-\eta)}\} < \infty \quad (20)$$

According to the definition of $R(k)$ and $0 < r < 1$, it implies that $\lim_{k\to\infty} \|\Delta(k)\|_\infty = 0$, so the multi-agent system (1) can realize the bipartite average consensus. □

Define the bipartite average consensus convergence rate as $r_{conv} = \sup_{\Delta(0)\ne 0} \lim_{k\to\infty} (\|\Delta(k)\| / \|\Delta(0)\|)^{1/k}$. From (16) and (17), for $\forall \Delta(0) \ne 0$, we get

$$\frac{\|\Delta(k+1)\|}{\|\Delta(0)\|} \le M(r\eta)^{k+1} + \frac{l_0 n\sqrt{N}\lambda_N M \|BK\|_\infty}{2r(1-\eta)\|\Delta(0)\|} r^k \quad (21)$$

Take the natural logarithm on both sides

$$\ln\left(\frac{\|\Delta(k+1)\|}{\|\Delta(0)\|}\right) \le \ln\left(\frac{l_0 n\sqrt{N}\lambda_N M \|BK\|_\infty}{2r(1-\eta)\|\Delta(0)\|} r^k\right) + \ln(1 + o(1))$$
$$= k\ln r + O(1), k \to \infty \quad (22)$$

Thus,

$$\lim_{k\to\infty}\left(\frac{\|\Delta(k+1)\|}{\|\Delta(0)\|}\right)^{1/(k+1)} \le \exp\left(\lim_{k\to\infty}\frac{1}{k+1}k\ln r + O(1)\right) = r \quad (23)$$

i.e., the bipartite average consensus convergence rate satisfies $r_{conv} \le r$.

3.2. *Structurally unbalanced graph*

The main result in this subsection is similar to Theorem 1. Please refer to theorem 1 for specific proof.

4. Conclusion

The consensus control problem of multi-agent systems with limited communication rate based on the signed connected undirected network topology is solved in this paper. For structurally balanced and unbalanced topology, the bipartite average consensus can be achieved by employing the proposed control scheme, respectively. In addition, the bipartite average consensus convergence rate is given explicitly.

Acknowledgements

This work was supported by the NSFC (62133001, 61520106010) and the National Basic Research Program of China (973 Program: 2012CB821201).

References

1. T. Li, M. Fu, L. Xie, J. Zhang: Distributed Consensus with Limited Communication Data Rate, IEEE Transactions on Automatic Control, vol. 56, No.2, pp.279-292, 2011.
2. K. You, L. Xie: Network Topology and Communication Data Rate for Consensusability of Discrete-Time Multi-agent Systems, IEEE Transactions on Automatic Control, vol. 56, No.10, pp.2262-2275, 2011.
3. Y. Meng, T. Li, J. Zhang: Coordination over Multi-agent Networks with Unmeasurablr States and Finite-Level Quantization, IEEE Transactions on Automatic Control, vol. 62, No.9, pp.4647-4653, 2011.
4. C. Altafini: Consensus Problems on Networks with Antagonistic Interactions, IEEE Transactions on Automatic Control, vol. 58, No.4, pp.935-946, 2013.
5. Y. Zhu, S. Li, J. Ma, Y. Zheng: Bipartite Consensus in Networks of Agents with Antagonistic Interactions and Quantization, IEEE Transactions on Circuits and Systems, vol. 65, No.12, pp.2012-2016, 2018.

Authors Introduction

Mr. Jintao Hu

He received the B.S. degree in automation in 2019 from Chongqing University in China. He is currently pursuing the Ph.D. degree with the Seventh Research Division and the Center for Information and Control, School of Automation Science and Electrical Engineering, Beihang University.

Prof. Yingmin Jia

He received the B.S. degree in control theory from Shandong University, China, in 1982, and the M.S. and Ph.D. 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. From February 1995 until February 1996 he was a Visiting Professor with the Institute of Robotics and Mechatronics of the German Aerospace Center (DLR), Oberpfaffenhofen, Germany. He held an Alexander von Humboldt (AvH) research fellowship with the Institute of Control Engineering at the Technical University Hamburg-Harburg, Hamburg, Germany, from December 1996 until March 1998, and a JSPS research fellowship with the Department of Electrical and Electronic Systems at the Osaka Prefecture University, Osaka, Japan, from March 2000 until March 2002. He was a Visiting Professor with the Department of Statistics at the University of California Berkeley from December 2006 until March 2007. Dr. Jia was the recipient of the National Science Fund for Distinguished Young Scholars in 1996, and was appointed as Chang Jiang Scholar of the Ministry of Education of China in 2004. He has been Chief Scientist of the National Basic Research Program of China (973 Program) since 2011, and in particular he won the Second Prize of National Technology Invention Award in 2015. His current research interests include robust control, adaptive control and intelligent control, and their applications in robots systems and distributed parameter systems.

Mr. Yaxin Li

He received the B.S. degree in automation from Shandong University in 2019. He is currently pursuing the Ph.D. degree with the Seventh Re-search Division and the Center for Information and Control, School of Automation Science and Electrical Engineering, Beihang University.

Formation control for rectangular agents with communication maintenance and collision avoidance

Yaxin Li, Yingmin Jia*, Jintao Hu

The Seventh Research Division and the Center for Information and Control, School of Automation Science and Electrical Engineering, Beihang University (BUAA), 37 Xueyuan Road, Haidian District, Beijing, 100191, China

E-mail: lyx2018@buaa.edu.cn, ymjia@buaa.edu.cn

www.buaa.edu.cn

Abstract

This paper mainly discusses the rectangular agents, which is not often considered in other papers. Coordinate transformation is used to describe the location relationship between two agents. Obstacle potential function between agents, obstacle function between agents and obstacles are designed to ensure no collisions. The distributed control protocol is designed to achieve desired formation with no collisions and communication maintenance between agents. Stability analysis proves the effectiveness of the algorithm.

Keywords: rectangular agents, formation control, communication maintenance, collision avoidance

1. Introduction

With the progress of science and technology, especially the unprecedented development of robots, sensors, artificial intelligence, the coordinated control system of multi-agent system has also become a research hotspot in the field of automatic control[1-3].

In the application scenario of coordinated control of multi-agent system, due to the limited communication and measurement range of a single agent, it is necessary to use multi-agent system to form a specific formation to achieve the maximum coverage and detection, such as seabed exploration and disaster search and rescue or in some military exercises and operations. Due to the rich application prospect and the challenge of group coordinated control, formation control has become one of the multi-agents coordinated control[4-6].

In the process of formation or other tasks, designing obstacle avoidance strategies between agents and between agents and environmental obstacles is the prerequisite for the safe and reliable operation of multi-agent system. Artificial potential function is often used to solve formation problems[7]. Formation and obstacle avoidance of first-order rectangular multi-agent systems have been studied[8]. Formation Maneuvering with collision avoidance and communication maintenance is also studied[9].

This article mainly solves two problems: For second-order rectangular multi-agents system

(i). How to design the control law to make the system reach the desired formation and avoid collision?

(ii). How to design the control law to make the system reach the desired formation, avoid collision and maintain communication?

Notation. *The length of the rectangle agent i is l_i, the width of the rectangle agent i is w_i. L_s denotes the Laplacian matrix of sensing graph. L_c denotes the Laplacian matrix of communication graph.*

2. Preliminaries and Problem Statement

2.1. *Graph theory*

A directed graph $G(v,\varepsilon)$ consists of a node set $v=\{1,2,\cdots n\}$ and a set of edges $\varepsilon \subset v\times v$. If the edge (j,i) is in edge ε, *j* is the in-neighbor of agent *i* and *i* is the out-neighbor of agent *j*.
For a directed graph, if there exists a path from node q_2 to node q_1, we say that node q_1 is reachable from node q_2. If there exists a node q_1, which is reachable from any other node q_i, we say the directed graph is rooted. Supposed existing a non-singleton subset of nodes Q:{Q_1, Q_2, …, Q_n}, if the exists a path from nodes {Q_1, Q_2, …, Q_n} to q_1 after removing any node in Q, we say q_1 is 2-reachable from Q. If there exists a subset of two nodes, which is 2-reachable from any other node, we say the directed graph is 2-rooted.

2.2. *Coordinate transformation*

Consider two rectangle agents 1 and 2 and one obstacle (obs) as shown in Fig. 1, where the coordinate of center of rectangle 1 (O_1) is denoted by (x_1, y_1) and vertex P_{11} (x_{11}, y_{11}) is denoted by P_{11} in the frame OXY, the coordinate of the center of rectangle 1 ($O^1{}_1$) is denoted by ($x^1{}_1$, $y^1{}_1$) and vertex P_{11} ($x^1{}_{11}$, $y^1{}_{11}$) is denoted by $P^1{}_{11}$ in the frame $O_1X_1Y_1$.
Denote $P_1=\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, P_{1k}=\begin{bmatrix} x_{1k} \\ y_{1k} \end{bmatrix}$ in the frame OXY, and

$P^1{}_{1k}=\begin{bmatrix} \overset{1}{x}_{1k} \\ \overset{1}{y}_{1k} \end{bmatrix}$ in the frame $O_1X_1Y_1$ for k=1,2,3,4,

$$\overset{1}{P}_{1k}=R(\varphi_1)(P_{1k}-P_1) \tag{1}$$

where $R(\varphi_1)=\begin{bmatrix} \cos(\varphi_1) & \sin(\varphi_1) \\ -\sin(\varphi_1) & \cos(\varphi_1) \end{bmatrix}$

The distance between two agents is defined as follows.

$$d_{x1}(p_{2k})=\begin{cases} \left|\overset{1}{x}_{2k}\right|-\dfrac{l_1}{2}, & if \left|\overset{1}{x}_{2k}\right|>\dfrac{l_1}{2} \\ 0, & otherwise \end{cases}$$

$$d_{y1}(p_{2k})=\begin{cases} \left|\overset{1}{y}_{2k}\right|-\dfrac{w_1}{2}, & if \left|\overset{1}{y}_{2k}\right|>\dfrac{w_1}{2} \\ 0, & otherwise \end{cases}$$

$$d_1(p_{2k})=\sqrt{d_{x1}(p_{2k})^2+d_{y1}(p_{2k})^2}$$

We define the distance between of agent 1 and agent 2 is $d_{12}=\min\left(\min_k\left(d_1(p_{2k})\right),\min_k\left(d_2(p_{1k})\right)\right)$.

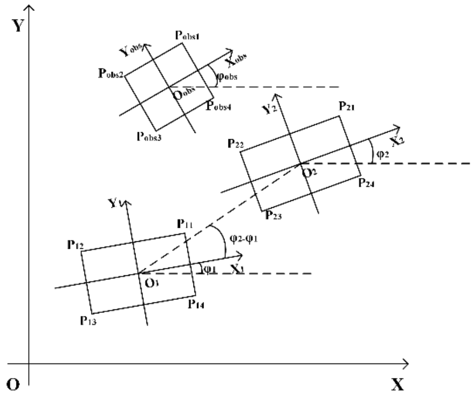


Fig. 1. Coordinate transformation between agents and obstacles

2.3. *Collision and connection region*

We define two different areas for obstacle avoidance (Ψ_i^a) and maintaining communication state (Ψ_i^c).
$\Psi_i^a=\left\{ j\in\mathbb{C}: r_a\le\left\|d_{ij}\right\|\le R_a\right\}$ is safe range of obstacle avoidance. If $\left\|d_{ij}\right\|\le r_a$, obstacle avoidance may occur.
$\Psi_i^a=\left\{ j\in\mathbb{C}: r_m\le\left\|d_{ij}\right\|\le R_m\right\}$ is range of maintaining communication state. If $\left\|d_{ij}\right\|\ge R_m$, agents will lose contact.

2.4. *Desired formation*

Given a desired relative position formation β: $\{0,\beta_2,\cdots,\beta_n\},\beta_i\in\mathbb{C}$, if the final position can be expressed as $\lim_{t\to\infty}p(t)=\varpi_1\beta+\varpi_2\mathbf{1_n}+v_n t1_n,\ \varpi_1,\varpi_2\in\mathbb{C}$, We say the target formation is formed.

3. Control Design

We consider a system with n agents, and the dynamic model is as follows:

$$\begin{cases} \dot{p}_i=v_i \\ \dot{v}_i=a_i \end{cases} \tag{2}$$

where $p\in\mathbb{C}$ donates the position, $v\in\mathbb{C}$ donates the velocity, $a\in\mathbb{C}$ denotes the acceleration.
Now we consider the problem (i).
The control protocol is designed as follows:

$$u_i=u_{if}+u_{ic}, i=1,\ldots,n \tag{3}$$

where u_{if} is for formation maneuvering and u_{ic} is for collision avoidance.
u_{if} is obtained from the following equation:

$$\begin{cases} \dot{\tilde{v}}_i = \sum\limits_{j\in M_i} a_{ij}\left(\tilde{v}_j - \tilde{v}_i\right) \\ \dot{\tilde{p}}_i = -\sum\limits_{j\in \mathcal{N}_i^+} s_{ij}\left(p_j - p_i\right) - \alpha\tilde{p}_i \\ u_{if} = -\sum\limits_{j\in \mathcal{N}_i^+} s_{ij}^*\tilde{p}_i + \sum\limits_{j\in \mathcal{N}_i^-} s_{ji}^*\tilde{p}_j + \tilde{v}_i - v_i \end{cases} \tag{4}$$

where $\tilde{p}_i$ is an auxiliary variable which is used to achieve the desired formation, $\tilde{v}_i$ is an auxiliary variable which is used to make the velocity synchronize. M_i is the neighbor of agent i in communication graph, N_i^+ is the in-neighbor of agent i in sensing graph, N_i^- is the out-neighbor of agent i in communication graph. a_{ij} can be any positive real number, $s_{ij} \in \mathbb{C}$ satisfies $\sum\limits_{j\in \mathcal{N}_i^+} s_{ij}(\zeta_j - \zeta_i) = 0,\ for\ i = 1,\ldots,n$.
The value of variable α satisfies

$$\alpha > \frac{\sqrt{1+4\lambda_{\max}\left(L_s^* L_s\right)} - 1}{2}$$

Proof.
Donate

$p = \left[p_1, p_2, \cdots p_n\right]^T, v = \left[v_1, v_2, \cdots v_n\right]^T, \tilde{p} = \left[\tilde{p}_1, \tilde{p}_2, \cdots \tilde{p}_n\right]^T,$

$$v = \left[\tilde{v}_1, \tilde{v}_2, \cdots \tilde{v}_n\right]^T, K_1 = \begin{bmatrix} 0 & I_n & 0 \\ 0 & -I_n & -L_s^* \\ L_s & 0 & -\alpha I_n \end{bmatrix}, K_2 = \begin{bmatrix} 0 \\ I_n \\ 0 \end{bmatrix},$$

$$\sigma_1 = p - \bar{v}t\mathbf{1}_n, \sigma_2 = v - \bar{v}\mathbf{1}_n, \sigma_3 = \tilde{v} - \bar{v}\mathbf{1}_n, \bar{v} = \frac{c^T\tilde{v}(0)}{c^T\mathbf{1}_n}$$

Then the system can be described as:

$$\begin{bmatrix} \dot{p} \\ \dot{v} \\ \dot{\tilde{p}} \\ \dot{\tilde{v}} \end{bmatrix} = \begin{bmatrix} 0 & I_n & 0 & 0 \\ 0 & -I_n & -L_s^* & I_n \\ L_s & 0 & -\alpha I_n & 0 \\ 0 & 0 & 0 & L_c \end{bmatrix} \begin{bmatrix} p \\ v \\ \tilde{p} \\ \tilde{v} \end{bmatrix}$$

We transform this system into a cascade system

$$\boxed{\dot{\tilde{v}} = -L_c\tilde{v}} \xrightarrow{+} \overset{\downarrow -}{\otimes} \rightarrow \boxed{\begin{bmatrix} \dot{p} \\ \dot{v} \\ \dot{\tilde{p}} \end{bmatrix} = K_1 \begin{bmatrix} p \\ v \\ \tilde{p} \end{bmatrix} + K_2\tilde{v}}$$

By analyzing the asymptotic convergence of the zero input system, we can solve the range of α.
We design a potential function Θ_{ij} :

$$\Theta_{ij} = \begin{cases} \left(\dfrac{\left|p_i - p_j\right|^2}{\sqrt{\tan\mu_1\left(\left|p_i - p_j\right|^2 - r_a\right)}}\right)^{1.4}, & \text{if } p_j \in \Psi_i^a \\ 0, & \text{otherwise} \end{cases}$$

where μ_1 is a constant.

$$\vartheta_i = -\sum_{p_j\in\Psi_i^a} \frac{\partial\Theta_{ij}}{\partial p_i},\ i = 1,\ldots,n \quad u_{ic} = \left(\left|u_{if}\right| + \mu_3\right)\mathrm{sgn}\left(\vartheta_i\right) \tag{5}$$

where μ_3 is a constant. □

Theorem 1. *Using control protocol in Eq. (3)(4)(5), the system can achieve the desired formation and avoid collision.*

Proof[7]. The proof is similar to Theorem 3.1. □
Now we consider the problem (ii).
The control protocol is designed as follows:

$$u_i = u_{if} + u_{ic} + u_{ia}, i = 1,\ldots,n \tag{6}$$

We design a potential function Λ_{ij} :

$$\Lambda_{ij} = \begin{cases} \left(\sqrt{\tan\mu_2\left(\left|p_i - p_j\right|^2 - r_m^2\right)}\right)^2, & \text{if } p_j \in \Psi_i^c \\ 0, & \text{otherwise} \end{cases}$$

where μ_2 is a constant.

$$\zeta_i = -\sum_{p_j\in\Psi_i^a} \frac{\partial\Lambda_{ij}}{\partial p_i},\ i = 1,\ldots,n \quad u_{ia} = \left(\left|u_{if}\right| + \mu_4\right)\mathrm{sgn}\left(\zeta_i\right) \tag{7}$$

where μ_4 is a constant.

Theorem 2. *Using control protocol in Eq. (4)(6)(7), the system can achieve the desired formation, avoid collision and maintain.*

Proof. The proof is similar to Theorem 3.1. □

4. Conclusion

This paper solves two problems for second-order rectangle agents: communication maintenance and collision avoidance. We use the coordinate transformation of different coordinate systems to obtain the distance information between agents. We design the potential function to keep the agents at a corresponding distance from each other. The corresponding control law is designed to achieve the desired goal.

Acknowledgements

This work was supported by the NSFC (62133001, 61520106010) and the National Basic Research Program of China (973 Program: 2012CB821201).

References

1. P. Ogren, E. Florelli, N. E. Leonard. Cooperative control of mobile sensor networks : Adaptive gradient climbing in a distributed environment. IEEE Transactions on Automatic Control, vol. 49, No. 8, pp. 1292-1302, August, 2004.
2. X. Wang, Z. Zeng, Y. Cong. Multi-agent distributed coordination control: Developments and directions via graph viewpoint. Neurocomputing, vol.199, pp. 204-218, July, 2016.
3. W. Ren. Distributed Consensus Algorithms and Their Applications in Multi-vehicle Cooperative Control. 2007 International Conference on Mechatronics and Automation, 2007.
4. X. Lu, S. Chen, F. Austin. Formation control for second-order multi-agent systems with time-varying delays under directed topology. Seventh International Conference on Natural Computation, August 2011.
5. K. K. Oh, H. S. Ahn. Formation Control of Mobile Agents Based on Distributed Position Estimation., IEEE Transactions on Automatic Control, Vol.58, No. 3, pp.737-742, March, 2013.
6. F. Chen, D. V. Dimarogonas. Leader–Follower Formation Control With Prescribed Performance Guarantees. IEEE Transactions on Control of Network Systems, vol. 8, no. 1, pp. 450-461, March 2021.
7. Urakubo, Takateru. Feedback stabilization of a nonholonomic system with potential fields: application to a two-wheeled mobile robot among obstacles. Nonlinear Dynamics, vol. 81, no. 3, pp.1478-1487, 2015.
8. T. Nguyen, H. M. La, T. D. Le and M. Jafari. Formation Control and Obstacle Avoidance of Multiple Rectangular Agents With Limited Communication Ranges. IEEE Transactions on Control of Network Systems, vol. 4, no. 4, pp. 680-691, Dec. 2017.
9. Z. Han, Z. Lin, Z. Chen and M. Fu. Formation maneuvering with collision avoidance and connectivity maintenance. 2015 IEEE International Symposium on Intelligent Control (ISIC), pp. 652-657, 2015.

Authors Introduction

Mr. Yaxin Li

He received the B.S. degree in automation from Shandong University at 2019. He is currently pursuing the Ph.D. degree with the Seventh Research Division and the Center for Information and Control, School of Automation Science and Electrical Engineering, Beihang University.

Prof. Yingmin Jia

He received the B.S. degree in control theory from Shandong University, China, in 1982, and the M.S. and Ph.D. 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. From February 1995 until February 1996 he was a Visiting Professor with the Institute of Robotics and Mechatronics of the German Aerospace Center (DLR), Oberpfaffenhofen, Germany. He held an Alexander von Humboldt (AvH) research fellowship with the Institute of Control Engineering at the Technical University Hamburg-Harburg, Hamburg, Germany, from December 1996 until March 1998, and a JSPS research fellowship with the Department of Electrical and Electronic Systems at the Osaka Prefecture University, Osaka, Japan, from March 2000 until March 2002. He was a Visiting Professor with the Department of Statistics at the University of California Berkeley from December 2006 until March 2007. Dr. Jia was the recipient of the National Science Fund for Distinguished Young Scholars in 1996, and was appointed as Chang Jiang Scholar of the Ministry of Education of China in 2004. He has been Chief Scientist of the National Basic Research Program of China (973 Program) since 2011, and in particular he won the Second Prize of National Technology Invention Award in 2015. His current research interests include robust control, adaptive control and intelligent control, and their applications in robots systems and distributed parameter systems.

Mr. Jintao Hu

He received the B.S. degree in automation in 2019 from Chongqing University in China. He is currently pursuing the Ph.D. degree with the Seventh Research Division and the Center for Information and Control, School of Automation Science and Electrical Engineering, Beihang University.

Research on Sign Language Recognition Algorithm Based on Improved R(2+1)D

Yueqin Sheng, Qunpo Liu, Ruxin Gao
School of Electrical Engineering and Automation, Henan Polytechnic University, 2001 Century Avenue Jiaozuo, Henan 454003, China

Hanajima Naohiko
College of Information and Systems, Muroran Institute of Technology, 27-1 Mizumoto-cho Hokkaido, Hokkaido 050-8585, Japan
E-mail: 212007010029@home.hpu.edu.cn, lqpny@hpu.edu.cn, gaoruxin@hpu.edu.cn
hana@mondo.mech.muroran-it.ac.jp
www.hpu.edu.cn, www.muroran-it.ac.jp

Abstract

Sign language recognition based on deep learning has advantages in processing large scale dataset. Most of them use 3D convolution, which is not conducive to optimization. In this paper, an improved R(2+1)D model is proposed for isolated word recognition. The model convolves the video frame sequence in space and time dimensions and optimizes the parameters respectively. Based on CELU activation function, the accuracy of sign language recognition is improved effectively. The validity of proposed algorithm is verified on CSL dataset.

Keywords: Sign Language Recognition; R(2+1)D Convolution; 3D Convolution; CELU Activation Function

1. Introduction

Sign language is an important tool for deaf-mutes to communicate, but most normal people have not learned it, which makes it difficult for deaf-mutes to communicate with others.

Different countries and regions use different sign language. Even under the same standard, there are great difference in action made by different signers because of left-handed or right-hander and speed of motion. Besides, part of sign language motion is obscured by hands, so sign language recognition (SLR) is a very challenging task. According to the type of sign language motion, the study of SLR can be divided into isolated word recognition and sentence recognition. This paper studies SLR based on isolated words. Fig. 1 shows a partial frame of the sign language “situation”.

Fig. 1. Example diagram of sign language

2. Study on Sign Language Recognition

Traditional SLR methods mainly include Hidden Markov Model, Dynamic Time Warping (DTW) and Conditional Random Field. Wang et al.[1] achieved 91% recognition accuracy in a data set containing 370 words based on hidden Markov model and gaussian mixture model. Yan et al.[2] improved the traditional DTW by combining dynamic trajectory with type information of key sign language. It is better than traditional DTW in speed and accuracy.

Traditional methods can only solve the problem of SLR in a certain scale dataset. In the current era of big data, SLR based on deep learning is mainstream research trend.

Liu et al.[3] proposed a SLR model based on long short-term memory, which took the motion trajectories of four joints as input. Using skeleton data alone may ignore facial features. Pu et al.[4] obtained the gesture changes of the video through 3D-Convolutional Neural Network (CNN) and used the shape context to describe the trajectory characteristics of the joint to construct a SLR system with two-channel data. However, 3D convolution is difficult to optimize, slow and requires high hardware.

3. Sign Language Recognition Model Based on Improved R(2+1)D

3.1. *(2+1)D convolution*

In static SLR, 2D-CNN plays an irreplaceable role. 3D-CNN that introduces space-time dimension promotes the progress of dynamic SLR. However, both of them have shortcomings. 2D-CNN cannot process the information of time series. 3D-CNN has many parameters, large computation, slow speed and high requirements for hardware.

Based on the above problems, Tran et al.[5] proposed a spatio-temporal feature extraction method that optimizes the 3D convolution kernel into (2+1)D convolution kernel under the situation that 3D convolution has been applied to ResNet. Each residual block consists of two convolution layers followed by a ReLU activation function. If x represents the input data size of $3\times L\times H\times W$, where L represents the number of frames, H and W represent the height and width of video frames respectively, and 3 is the RGB channel of image, the output of ith residual block is:

$$z_i = z_{i-1} + F\left(z_{i-1};\theta_i\right), \tag{1}$$

where z_{i-1} is the output of (i-1)th residual block; $F\left(z_{i-1};\theta_i\right)$ is the output obtained through two convolution layers and two activation functions.

R(2+1)D introduces hyperparameter M_i and uses M_i two-dimensional space convolution kernels with size of and $N_{i-1}\times1\times d\times d$ and N_i one-dimensional time convolution kernels with size of $M_i\times t\times1\times1$ to replace N_i three-dimensional convolution kernels with size of $N_{i-1}\times t\times d\times d$, so as to maintain approximately the same number of parameters as the three-dimensional residual network. The following relation can be obtained:

$$N_{i-1}\times t\times d^2\times N_i = N_{i-1}\times d^2\times M_i + M_i\times t\times N_i, \tag{2}$$

$$M_i = \frac{td^2 N_{i-1}N_i}{d^2 N_{i-1} + tN_i}. \tag{3}$$

When the input is single channel, the 3D convolution kernel and (2+1)D convolution kernel are shown in Fig. 2. The left is the 3D convolution kernel with the size of $t\times d\times d$, where t represents time depth and d represents the height or width of the space. The right is the (2+1)D convolution kernel formed by decomposing 3D convolution kernel. The number of 2D convolution kernel after decomposition is M_i.

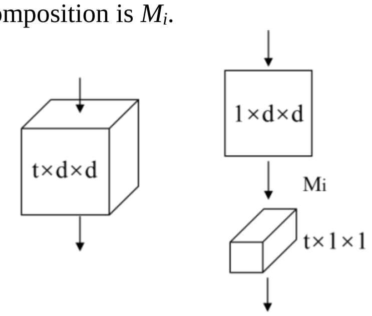


Fig. 2. 3D convolution kernel and (2+1)D convolution kernel

3.2. *Optimization of activation function*

The R(2+1)D model proposed by Tran et al.[5] uses ReLU activation function. ReLU is an activation function commonly used in neural networks, characterized by fast computing speed and good performance. However, when input x<0, the function output is 0. The loss gradient disappears during back propagation, resulting in the failure of parameter updating. To solve this problem, the improved R(2+1)D model in this paper selects CELU[7] as the activation function. CELU is a continuous and differentiable exponential smoothing function with nonlinear turning point which is beneficial to the convergence and generalization of neural networks. The calculation formula of ReLU activation function is shown in Eq. (4). The calculation formula of CELU activation function is shown in Eq. (5). In this paper, the value of α of CELU activation function is 0.05. The output comparison between ReLU and CELU is shown in Fig. 3.

$$\mathrm{ReLU}(x) = max\{0, x\} \tag{4}$$

$$\mathrm{CELU}(x,\alpha) = max\left\{\alpha\left(\exp\left(\frac{x}{\alpha}\right)-1\right), x\right\} \tag{5}$$

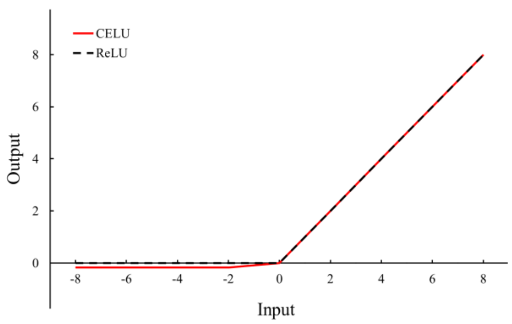


Fig. 3. Activation function curves of CELU and ReLU

3.3. *Improved (2+1)D-ResNet18 model*

The structure of the improved (2+1)D-ResNet18 model proposed in this paper is shown in Fig. 4. The video frame sequence first enters the fully connected layer and the max pooling layer, then enters four improved (2+1)D residual convolution blocks. After that, the average pooling layer and the fully connected layer are entered successively. Finally, the classifier outputs the classification results.

4. Experimental Results and Analysis

The data set we used is CSL isolated word sign Language dataset from University of Science and Technology of China, which contains 500 commonly used sign language words. Considering the training time and hardware requirement, we selected 100 of these words to conduct experiments on 3D-ResNet18, (2+1)D-ResNet18 and improved (2+1)D-ResNet18 in this paper. The data set was divided into training set, validation set and test set in a ratio of 7:2:1. Each continuous video was extracted into discrete video frames. 16 frames were sampled from each video frame set as the input of the model using uniform sampling method. The experiments were carried out in the same experimental environment. Fig. 5 and Fig. 6 show the validation result curves of the three models. Table 1 shows the accuracy of the three models on the test dataset.

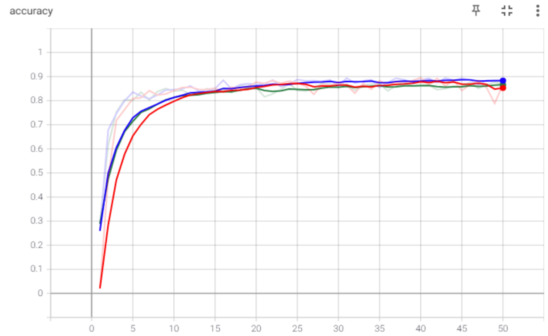


Fig. 5. Validation accuracy curves of 3D-ResNet18(green), (2+1)D-ResNet18(red) and improved (2+1)D-ResNet18(blue)

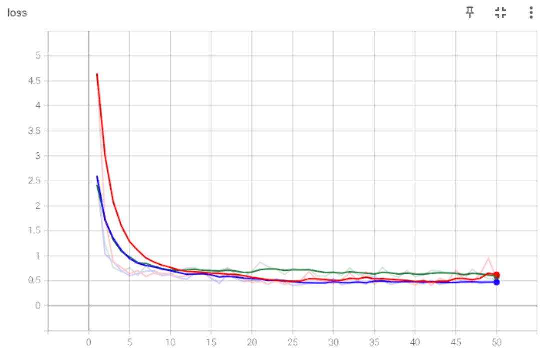


Fig. 6. Validation loss curves of 3D-ResNet18(green), (2+1)D-ResNet18(red) and improved (2+1)D-ResNet18(blue)

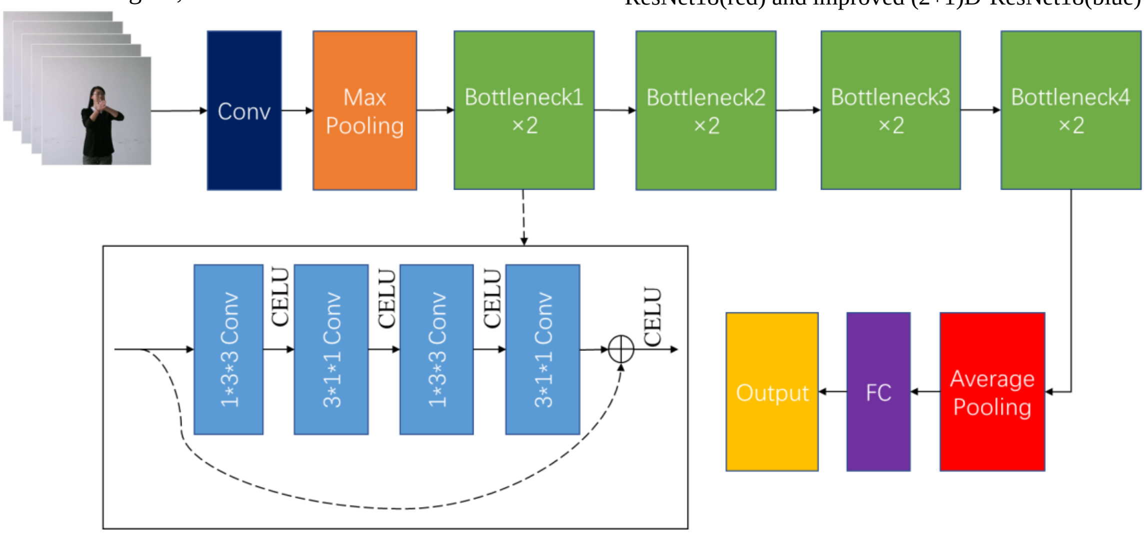


Fig. 4. Structure diagram of improved (2+1)D-ResNet18 model

As can be seen from the validation curves, the improved (2+1)D-ResNet18 has the fastest speed of accuracy increase and loss reduction. It is the first one to reach the minimum value of loss, and its curve is the smoothest.

Table 1. Test results of models

Model	Test Accuracy
3D-ResNet18	86.94%
(2+1)D-ResNet18	87.76%
Improved (2+1)D-ResNet18	88.92%

It can be seen from Table 1 that the test accuracy of 3D-ResNet18 is 86.94%. The accuracy of (2+1)D-ResNet18 obtained by separating spatial dimension and time dimension is 87.76%. The improved (2+1)D-ResNet18 has the highest accuracy of 88.92%. The CELU activation function in the improved model solves the problem of gradient disappearance during back propagation, thus further improving the accuracy.

5. Conclusion

This paper proposes an improved R(2+1)D model for isolated word recognition. The model separates spatial convolution and time convolution and use CELU as the activation function, reaching the accuracy of 88.92%. In the future, sign language in complex environments will be further studied.

References

1. Hanjie Wang, Xiujuan Chai, Xilin Chen. Sparse Observation (SO) Alignment for Sign Language Recognition. Neurocomputing, Volume 175, Part A, 2016, Pages 674-685.
2. Y. Yan, Z. Li, Q. Tao, C. Liu and R. Zhang, "Research on Dynamic Sign Language Algorithm Based on Sign Language Trajectory and Key Frame Extraction," 2019 IEEE 2nd International Conference on Electronics Technology (ICET), 2019, pp. 509-514.
3. T. Liu, W. Zhou and H. Li, "Sign language recognition with long short-term memory," 2016 IEEE International Conference on Image Processing (ICIP), 2016, pp. 2871-2875.
4. Pu J., Zhou W., Li H. (2016) Sign Language Recognition with Multi-modal Features. In: Chen E., Gong Y., Tie Y. (eds) Advances in Multimedia Information Processing - PCM 2016. PCM 2016. Lecture Notes in Computer Science, vol 9917. Springer, Cham. https://doi.org/10.1007/978-3-319-48896-7_25.
5. Tran, D., Wang, H., Torresani, L., Ray, J., LeCun, Y., and Paluri, M., “A Closer Look at Spatiotemporal Convolutions for Action Recognition”, <i>arXiv e-prints</i>, 2017.
6. Barron, J. T., “Continuously Differentiable Exponential Linear Units”, <i>arXiv e-prints</i>, 2017.

Authors Introduction

Ms. Yueqin Sheng

She received her Bachelor's degree from the Department of Automation, Henan Polytechnic University (China) in 2020. She is currently a Master’s course student in Henan Polytechnic University (China).

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 and master tutor at the School of Electrical Engineering and Automation, Henan Polytechnic University (China). He is mainly engaged in teaching and research work in robotics, intelligent instruments and industrial automation equipment.

Dr. Ruxin Gao

He graduated from Huazhong University of Science and Technology (China) with a ph. D. in Pattern Recognition and Intelligent Systems. He is an associate professor and master tutor at the School of Electrical Engineering and Automation, Henan Polytechnic University (China). He is mainly engaged in teaching work of computer application and research work of image processing and visual related.

Dr. Hanajima Naohiko

He graduated from the Hokkaido University (Japan) of Technology in Japan with a Ph.D. He is a professor at the College of Information and Systems at Muroran Institute of Technology (Japan). He is mainly engaged in the research work of robotics and intelligent equipment.

Directional Flocking of Multi-Agent system Caused by Limited Visual Field

Yongnan Jia
School of Automation and Electrical Engineering, University of Sciences and Technology Beijing, Beijing, 100083, P. R. China
Jiali Han
School of Automation and Electrical Engineering, University of Sciences and Technology Beijing, Beijing, 100083, P. R. China
Yong Xie
The System Design Institute of Mechanical-Electrical Engineering, Beijing, 100854, China.
Weicun Zhang
School of Automation and Electrical Engineering, University of Sciences and Technology Beijing, Beijing, 100083, P. R. China
E-mail: ynjia@pku.edu.cn

Abstract

Experiment evidence has proved that the visual field of each individual in biological swarms is usually non-omnidirectional. Therefore, we introduce limited visual field to the egalitarian flocking model. The directional flocking problem refers to the flocking problem that all the individuals are expected to move in a specified direction, which is decided by the leader. This paper mainly compared the limited-visual-field flocking model with the classic flocking model (that is the egalitarian one) from the point view of rate of convergence. Experimental results indicated that limited-visual-field flocking model is more efficient than the omnidirectional one for the directional flocking problem.

Keywords: Flocking model, limited visual field, egalitarian flocking model, rate of convergence.

1. Introduction

Flocking phenomena widely exist among social animals, such as bacteria, birds, bees, herds, and fishes. In the colony, each individual just relies on local interaction and simple decision−making rules to emerge a global dynamical behavior for the sake of obtaining the chance of survival. Several models, such as Boid model[1], Vicsek model[2], Cucker-Smale model [3], have been developed for years in order to reveal the mechanism or principle behind these flocking behaviors.

During these classical flocking models, there is a common prerequisite, that is, each individual has an omnidirectional visual field. However, in nature, individuals with limited visual fields in a flock is a more universal phenomenon. For example, the visual field of starlings is 143°, the visual field of pigeons is 158°, and the visual field of owls is 100.5° [4-6]. Therefore, limited visual field should be included into flocking models.

Compared with the flocking model with omnidirectional visual field, limited visual field not only led to limited environment information, but also brings in a polarity for the emergence of the group behavior. Thus, we define a new flocking model called directional flocking. For each pair of individuals, if they can be seen by each other, then they are neighbors and their information can be interacted bidirectional; if agent *i* can be seen by agent *j* but agent *j* cannot be seen by agent *i*, then there exists an unequal/polar relationship. This paper will quantitatively discuss the classical flocking model and the directional flocking model caused by the limited visual field. We would like to know that whether the directional flocking is better than the classical one in the respects of rate of convergence and stability, especially for these tasks that require individuals moving from one place to another following an empirical route, such as migration of wild goose. Besides, we will try to give an explanation on the evolution of the limited visual field.

The rest of this paper is organized as follows. The flocking model is given and the directional flocking problem is formulated in section II. Section III presents the simulation results of the two flocking models. Finally, the conclusions are drawn in section VI.

2. Modeling

2.1. *Flocking model*

Consider N particles moving continuously (off lattice) in a free area without any boundary limitation. Without loosing generality, suppose that the time interval between two updates of the directions and positions is $\Delta t = 1$.

At $t = 0$, Each particle is randomly distributed within an area of a given size and has the same absolute velocity v_d as well as randomly distributed direction $\theta_i, i, ..., N$. At each time step, the position of the ith particle is updated according to

$$x_i(t+1) = x_i(t) + v_i(t)\Delta t \tag{1}$$

In each time step, the velocity of a particle $v_i(t+1)$ is updated according to the following equation

$$v_i(t+1) = v_i^{align}(t+1) \tag{2}$$

where $v_i^{align}(t+1)$ is the alignment term. The alignment term was constructed to have an absolute value v_d and a direction given by the angle $\theta_i^{align}(t+1)$. The angle was obtained from the expression

$$\theta_i^{align}(t+1) = <\theta_i(t)> + \Delta\theta(t) \tag{3}$$

where $\Delta\theta(t)$ represents noise, which is a random number chosen with a uniform probability from the interval $[-\eta/2, \eta/2]$. $<\theta_i(t)>$ denotes the average direction of the velocities of neighbors of the given particle i. The average direction is given by the angle

$$<\theta_i(t)> = \arctan\left(\frac{\sum_{j=1}^{N} l_{ij}(t)\sin(\theta_j(t))}{\sum_{j=1}^{N} l_{ij}(t)\cos(\theta_j(t))}\right). \tag{4}$$

Neighbor matrix $L_N(t) = [l_{ij}(t)]_{N\times N}$ describes the neighbor relationships of particles at time t [7], where

$$l_{ij}(t) = c_{ij} * b_{ij} * a_{ij}(t), \forall i, j = 1, ..., N. \tag{5}$$

Therein, $A_N(t) = [a_{ij}(t)]_{N\times N}$ is the adjacency matrix, $C_N(t) = [c_{ij}(t)]_{N\times N}$ is the contribution matrix, and $B_N(t) = [b_{ij}(t)]_{N\times N}$ is the dominance matrix.

The contribution matrix $C_N(t) = [c_{ij}]_{N\times N}$ (c_{ij} >0) is defined to describe the contribution strength of each particle during the decision making process regarding the new preferred directions of the particles.

The dominance matrix $B_N = [b_{ij}]_{N\times N}$ is defined to describe the direction of information flow between each pair individuals. The direction of information flow specifies the set of particles whose behavior influences the decision of a given particle at each time step. For the classical flocking model, that is the egalitarian model, the dominance matrix is an undirected graph. For the directional flocking model, the dominance matrix is the mixed graph, since sometimes the information flow of the pairwise particles are bidirectional while sometimes the information flow of the pairwise particles is directional, which depends on whether the other one is located in its visual field or not.

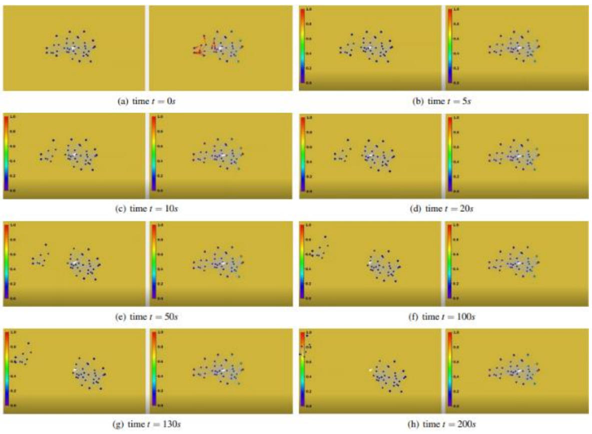


Fig. 2. The critical snapshots of the evolutionary process of the egalitarian flocking model and directional flocking model with hierarchical structure. The left one is the egalitarian flocking model and the right one is the directional flocking model.

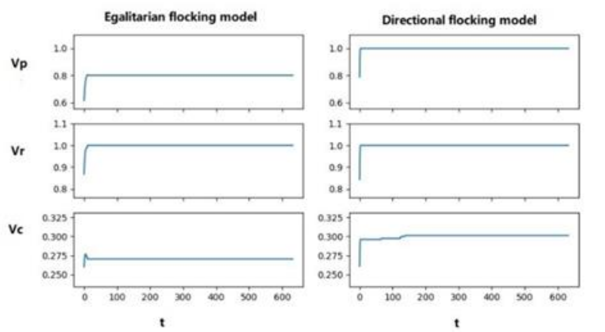


Fig. 3. Three order parameters of the two flocking models versus time.

The definition of adjacency matrix $A_N(t)$ is

$$a_{ij}(t)\begin{cases}1, i=1,...,N,\ j\in N_i(t)\\ 0, otherwise\end{cases}, \tag{6}$$

where $N_i(t)$ is the set of *j* satisfying that the distance between *i* and *j* is no larger than the interaction radius *r*. Using the above expressions the update strategy of the velocity of each particle can be written as

$$v_i^{align}(t+1)=c^{align}v_d e_i(t) \tag{7}$$

where c^{align} is the coefficient of the alignment term. $e_i(t)$ is a unit vector with direction angle $\theta_i^{align}(t+1)$.

2.2. *Order Parameters*

In order to thoroughly investigate the differences between the classical flocking model with omnidirectional visual field (that is, $\theta_i=\theta_d=180°$) and the directional flocking model with limited visual field $\theta_i=\theta_d=150°$ in a quantitative way, three order parameters are given to characterize the evolutionary process of these particles governed by the above two flocking models from different point views.

The first order parameter V_p is the average value of normalized velocity of all individuals, and its formula is as follows:

$$V_p=\left|\frac{1}{N}\sum_{i=1}^{N}v_i\right| \tag{8}$$

of particle *j*. The color bar reveals the weight of the contribution of the given particle. The red particle is the strongest one, while the purple one is the weakest particle.

where $V_p\in[0,1]$. $V_p=1$ denotes that all the individuals in one group move with the same velocity (including value and direction), while when $V_p=0$ equals to the initial state that the average value of the velocities of all the individuals become zero (like random distribution).

The second order parameter V_r is the cluster-dependent velocity correlation, whose definition is

$$V_r=\frac{1}{T}\frac{1}{N}\int_0^T\sum_{i=1}^{N}\frac{1}{N_i-1}\sum_{j\in J_i}\frac{v_i\bullet v_j}{|v_i||v_j|}dt \tag{9}$$

where N_i is the number of particles in the cluster that contains the *i*th particle, and J_i refers to the set of indices of particles that are in the same cluster as the *i*th particle.

The third one V_c is also a velocity correlation but not depends on cluster, whose definition is

$$V_c=\frac{1}{T}\frac{1}{N(N-1)}\int_0^T\sum_{i=1}^{N}\sum_{j\in J_i}\frac{v_i\bullet v_j}{|v_i||v_j|}dt \tag{10}$$

3. Simulation Analysis

The simulations were carried out in a free two-dimensional area. We considered groups of N=40 particles. We aimed at comparing the rate of convergence and stability of egalitarian flocking model versus directional flocking model with hierarchical structure. In order to keep the comparability of these simulation results, both models have the same initial positions and the same initial velocities. For the directional flocking model, the visual field of each particle is $\theta_d=150°$, while that of each particle of egalitarian flocking model is $\theta_d=180°$. Besides, the element c_{ij} of the contribution matrix of directional flocking model satisfies log-normal distribution with mean value 0 and standard deviation 1, while that of egalitarian flocking model is the average value of these c_{ij} of directional flocking model. For egalitarian flocking model, pairwise particles are equal, that is, if particle *i* is the neighbor of particle *j*, then *j* is also the neighbor of particle *i*. However, for directional flocking model, pairwise particles sometimes are equal, sometimes are non-equal. That is to say, we cannot obtain the conclusion that *j* is the neighbor of particle *i* according to the condition that particle *i* is the neighbor

Furthermore, we record the three order parameters during the evolutionary process of 40 particles versus time in Fig. 3. It is clearly to see that the directional flocking model has a better consensus properties and a faster convergent rate according to V_p and V_c. V_p denotes that these particles in each cluster converge to consensus finally.

4. Conclusion

We have investigated the directional flocking problem of multiagent system governed by hierarchical structure. The limited visual field of each particle produces a polar for the interaction network of these particles. The well ordered initial states and limited visual field force the flocking model has better convergent rate and consensus property than the classical flocking model (the egalitarian one), especially for these navigation tasks of multi-agent system. It is amazing that the interaction network of these particles in directional flocking model caused by limited visual field gradually from a mixed state (including bidirectional interaction and directional interaction) to a pure directional one. We cannot explain why they evolute like this. However, the conclusion conform to the collective behaviors of these real social animals in nature. Displayed equations should be numbered consecutively in each section, with the number set flush right and enclosed in parentheses.

References

1. Reynolds, C.W., "Flocks, herds, and schools: a distributed behavioral model", ACM SIGGRAPH Computer Graphics, 21(4), 25–34,1987.
2. Vicsek, T., Czirok, A., Ben-Jacob, E., Cohen, I., and Shochet, O. "Novel type of phase transition in a system of self-driven particles", Physical Review Letters, 75(6), 1226,1995.
3. Cucker, F. and Smale, S.,"Emergent behavior in flocks. IEEE Transactions on Automatic Control", 52(5), 852–862,2007.
4. Couzin, I., Krause, J., and James, R., "Collective memory and spatial sorting in animal groups", Journal of Theoretical Biology, 218(1), 1–11,2002.
5. Martin, G.R., " Visual fields in woodcocks scolopax rusticola (scolopacidae; charadriiformes)" , Journal of Comparative Physiology A, 174(6), 787-793,1994.
6. Andreas, H. and Christian, W., "The simulation of the movement of fish schools " , Journal of Theoretical Biology, 156(3), 365-385, 1992.
7. Jia, Y.N. and T., V. "Modelling hierarchical flocking", New Journal of Physics, 21, 093048, 2019.

Authors Introduction

Dr. Yongnan Jia

She is currently an associate professor of University of Science and Technology Beijing. She received her Ph.D degree in Intelligent Control Laboratory, College of Engineering, Peking University in 2014. She is currently a post-doctoral lecturer in the School of Automation and Electrical Engineering, University of Sciences and Technology Beijing. Her research interests include intelligent robots, biomimetic robots, and collective behaviors of multi-robot system.

Ms. Jiali Han

She received her B.S. degree in automation in 2021 from the School of Engineering, Beijing Forestry University in China. He is acquiring the M.E. in University of Science and Technology Beijing. Her research interests include collective behavior analysis, and mutli-agent system.

Mr. Yong Xie

He is currently an engineer in the System Design Institute of Mechanical-Electrical Engineering. He received his Master's degree in 2015 from the School of Aeronautic Science and Engineering, BUAA University in China. His research interests include intelligent control theory, positioning, and navigation of unmanned arerial vehicles.

Dr. Weicun Zhang

He is currently an associate professor of University of Science and Technology Beijing. He received his Doctor's degree in Control Theory from Tsinghua University (Beijing) in 1993. He has been a visiting research fellow at the University of Michigan from 1997 to 1998. His research interests include adaptive control, multi-agent system, positioning and navigation of unmanned vehicles.

An Object Acquisition Based on Human-Robot Cooperation

Kota Ito
Graduate of Engineering, Kyushu Institute of Technology, Sensui-cho Tobata-ku
Kitakyushu, 804-8550, Japan

Masuhiro Nitta, Seiji Ishikawa[†], Joo Kooi Tan
Faculty of Engineering, Kyushu Institute of Technology, ‡N&N Inc. Sensui-cho1-1 Tobata-ku
Kitakyushu, 804-8550, Japan
E-mail: {ito-kota ,etheltan}@ss10.cntl.kyutech.ac.jp
http://lab.cntl.kyutech.ac.jp/~etheltan/

Abstract

In this paper, we propose a human-robot cooperative system to support shopping refugees. In the system, a robot acquires an object specified by a person in a distant site. The normal vector is calculated from the depth image, and the region is segmented using GBS on an RGB image. The two obtained clues are used to accurately detect the position of the specified object. The effectiveness of the proposed method was verified by experiments.

Keywords: Region Division, Normal Vector, Newton-Raphson Method, GBS

1. Introduction

According to the report by Ministry of Internal Affairs and Communications in 2017, the number of shopping refugees in Japan has been increasing in recent years; as of 2014, the number of shopping refugees in Japan was about 7 million. It also increased by about 1 million people in the six years between 2008 and 2014.[1] The term "shopping refugees" refers to "the people who have difficulty in making daily purchases such as groceries". The number of shopping refugees is on the rise, making it necessary to take countermeasures.

According to a survey by the Ministry of Economy, Trade and Industry (METI) in 2012, the market size of the Japanese robotics industry is expected to expand more than tenfold between 2012 and 2035.[2] Among them, robots in the service field are expected to account for about 51% of the total robot market in Japan by 2035. Looking at the world, the market for household and personal service robots is growing at a high rate of 16.3 million units in 2018, a 59% increase over the previous year, and the service robot market is also expanding year by year, doubling in size.[3] The current robot market is dominated by industrial and manufacturing robots used in limited locations, but service robots that help people in their daily lives are expected to increase in the future. In this paper, we propose an object grasping robot system for assisting shopping refugees among service robots.

In previous researches, various robots have been proposed to take the place of humans for object grasping [4-8]. However, these methods do not deal with a large number of unknown objects, such as methods for grasping objects at specific locations[4], methods limited to specific objects or environments[5,6], or methods that require the robot to learn unknown objects[7,8]. The proposed method is an extension of the method of Sato et al[9].

This paper proposes a method of object grasping by a human-robot cooperative system using an RGB-D camera. The proposed method detects the position and pose of an object specified by a person from two pieces of information: the normal vector calculated from the depth image and the result of region segmentation obtained using Graph Based Segmentation (GBS)[10] for RGB images. In the proposed method, a person instructs a robot to go to a specified site. The person selects an object he/she wants on the images sent from the robot. Then the robot knows which object the person selected and it acquires the object. The proposed system is adaptable to unknown environments and objects.

2. Outline of the Proposed System

In this section, an outline of the proposed human-robot cooperative system is described. In the system, an instruction screen, as shown in **Fig. 1**, is displayed to the user from the image transmitted to the user from a robot in a remote location. The user instructs the robot to move by touching the instruction screen, and the robot moves around the site. Depending on the area touched by the user in the instruction screen, the robot's behavior is determined as follows;

(i) Original image area: Move forward or object selection,
(ii) Green area: Retreat,
(iii) Blue area: Turn left,
(iv) Red area: Turn right,
(v) STOP button: Stop.

When the original image area is touched, it judges whether to move forward or select an object based on the average distance from the surrounding area of the touched pixel. If the selected object is found, the object is separated into regions based on the coordinates touched by the user. The details of region segmentation are described in Section 3.

3. Region segmentation of an object

In this section, we describe a method of region segmentation of user-specified objects. The depth image used in this method is characterized by the fact that the distance of the object boundary is often not obtained. Therefore, if we perform region segmentation using only distance information, the estimated object region will be smaller than the actual object. The proposed method can accurately detect objects even in the area where it is difficult to obtain distance information, such as object boundaries, by using two types of information; normal vector information from depth images and region segmentation results from RGB images.

Fig. 1. Robot motion instruction screen

3.1. *Normal vector*

In the proposed method, the normal vector is first calculated for each pixel in the entire depth image. Then, the object region is expanded by calculating the inner product of the normal vector of the pixel specified by the user and the normal vectors of its surroundings[11]. In this section, we describe a method of estimating the candidate region of an object using normal vectors.

3.1.1. *Calculating normal vector*

To calculate the normal vector, we use the position vectors of the three neighboring points of the pixel of interest for all pixels in the depth image to calculate the normal vector as shown in the following equation.

$$\boldsymbol{n} = (\boldsymbol{b} - \boldsymbol{a}) \times (\boldsymbol{c} - \boldsymbol{a}) \quad (1)$$

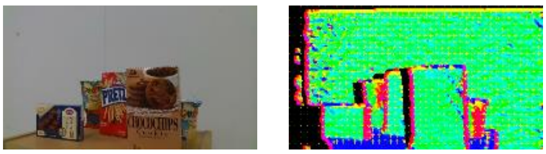
Then, the length of the calculated normal vector is normalized to 1 by the following equation.

$$\hat{\boldsymbol{n}} = \boldsymbol{n}_G / \|\boldsymbol{n}_G\|_2 \quad (2)$$

The result of calculating the normal vectors for all pixels in the depth image is shown in **Fig. 2**.

3.1.2. *Region extension*

Using the normal vectors calculated in 3.1.1, we expand the candidate object region. In the region expansion, an



(a) (b)

Fig. 2. The result of normal vector calculation: (a) An input image, (b) obtained normal vectors.

$n \times n$ pixels region centered at the coordinates obtained from the user is set as the initial object region. Then, for each of the 8-neighbor pixels of the object region, it is determined whether the pixel is an object region or an edge region.

For the judgement, the inner product of the normal of the center pixel and the normal of the pixel which is p pixels apart from the center pixel is calculated in each of the 8 directions. The inner product obtained in each direction is then used to perform thresholding as shown in the following;

$$\delta_{min} = \min\{\delta_N, \delta_{NE}, \cdots, \delta_W, \delta_{NW}\}$$
$$L = \begin{cases} 1 & \delta_{min} > \delta_{th} \\ 0 & otherwise \end{cases} \quad (3)$$

The pixels with L=1 are judged as the pixels in the object region, whereas the pixels with L=0 are judged as those in the edge region.

In the proposed method, the new pixel identified as the pixel in the object region is chosen as the next search point. The region expansion is iterated until all the object regions are surrounded by edge regions.

3.2. *Region correction using color information*

The object candidate regions estimated using normal vectors in Section 3.1 are corrected using the color information of the image. In the proposed method, GBS is performed on RGB images and the obtained segmentation results are used for region correction.

3.2.1. *Graph based segmentation (GBS)*

In the proposed method, GBS is used as a method for segmenting RGB images. GBS is a method of dividing an image into multiple regions by grouping the regions into pixels that have similar pixel values. In the method, GBS is used for the segmentation because it is fast and the parameters for the likelihood of region segmentation can be set.

3.2.2. *Region correction*

The object candidate regions estimated in Section 3.1 are corrected using the segmentation results obtained by GBS described in Section 3.2.1. Since the RGB-D camera used in this method has the characteristic that it is difficult to obtain the distance information of the object boundary, GBS is performed on n pixels around the object candidate region estimated in Section 3.1. By limiting the range of the GBS, the processing time can be reduced and the number of false estimates can also be reduced. Then, for each region created by GBS, we calculate the percentage of the region occupied by the candidate object region estimated in Section 3.1, and if the percentage is greater than a threshold, we estimate the region as the object region.

4. Object Grasping

In the proposed method, the object is grasped based on the estimated position and posture information of the object after the object is divided into regions specified by the user. To estimate the position of an object, an $n \times n$ rectangle is placed at the center of gravity of the detection area, and the average of the 3D positions is calculated. The calculated 3D position is then converted from the camera coordinate system to the robot arm coordinate system to estimate the position of the object.

For object pose estimation, the object pose is obtained from the normal vector of the center of gravity pixel in the detection area. From the estimated position and posture of the object, the joint angle of the robot arm is determined by solving the inverse kinematics problem using the Newton-Raphson method, and the object is grasped.

5. Experiment

Experiments were conducted to show the performance of the proposed method. The camera used in the experiment was an Intel RealSense D415. The camera was placed in front of the robot.

5.1. *Region segmentation experiment*

The first experiment was an object segmentation experiment. In the experiment, we performed region segmentation on the images taken from three different viewpoints for three types of objects with different arrangements and evaluated the accuracy. For objects A and C, the whole shape is visible in the image, but object B is partially hidden by other objects. The accuracy of region segmentation is evaluated using the F value.

As a result, the F value was 90.9% for object A, 77.7% for object B, and 77.5% for object C.

An example of the experimental results is shown in **Fig. 3**. Figure 3(a) is an original image, Fig. 3(b) shows

(a) (b) (c) (d)

Fig. 3. An experimental result on region segmentation: (a) Input image, (b) normal vectors, (c) after correction, (d) the result.

Fig. 4. Scenes on object grasping experiment

the object candidate region estimated by the normal vector, Fig. 3(c) shows the result of correction using color information, and Fig. 3(d) shows the result on the object region detection. Note that the yellow frame shows the detected object.

5.2. *Object grasping*

The second experiment was object grasping. A user gave instruction to the robot to move and acquire the object specified by the user at two different locations. In the object grasping, the user controlled the robot from a remote location. The experiments were conducted five times for a total of ten object grasps, and the accuracy was evaluated.

As a result of the experiment, the success rate of object grasping was 80% in five experiments. The experimental scene is shown in **Fig. 4**.

6. Conclusion

In this paper, we proposed a method of object acquisition by a human-robot cooperative system using an RGB-D camera. The effectiveness of the proposed method was verified by experiments. In the object region segmentation experiment, the *F* value was 82.0%, whereas, in the object grasping experiment, the success rate of object acquisition was 80%. In the future, we aim at developing a system that can handle complex environments with obstacles, such as an actual store.

References

1. Ministry of Internal Affairs and Communications: "Fact-finding survey on measures for shopping refugees", pp.2-4, 2017. (In Japanese)
2. Ministry of Economy, Trade and Industry: "Market Trends in the Robotics Industry", pp.17-19, 2012. (in Japanese)
3. Japan Machinery Federation: "FY2019 Research Report on the Promotion of Robotics Industry and Technology", pp.18-19, 2020.
4. H. Masuta, H. Lim, T. Motoyama, K. Koyanagi, T. Oshima: "Invariant perception for grasping an unknown object using 3D depth sensor", *2015 IEEE Symposium Series on Computational Intelligence*, pp.122-124, 2015.
5. S. Takaki, J. K. Tan, S. Ishikawa: "A human care system by a mobile robot employing cooperative objects recognition", *Proceedings of the 2017 IEEE Region 10 Conference*, pp.1148-1152, 2017.
6. J. Stückler, R. Steffens, D. Holz, S. Behnke: "Efficient 3D object perception and grasp planning for mobile manipulation in domestic environments" *Robotics and Autonomous Systems*, Vol.61, No.10, pp.1106-1115, 2013.
7. K. Suzuki, Y. Yokota, Y. Kanazawa, T. Takebayashi: "Online self-supervised learning for object picking: detecting optimum grasping position using a metric learning approach", *2020 IEEE/SICE International Symposium on System Integration (SII)*, pp.1-8, 2020.
8. E. Maiettini, G. Pasquale, L. Rosasco, L. Natale: "Interactive data collection for deep learning object detectors on humanoid robots", *Proceedings of the 2017 IEEE-RAS 17th International Conference on Humanoid Robotics (Humanoids)*, pp.862-868, 2017.
9. Joo Kooi Tan, Akitoshi Sato: "Human-robot Cooperation Based on Visual Communication", *International Journal of Innovative Computing, Information and Control*, Vol.16, No2, pp.543-554, 2020.
10. P. F. Felzenszwalb, D. P. Huttenlocher: "Efficient graph-based segmentation", *International Journal of Computer Vision*, Vol.59, No.2, pp.167-181, 2004.
11. Akitoshi Sato, Joo Kooi Tan, Yuta Ono: "Development of a Human-Robot Cooperative System Based on Visual Information, *Proceedings of SPIE International Workshop on Advanced Image Technology (IWAIT2019)*, 6pages, 2019.

Authors Introduction

Mr. Kota Ito

He received his B.E. degree in Department of Control Engineering in 2020 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, object detection, robot vision and robotics.

Dr. Masuhiro Nitta

He received his PhD from Nara Institute of Science and Technology. He is now an assistant professor in Faculty of Engineering at Kyushu Institute of Technology in Japan .

Emeritus Prof. Seiji Ishikawa

He graduated from Tokyo University and was awarded BE, ME and PhD there. He is now Emeritus Professor of Kyushu Institute of Technology. He was Visiting Researcher of the University of Sheffield, UK, and Visiting Professor of Utrecht University, NL. His research interests include visual sensing & 3-D shape/motion recovery. He was awarded The Best Paper Award in 2008, 2013 and 2015 from Biomedical & Fuzzy System Association, Japan. He is a member of IEEE.

Prof. Dr. Joo Kooi Tan

She is currently with Department of Mechanical and Control Engineering, Kyushu Institute of Technology, as Professor Her current research interests include ego-motion, three-dimensional shape and motion recovery, human detection and its motion analysis from videos. She was awarded SICE Kyushu Branch Young Author's in 1999, the AROB Young Author's Award in 2004, Young Author's Award from IPSJ of Kyushu Branch in 2004 and 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 musculoskeletal walking simulator for analysis of human walking and rehabilitation

Nobuo Sakai*, Ryu Yukiho, Tsubasa Ikeda
Mechanical and Control Engineering, Kyushu Institute of Technology
1-1 Sensucho, Tobata-ku, Kitakyushu city, 804-8550, Japan
E-mail: sakai.nobuo@ise.kyutech.ac.jp

Masako Fuchi
Ocupational therapy, Kyushu Nutrition Welfare University
1-5-1 Katsuhara-Takamatsu, Minami-ku, Kitakyushu city, 800-0298, Japan

Katsuki Hayashi
Rehabilitation division, Seiai Rehabilitation Hospital
2-7-1 Minamioori, Onojyo city, 816-1199, Japan

Mochimitsu Komori
Electrical and Electric Engineering, Kyushu Institute of Technology
1-1 Sensucho, Tobata-ku, Kitakyushu city, 804-8550, Japan

Abstract

In the field of rehabilitation, the explanation of walking motion, called ‘Rocker function’, is one of the practical benchmarks for the facilitation of human walking in clinical settings. If this explanation is true, we can reconstruct it by artificial materials. In this study, we tried to develop the musculoskeletal walking simulator, which can actually reproduce bipedal walking according to the rocker function. Muscles and tendons including biarticular arrangements were represented by springs and cables. In this report, 3 muscles were actuated by servomotors. The simulator reproduced the human musculoskeletal walking motion generated from its intellectual structure in nature. The results would support the insight of the explanation of the rocker function in the rehabilitative treatments.

Keywords: Human walking, Bipedal robot, Musculoskeletal system, Biomechanism, Biomimetics

1. Introduction

Musculoskeletal simulators have been one of the useful tools for considering the motions and their functions of the human body. The spread of computer calculation systems enabled the dynamic mechanical simulation [1]. Around the same time, researches on bipedal robots were carried out by the interests in biomechanics [2] and the powered prostheses [3], as well as mechanical system and control theory. Well-known walking principals would be LIP (Linear Inverted Pendulum) [4] and SLIP (Spring–Loaded Inverted Pendulum) [5] model. Along with these robots, a passive dynamic walking robot [6], which can walk on a slight slope, was advocated.

As the musculoskeletal system of the human leg has peculiar structures, a musculoskeletal legged robot was developed by the rubber artificial muscles [7]. The biarticular arrangement of the muscles can produce high-power output and precise motion simultaneously by its mechanical structure, which was applied in a vertical jumping robot [8]. The robotics on the study of musculoskeletal system is an interesting tool to recognize living bodies and human.

Various biomechanics studies on the analysis of human walking have been reported for the knowledge of clinical application. Perry described the human normal and pathological walking by integrating the elements of human walking, which called ‘Rocker function’ [9]. Today, the rocker function has become an essential model in clinical gait analysis in Japan.

In concerning the walking facilitation in post stroke therapy, it is important to recognize the normal human walking. If the musculoskeletal structure and its mechanical function are understood, it should be possible to reproduce it in a physical simulator [10]. The aim of this study is to construct the physical simulator of the lower limb to understand the human musculoskeletal structure, and to generate the walking motion according to the rocker function. In this paper, the gait of the robotic walking simulator is described in detail according to the rocker function, than the parametric study for the gait construction [11].

2. Materials and Methods

2.1 Design of the musculoskeletal simulator

A simple description of the rocker function is given in Figure 1. First, the knee joint is flexed from the beginning of the load-response phase, which appears inefficient from an engineering view because the quadriceps muscles needs energy to maintain the knee joint in flexion under load. In addition, direct braking by the soleus muscle during ankle rocker to prevent knee breakage also leads to energy consumption.

In the human musculoskeletal structure, muscles are efficient in isometric to centrifugal contraction. Recent muscle research has found that the Titin is a passive spring mechanism that can change its anchor position in muscle force generation [12]. Although its properties are not fully understood, it was reported to be necessary for fast centrifugal contraction [13]. The elastic recoil of the muscle-tendon complex is thought to contribute to the high energy efficiency of walking in terms of elastic energy recovery. Therefore, the basic idea of this study is to represent each muscle as a spring. The springs were attached via a cable and do not exert force in the compression direction. In the case of active driving, a motor pulls a cable attached to the spring in a spring-motor series configuration [8].

2.2 Development of the simulator

A schematic of the model is shown in Fig. 2. The configuration is a common 9-muscles model including biarticular muscles [1]. The robotic simulator does not have any foot structures such as foot arches or toes. The total height of the robot is 800 mm, the hip joint width is 200 mm, and the weight is 5.66 kg including 6 motors for both legs and 500 g weight for the waist to improve the gait. At this stage, the robot is designed by two-dimensional walking with ball rollers between two acrylic plates placed 452 mm apart.

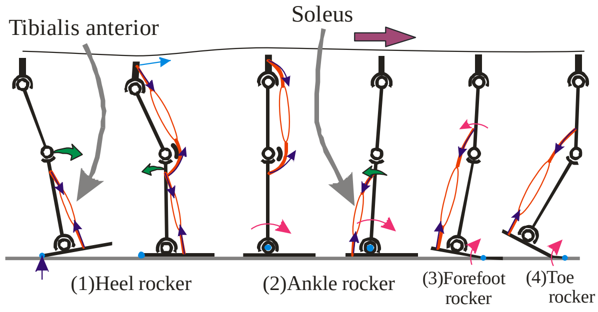


Fig.1 Schematic drawing of the Rocker function.

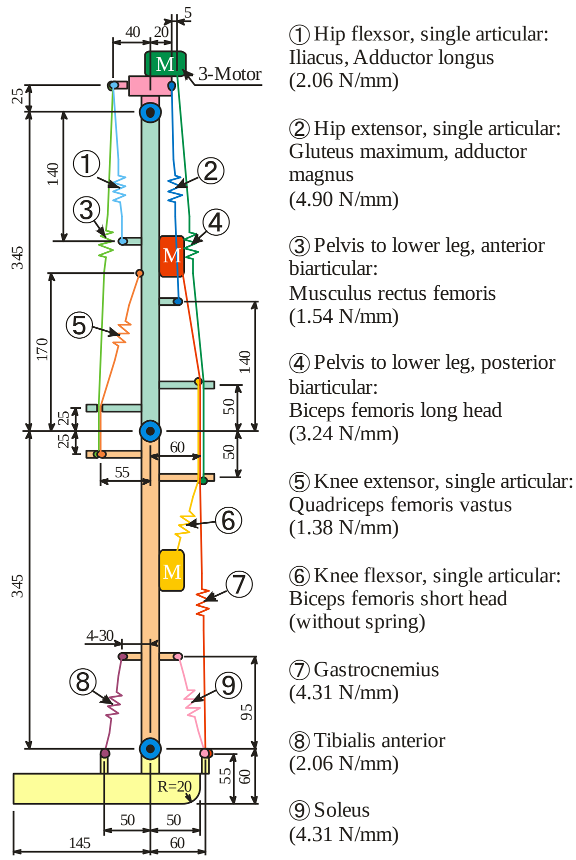


Fig. 2 Configuration of the robotic walking simulator.

The springs, which simulate muscles, are placed in the order of Origin-Cable-Spring-Cable-Insertion, through a cable passage hole in order to adjust the moment arm. The corresponding muscles are shown in Fig. 2 as a reference. The position of some springs seems to be different from the common 9-muscles model, but the function is physically the same no matter where the springs are placed on the spring-cable series. The motor module (Dynamixel MX-106, Robotis Inc.) is placed at the origin or insertion of the cable in series to the spring.

2.3 Actuation of the simulator and evaluation

In this paper, we report the results of active driving of the three posterior muscles (④⑥⑦ in Fig. 2). With regard to the spring constant, each muscle does not generate a force proportional to its muscle cross-sectional area during walking. The motor arrangement and its drive timing, and especially the spring constant, were adjusted by trial and error. The heuristic search for drive timing and algorithms with reference to the living body is commonly shown in robotics researches [7,8,14,15]. The evaluation was based on the similarity of the obtained gait to that of a human, based on joint angles and functionalities on rocker function. The start of walking is a very interesting topic, but in this paper it was initiated with the assistance of the experimenter.

3. Results

A part of the results of the rocker function is shown in Fig. 3. The explanation of the rocker function by Perry starts from the initial contact (Fig. 3 (1)). The loading response immediately after the initial contact causes increase of the knee joint flexion by the rapid contraction of the anterior Tibialis anterior muscle. Ongoing Tibialis anterior contraction limits the plantar flexion of the ankle joint and produces a rolling action around the heel as shown in Fig. 3 (2). After the heel rocker, the simulator body rolls forward by the ankle joint with keeping the knee joint flexed at same level over the mid stance as shown in Fig. 3 (3). In the latter half of the ankle rocker into the terminal stance, the knee joint is starting extension while the ankle joint continues rolling forward. This knee joint extension is supported by the ⑨Soleus muscle contraction. Before the falling of the forward foot as shown by the red doted circle in Fig. 3 (9), the contraction of the Gastrocnemius muscle (⑦) raises the heel of the left leg, which called as “forefoot rocker”, to keep the position of the center of mass. Then, the heel of the forward foot falls to the ground which leads the loading response.

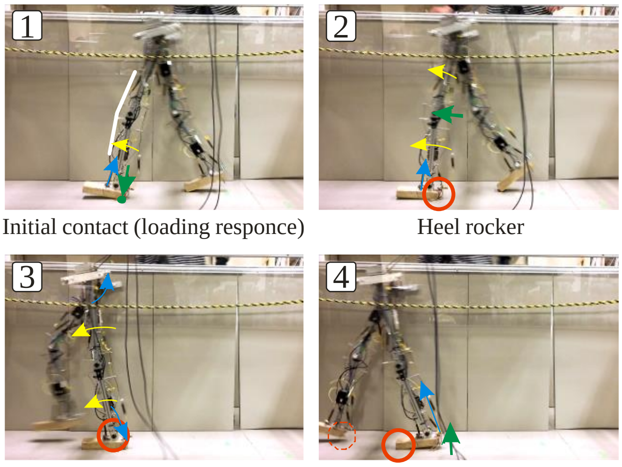


Fig. 3 Sequential photographs of the musculoskeletal walking simulator.

4. Discussion

There are approaches to improve the performance of bipedal robots by applying musculoskeletal configurations and locomotion patterns related to human walking [16]. In this study, a musculoskeletal physical simulator of the lower limb was constructed by replacing muscles with springs. Some parts of the posterior muscles were driven and walked by pulling cables placed in series with the spring. As a result, the robot was able to walk while performing the various actions described in the rocker function description. Because the human knee joint is a roll-on-plate structure and the contact conditions are severe, one possible reason for knee bending in initial contact might be to reduce the impact load on articular cartilage as a biphasic gel layer [17].

We would like to discuss the problem that this walking simulator did not have a foot structure such as a toe structure along with an arch structure. The toe structure is necessary for the forefoot rocker to flex at the metatarsophalangeal joint. In other words, it is thought to have an effect of switching the foot length. At the same time, elastic recoil of the gastrocnemius and flexor digitorum longus is expected. The elastic recoil energy at the end of the stance phase is used for generation of the next swing leg, rather than contributing to the acceleration of the hip. This swing leg generation is called “push-off” [18].

One interesting theme should be how we constructed the gait by trial-and-error method in the physics simulator. At first, we constructed the gait in order from the loading response as Perry’s description. While we noticed the movement on the opposite side, we had to consider the walking motion as a cyclic movement, so we needed to consider the gait as the whole, not in order. It is also interesting to know which muscle should be driven by the motor. Before the three-motor model shown in this report, the authors had confirmed a certain walking in a two-motor model. In these two physical simulators, it was very difficult to generate walking motion without the drive of the single articular muscle for knee flexor (Fig. 2 ⑥). The role is to maintain knee flexion during the swing phase. From a physical point of view, it is thought that the generation of the swing leg should be one of the energy-consuming factors. The importance of the 'push-off' action could also be found here.

In the development of the musculoskeletal walking simulator, the authors felt that the tumbling motion and failure of the simulator was interesting to bring a new idea for walking generation. Currently, there are many issues to be solved, such as actual measurement of COP and long walking time, in addition to the fact that it is a two-dimensional gait and the moment arm is not physiological, but the reproduction of foot structure and push-off will be the next issue to be addressed.

5. Conclusion

In this study, we developed a musculoskeletal robotic walking simulator for the study of biomechanics in rehabilitation. By driving some muscles of the simulator with servomotors, the simulator was able to generate actual walking motion based on the rocker function. Then, the generated walking motion was compared with the motion description of the rocker function. It was confirmed that the simulator was able to generate most of the actions of the rocker function. In addition, the joint angles were compared with the standard human walking motion. In the future, we plan to introduce the forefoot and the arch structure to the musculoskeletal simulator.

Acknowledgements

This work is partly supported by JSPS KAKENHI 18K10747.

References

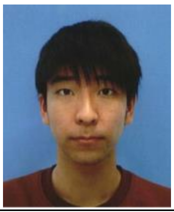
1. N. Yamazaki, Analytical model and simulation of walking, Biomechanism, 3, pp1219–1223, 1974. (in Japanese)
2. I. Kato, The bipedal walking robot - its history and challenges, Journal of the Robotics Society Japan, 1(3), pp4–6, 1983. (in Japanese)
3. M. Vukobratovic, et al., Human and Humanoid Dynamics, Journal of Intelligent and Robotic Systems, 41, pp65–84, 2004.
4. S. Kajita, The humanoid robot, Ohmsha Ltd., 2005. (in Japanese, book: ISBN978-4-274-22602-1)
5. M.A. Sharbafi, A. Seyfarth, Bioinspired Legged Locomotion, Elsevier, OX UK, MA US, Butterworth – Heinemann, 2017. (book: ISBN 978-0-12-803766-9)
6. T. Kinugsa, et al., Recommendation of passive walking robot, Corona publishing Co. Ltd., 2016. (in Japanese, book: ISBN 978-4-339-04649-6)
7. H. Shin, et al., Constructive understanding and reproduction of functions of gluteus medius by using a musculoskeletal walking robot, Advanced Robotics, 32(4), pp202–214, 2018.
8. M.A. Sharbafi, et al., A new biarticular actuator design facilitates control of leg function in BioBiped3, Bioinspiration & Biomimetics, 11, ID046003, 14 pages, 2016.
9. J. Perry, J.M. Burnfield, Gait analysis: normal and pathological function, Second edition, NJ US, SLAK Incorporated, 2010. (book: ISBN 978-1556427664)
10. N. Sakai, et al., A development of joint mechanism of robot arm based on human shoulder morphology, IEEE/RAS-EMBS BioRob 2006, pp982–987, 2006.
11. N. Sakai, et al., Development of robotic musculoskeletal walking simulator using rocker function, Journal of clinical gait analysis, 6(2), pp591–600, 2019. (in Japanese)
12. H. Granzier, et al., Titin: Physiological Function and Role in Cardiomyopathy and Failure, Heart Failure Reviews, 10, pp211–223, 2005.
13. J.A. Monroy, et al., Effects of activation on the elastic properties of intact soleus muscles with a deletion in titin, Journal of Experimental Biology, 220, pp828–836, 2017.
14. S. Soek, et al., Design Principles for Energy -Efficient Legged Locomotion and Implementation on the MIT Cheetah Robot, IEEE/ASME Transactions on Mechatronics, 20(3), pp1117–1129, 2015.
15. D. Hobbelen, et al., System overview of bipedal robots Flame and TUlip: tailor-made for Limit Cycle Walking, IEEE/RSJ IROS 2008, pp2486–2491, 2008.
16. D. Torricelli, et al., Human-like compliant locomotion: state of the art of robotic implementations, Bioinspiration & Biomimetics, 11, ID051002, 20 pages, 2016.
17. N. Sakai, et al., Analysis of biphasic lubrication of articular cartilage loaded by cylindrical indenter, Tribology International, 46, 225-236, 2012.
18. H. Houdijk, et al., Energy storing and return prosthetic feet improve step length symmetry while preserving margins of stability in persons with transtibial amputation, Journal of NeuroEngineering and Rehabilitation, 15(1), 41–48, 2018.

Authors Introduction

Prof. Nobuo Sakai

Associate professor of Mechanical and Control Systems, Kyushu Institute of Technology in Japan. He graduated Intelligent Machinery and Systems, Kyushu University in 2002. His research interests are Robotics, Bioengineering, Tribology.

Mr. Yukiho Ryu



He is currently a Master course student of Mechanical and Control Systems, Kyushu Institute of Technology. He graduated Integrated System Engineering, Kyushu Institute of Technology, 2020.

Mr. Tsubasa Ikeda

He is currently a Master course student of Mechanical and Control Systems, Kyushu Institute of Technology. He graduated Integrated System Engineering, Kyushu Institute of Technology, 2021.

Prof. Masako Fuchi

Professor, Occupational therapist, Department dean of Occupational Therapy, Kyushu Nutrition Welfare University. She graduated National Kyushu Rehabilitation College, 1981. After the working career of Kyushu Rosai Hospital, she graduated English Course of Foreign Studies, University of Kitakyushu. After 1990, she worked in Seiai Rehabilitation Hospital, then Vice-president of the hospital. Also, she received Master degree in Social and Information Science, Nihon Fukushi University in 2000. She moved to the present post in 2016.

Mr. Katsuki Hayashi

Occupational therapist. He graduated National Kyushu Rehabilitation College, 1975. After the working careers of Yamaguchi Rosai Hospital and Kyushu Rosai Hospital, he moved in Seiai Rehabilitation Hospital, then Vice-president of the hospital. He is now Specially Appointed Vice-president of the hospital.

Prof. Mochimitsu Komori

Professor of Electrical Engineering and Electronics, Kyushu Institute of Technology in Japan. He graduated Osaka University in 1984. After the career of National Railway Technical Research Institute, he moved in Kyushu Institute of Technology as an associate professor. Then, he is a professor from 2003.

Collision Avoidance in a Human-Robot Coexistent Food Preparation Environment Using Hands Area Extraction

Takaaki Yotsumoto

Graduate of Engineering, Kyushu Institute of Technology, Japan, Sensuicho 1-1, Tobata, Kitakyushu, Fukuoka 804-8550, Japan

Yuta Ono, Joo Kooi Tan*

**Faculty of Engineering, Kyushu Institute of Technology, Japan, Sensuicho 1-1, Tobata, Kitakyushu, Fukuoka 804-8550, Japan*

E-mail: {Takaaki-yotsumoto , etheltan}@ss10.cntl.kyutech.ac.jp

https://www.kyutech.ac.jp/~etheltan/

Abstract

In Japan, the population of the working-age between 15 and 64 years old is peaked in 1995 at about 87 million and is expected to continue to decline in the future. Therefore, in order to solve the labor shortage, the introduction of industrial robots that can perform the same level of work as humans is strongly requested.especially in a food preparation industry. In order to prevent danger to workers there, it is necessary for industrial robots to recognize workers and avoid them when there is fear of collision. In this paper, we propose a method of extracting hand regions based on the color distributions of a hand and GrabCut in an experimental environment to recognize human hands and detect their directions of approach. The performance of the proposed method is shown experimentally.

Keywords: Recognition, Hand Area Extraction, Lab Color System, Grab cut

1. Introduction

According to an announcement by the Small and Medium Enterprise Agency[1], the working-age population (15-64 years old) in Japan peaked in 1995 at about 87 million and has been declining ever since, reaching about 77 million in 2015. This trend is expected to continue into the future, and it is estimated that the population will decline to about 60% of the 2015 level by 2060. The working population (the total of "employed" and "totally unemployed" among the population aged 15 and over) declined by only about 420,000 between 1995 and 2015, due to an increase in the labor participation rate of women and people aged 65 and over, but this decline is not as large as that of the working-age population. However, it is easy to imagine that the working population will decline in proportion to the decline in the working-age population. For this reason, the introduction of industrial robots, such as robot arms capable of performing tasks equivalent to those of humans, will be necessary to solve the labor shortage.

According to the Industrial Safety and Health Regulations[2] and the Ministry of Labor Notification No. 51 of 1983 (machines specified by the Minister of Health, Labor and Welfare under the provisions of Article 36, Item 31 of the Industrial Safety and Health Regulations) [3], except for those that meet certain standards, in order to use industrial robots, it is basically necessary to enclose them with fences and take other safety measures to prevent humans from entering. However, according to the partial revision[4] of the enforcement notice of Article 150-4 of the Industrial Safety and Health Regulations concerning industrial robots, when it can be assessed that the risk of danger to workers due to contact with industrial robots has been eliminated by implementing measures based on the investigation of danger, etc., joint work between industrial robots and humans can be performed. In order to prevent danger to the worker, it is necessary for the industrial robot to recognize the worker and avoid the worker when necessary.

In this study, as the first step for an industrial robot to avoid a worker, we propose a method to recognize a human hand and to detect the direction of its approach by extracting the hand region.

One of the methods for hand region extraction is to use a distance image[5, 6, 7] to capture and recognize the shape in 3D. However, the infrared camera used in this method has a drawback that it is easily affected by sunlight, which limits the locations where it is used. Other methods that use multiple sensors[8,9] or expensive sensors[10] have been proposed, but they are considered to be difficult to use because of the high price of the system.

2. Summary of Research

This section gives an overview of the system proposed in this paper. The proposed system for detecting the proximity of a human hand or a robot arm recognizes the proximity of a human hand in front of a head-mounted camera attached to the user as shown in **Fig. 1**. Before recognizing the approaching hand, the system obtains the color information of the hand area in the lighting environment where the system is used. Then, the system extracts the hand region from the fingerprint input frame, and identifies the hand from the region.

2.1. *Acquisition of hand area color information*

The proposed system obtains the color information of the hand area in the system environment before judging the proximity of a human hand, and creates the color distribution of the hand at that place and time.

2.2. *Judging the proximity of a human hand*

First, the hand region is extracted from the input image based on the hand region color distribution obtained in

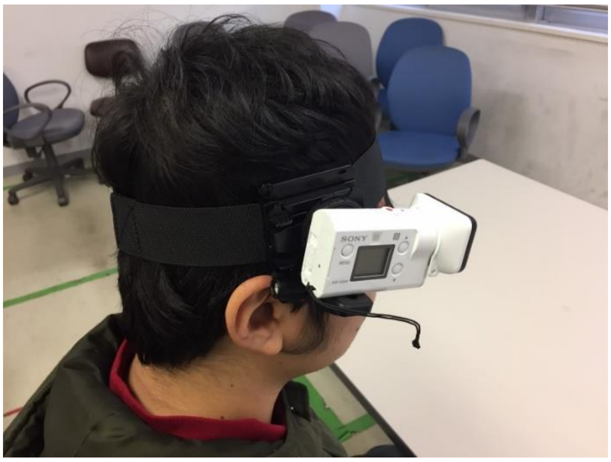


Fig.1. Head mounted camera

the system's operating environment as described in Section 2.1, and noise is removed. After that, the system identifies whether proximity is occurring, and if so, it displays a rectangle on the image where the proximity is occurring as the result of the recognition.

3. Acquisition of Hand Area Information

In this section, we describe a method to acquire color information of the hand region from the captured video.

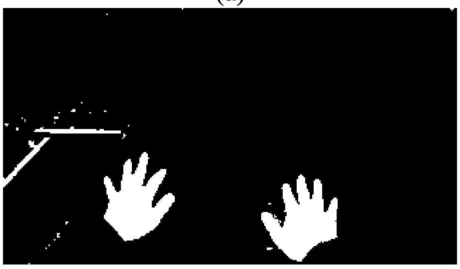
The first frame of the video to acquire the hand area information is the background frame without a hand. After that, while keeping the orientation of the camera unchanged, frames containing several patterns of hand regions are input for a few seconds as shown in **Fig.2**, which are used as the frames for extracting color information.

3.1. *Extraction of candidate hand region candidates*

As shown in **Fig. 3(a)**, most of the images acquired from the head-mounted camera used in this study are the background except for the hand region. Since these areas interfere with the acquisition of the color information of the hand area and are clearly background areas, we delete these areas. We detect the hand region candidates by extracting the difference region between each color information extraction frame and the background frame described above. For each pixel, a foreground region f is extracted to prevent false extraction due to changes in brightness. The $L^*a^*b^*$ color system is used for this process. After that, GrabCut and fine line segmentation

Fig.2. Color information extraction frame

(a)



(b)

Fig. 3 Extraction of a hand region candidate: (a) Input frame, (b) hand region candidates.

are applied to the hand region extraction image for possible improvement in detection accuracy.

Since areas other than skin color are not considered to be hand areas, the skin color areas are extracted by using the HSV color system.

3.2. *Extraction of candidate hand regions*

The color information is extracted from the color information extraction frame using the hand region candidate extracted in Section 3.1 as a mask. To select the color distribution of the hand region, we use the difference between the color distribution of the color information extraction frame and the color distribution of the background frame. The color distribution histograms of the two frames are compared to select the hand region color distribution in the processing frame. Finally, the color distributions in the hand region obtained from the color information extraction frame are combined by logical OR, and the distribution with the largest area is selected to remove noise in the case of selecting a color distribution other than the hand region by mistake. In the proposed method, the above process is used to obtain the color distribution of the hand region according to the usage environment.

4. Judging the Proximity of a Human Hand

In this section, we describe a method for judging the proximity of a human hand from an input image.

To judge the proximity of a human hand, the method identifies whether proximity is occurring, and if so, it displays a rectangle in the image where the proximity is occurring.

4.1. *Extraction of hand regions*

First, the hand region is extracted from the input video. The hand region color distributions of the camera wearer and his neighbors, which were created by acquiring hand region color information, are used to extract the fruits of a^* and b^* values that are included in the color distribution. Then, by using a median filter as a denoiser, the hand region is first extracted from the input video. Using the color distributions of the hand regions of the camera wearer and his neighbors, which are created by acquiring the hand region color information, the real part is extracted, although the values of a^* and b^* are included in the color distribution. After that, a median filter is used for denoising to obtain the hand region extracted image.

4.2. *Judging the proximity of a human hand*

The proposed method judges the proximity of a human hand to the obtained hand region extracted image.

First, labeling is performed, and among the labels obtained by labeling, judgment is made for the labels with an area of a certain area or more.

If the area of the hand is more than a certain area in the left or right danger area, it is judged to be dangerous, because there is approach on the left or the right. The judgment results of the hand region extracted images are ‘no approach’, ‘approach to the right side’, ‘approach to the left side’, ‘approach to both the right and the left side’,

Table 1 Accuracy evaluation

Accuracy evaluation frames[frame]	149
TP [frame]	114
FP [frame]	27
FN [frame]	9
$recall$ [%]	92.68
$precision$ [%]	80.85
F- measure [%]	86.36

'approach to the right side', 'approach to the left side'. There are four patterns of judgments from the camera wearer's side: 'no approach,' 'approach to the right side,' 'approach to the left side,' and 'approach to both the right and the left sides'. method judges the proximity of a human hand to the obtained hand region extracted image. First, labeling is performed, and among the labels obtained by labeling, judgment is made for the labels with an area of a certain area or more.

If the area of the hand is more than a certain area in the left or right danger area, it is judged to be dangerous because there is approach on the left or right. The judgment results of the hand region extracted images are "no approach", "approach to the right side", "approach to the left side", "approach to both the right and left side", "approach to the right side", "approach to the left side. There are four patterns of judgments from the camera wearer's side: "no approach," "approach to the right side," "approach to the left side," and "approach to both the right and left sides.

5. Experimental Results

The system recognizes the presence or absence of approach and the location of approach in images acquired by a head-mounted camera. The input images do not have a single background, and the recognition is performed on 149 images extracted every 10 frames from the input images. The accuracy is calculated by creating a ground truth image and calculating the recall, precision and *F*-measure. As the result of the experiment, recall was 92.68%, precision was 80.85% and *F*-measure was 86.36% as shown in **Table 1**.

6. Conclusion

The causes of false positives are discussed here. First, there are cases where the correct image is not created when viewed in a single frame, but it is created in the previous frame. In this case, the approach detection of a person cannot be performed correctly in a single frame. Second, it is considered difficult to acquire the hand region information every time. Therefore, it is necessary to devise a system that can detect a person, even if the hand region information is acquired only by the person wearing the camera.

References

1. The Small and Medium Enterprise Agency, "2019 White Paper on Small and Medium Enterprises in Japan", The Small and Medium Enterprise Agency, p.121, 2018. (In japanese).
2. Ministry of Health, Labour and Welfare, "Ordinance on Industrial Safety and Health", Ministry of Health, Labour and Welfare, 1972. (In japanese).
3. Ministry of Health, Labour and Welfare, "Ministry of Labor Notification No. 51 of 1983 (Machinery specified by the Minister of Health, Labour and Welfare pursuant to the provisions of Article 36, item 31 of the Industrial Safety and Health Regulations)", Ministry of Health, Labour and Welfare, 1983. (In japanese).
4. Ministry of Health, Labour and Welfare, "Partial Revision of the Enforcement Notice of Article 150-4 of the Industrial Safety and Health Regulations Concerning Industrial Robots", Ministry of Health, Labour and Welfare, 2012. (In japanese).
5. S. Tanaka, A. Okazaki, N. Kato, H. Hino, K. Fukui, "Spotting fingerspelled words from sign language video by temporally regularized canonical component analysis", 2016 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA), 7pages, 2016.
6. K.O. Rodriguez, G.C. Chavez, "Finger spelling recognition from RGB-D information using kernel descriptor", IEEE Conference on SIBGRAPI, 7pages, 2013.
7. N. Pugeault, R Bowden, "Spelling it out: Real-time ASL fingerspelling recognition", IEEE International Conference on Computer Vision Workshops (ICCV Workshops), p.1114-1119, 2011.
8. H. Brashear, V. Henderson, K. H. Park, H. Hamilton, S. Lee, T. Starner, "American sign language recognition in game development for deaf children", Proc. of the 8th International ACM SIGACCESS Conference on Computers and Accessibility, pp.79- 86, 2006.
9. H. Brashear, T. Starner, P. Lukowicz, H. Junker, "Using multiple sensors for mobile sign language recognition", Proc. of the 7th IEEE International Symposium on Wearable Computers, pp.45-52, 2003.
10. Yu Wang, Seiji Itai, Satoshi Ono, Nakayama Shigeru, "Human Recognition with Ear Image by Principal Component Analysis", Journal of Japan Society of

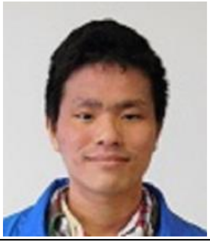
Information and Knowledge, Vol. 16, No. 1, pp.15-27, 2005. (In Japanese)

Authors Introduction

Mr. Takaaki Yotsumoto

He obtained the B.E. degrees in Department of Control Engineering from Kyushu Institute of Technology in 2019, Japan. He is presently a master course student in the same university. His research interests include human detection and recognition.

Mr. Yuta Ono



He received M.E. from Kyushu Institute of Technology. He is now in the Ph.D. course of the Graduate School of Engineering, Kyushu Institute of Technology. His research includes computer vision, machine learning, and rush-out pedestrian detection from videos.

Prof. Dr. Joo Kooi Tan

She is currently with Department of Mechanical and Control Engineering, Kyushu Institute of Technology, as Professor Her current research interests include ego-motion, three-dimensional shape and motion recovery, human detection and its motion analysis from videos. She was awarded SICE Kyushu Branch Young Author's in 1999, the AROB Young Author's Award in 2004, Young Author's Award from IPSJ of Kyushu Branch in 2004 and 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.

Supporting Safe Walk of a Visually Impaired Person at a Railway Station Platform Based on MY VISION

Yuki Kawaguchi*
Department of Mechanical and Control Engineering, Kyushu Institute of Technology, Sensui-cho Tobata-ku Kitakyushu, 804-8550, Japan

Seiji Ishikawa†, Takashi Shinomiya‡, Joo Kooi Tan

Faculty of Engineering, Kyushu Institute of Technology, Sensui-cho1-1 Tobata-ku Kitakyushu, 804-8550, Japan,
‡N&N Inc.
E-mail: {kawaguchi-yuki ,etheltan}@ss10.cntl.kyutech.ac.jp
http://lab.cntl.kyutech.ac.jp/~etheltan/

Abstract

For a visually impaired person, platforms of railway station are places where there is a risk of falling. This paper proposes a system that prevents fall of a visually impaired person from the platform by use of self-viewpoint images provided from MY VISION, an ego camera system attached to a user. The edges of the platform are detected by region segmentation and line segment detection on the self-viewpoint images. The performance of the proposed system was experimentally examined and satisfactory results were obtained.

Keywords: MY VISION, Depth image, Line Segment detection.

1. Introduction

According to the Ministry of Land, Infrastructure, Transport and Tourism (MLIT), an average of 3177 accidents involving falls from platforms of railway station occurred nationwide every year between 2010 and 2016, of which an average of 76 accidents involving person with visibility difficulties occurred annually [1]. According to a questionnaire conducted by a group of visually impaired persons, about 40% of the respondents said that they had fallen from a platform, and about 60% said that they had almost fallen from a platform, making platforms a dangerous place for visually impaired person to move [2]. However, the installation of platform doors requires several hundred million yen per railway station.

The number of railway stations with platform doors in Japan from the end of FY 2006 to the end of FY 2008 was 465. The number of railway stations with platform doors in Japan at the end of fiscal year 2008 was 783, and it is not realistic to install platform doors at all railway stations nationwide. Therefore, it is necessary to develop an inexpensive system that can prevent a person from fall. However, the use of a white cane requires training for a certain period of time and also has the disadvantage of a narrow range of probe. Therefore, in this study, the edge of the platform is detected with the aim of realizing safe movement for a visually impaired person when they use railway stations.

Conventional researches on detecting steps include the methods such as segmenting a region that is in the same plane [3][4], using the Hough transform [5]. In addition, researches using MY VISION include [7][8]. In contrast to these methods, we use an RGB-D camera, Line Segment Detector [9] and Graph Based Segmentation [10] to detect the edge of the platform.

2. Obstacle Detection

Initially, the coordinate system of the input depth image is transformed from the camera coordinate system to the world coordinate system by the following formula.

$$X(x,y)=\frac{(2x\text{-width})}{\text{width}}d\tan\left(\frac{\theta_h}{2}\right) \tag{1}$$

$$Y(x,y)=\frac{(\text{height-}2y)}{\text{height}}d\tan\left(\frac{\theta_v}{2}\right) \tag{2}$$

$$Z(x,y)=d \tag{3}$$

The gradients of Y and Z with respect to the change in y-coordinate and the gradients of X and Y with respect to the change in x-coordinate are then determined as follows;

$$\frac{dY(x,y)}{dy}=Y(x,y)-Y(x,y+k) \tag{4}$$

$$\frac{dZ(x,y)}{dy}=Z(x,y)-Z(x,y+k) \tag{5}$$

$$\frac{dX(x,y)}{dx}=X(x,y)-X(x+k,y) \tag{6}$$

$$\frac{dY(x,y)}{dx}=Y(x,y)-Y(x+k,y) \tag{7}$$

If the angle of the change of the gradient of the Z-coordinate to the gradient of the Y-coordinate in the coordinate (x,y) and the angle of the change of the gradient of the X-coordinate to the gradient of the Y-coordinate are larger than the threshold, respectively, it is judged that there is an obstacle in the area, and the area where the obstacle exists is excluded from the search area of the platform edge. The procedure is formulated by the following equations;

$$d\_YZ(x,y)=\arctan\left(\frac{\frac{dY(x,y)}{dy}}{\frac{dZ(x,y)}{dy}}\right) \tag{8}$$

$$d\_YX(x,y)=\arctan\left(\frac{\frac{dY(x,y)}{dx}}{\frac{dX(x,y)}{dx}}\right) \tag{9}$$

$$object=\begin{cases}1 \ \ if d_{YZ}>th \ \ OR \ \ d_{YX}>th \\ 0\end{cases} \tag{10}$$

3. Estimation of the Ground Area

Graph Based Segmentation (GBS) is applied to the obstacle-free area obtained in the previous step, and the area containing the point g_c (width, height/2) is estimated as the ground area.

GBS is a method of segmenting and merging images into areas with similar features.

4. Platform Edge Detection

The Line Segment Detector (LSD) is used to detect the line segments in the image of the area obtained in the previous step, which is free from obstacles and is the ground. The line segments obtained in this step are merged using the least squares method to detect the line segments that are the edges of the platform.

This line segment is represented by a linear equation $y=ax$ using the coordinates (x,y) in the image, and its slope determines the direction of the edge of the platform as seen by the visually impaired person.

5. Derivation of the Distance to the Edge of the Platform

We refer to the left and the right depths of the pixel g_p on the detected line segment closest to the point g_c (width, height/2), and consider the depth of the point closer to the viewpoint as the distance to the home edge D_{min}^{H}.

$$D=\min_{p=1,2,\dots,p}\left(|g_c-g_p|\right) \tag{11}$$

$$d(x,y,d_e)=D \tag{12}$$

$$D_{min}^{H}=\min_{d_{e^*}}\left(d(x+1,y,d_{e^*}),d(x-1,y,d_{e^*})\right) \tag{13}$$

The derivation of the distances is shown in **Fig. 1**.

6. Experiment

In this experiment, we acquired depth images using a forward-facing RGB-D camera outdoors and detected the edges of the platform using the proposed method. Then, we evaluated the detected straight line by comparing the detected platform edge with the ground truth. The images

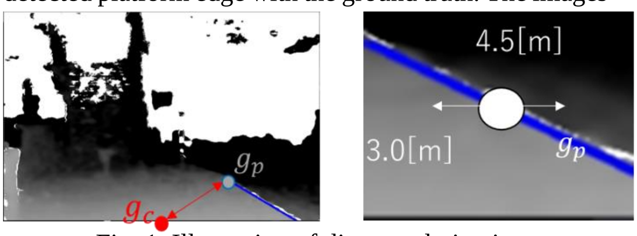
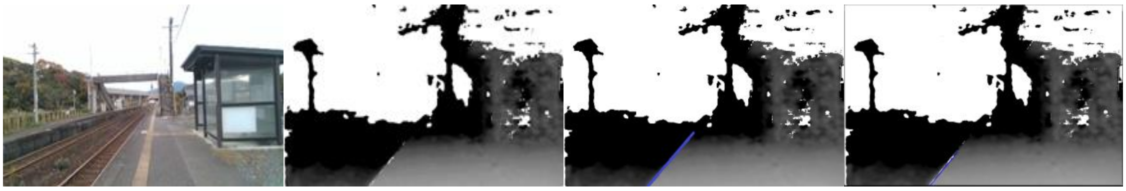


Fig. 1. Illustration of distance derivation.



(a) (b) (c) (d)

Fig. 2. (a) Color image (b) Depth image (c) Ground truth (d) Experimental result. The platform edges and ground-truths detected in the experimental results are marked with blue lines.

used in the experiment were those in which the edge of the platform was in front of or to the left or to the right of the visually impaired person. The images used in the experiment were taken at two railway stations, and the number of images used in the experiment was 88 and 228 at each railway station, giving a total of 316 images. The actual images used in the experiment, the results of the experiment and an example of the ground-truth are shown in **Fig. 2**.

The evaluation of the detected edge is based on the overlap and angle errors defined by equation (14) and equation (15).

$$\boldsymbol{overlap} = \frac{\boldsymbol{GT} \cap \boldsymbol{OA}}{\boldsymbol{GT}} > \boldsymbol{T} \tag{14}$$

$$\boldsymbol{Angular\ error} = |\boldsymbol{a_T} - \boldsymbol{a_d}| < \boldsymbol{T_q} \tag{15}$$

Here *OA* is the area of the home edge detected by the experiment, *GT* is the area of the ground truth, and *T* is the threshold value. In the experiment, we set *T*=0.5, and evaluate the detection success when the overlap value exceeds the threshold, and the detection failure when the overlap value is less than the threshold or cannot be detected. Figure 2 shows an example of an input image, a ground-truth image, and a successful detection. a_T is the angle of the ground-truth, a_d is the angle of the detected line segment, and T_q is the threshold. In this experiment, we set T_q=3°. If the angle error is above the threshold, we consider it false detection, and if it is below, we consider it successful detection.

7. Results

The results of the experiment are shown in **Table 1**.
The percentage of correct position and angle for all images is 92.5% and 81.1%, respectively. The average processing time per image is 664.7ms.

8. Conclusion

This paper proposed a method of detecting the edge of a railway station platform for a visually impaired person using self-viewpoint images, which provides the direction and position of the edge of a platform from the visually impaired person by region segmentation and line detection on depth images acquired by an RGB-D camera. The performance of the proposed method was examined by experiments and satisfactory results were obtained. Further refinement of the method needs to be done so that it may become robust to illumination and weather change, and to the obstacles such as persons on the platform.

Table 1. Experimental results

	Percentage of positions answered correctly[%]	Percentage of angles correct[%]	Average processing time[ms]
Station A	100	80.9	807.4
Station B	90.0	81.2	606.3

References

1. Ministry of Land, Infrastructure, Transport and Tourism, "Study Group for Improving Safety at Railway station Platforms."https://www.mlit.go.jp/tetudo/tetudo_fr7_000015.html,(2019-10-30)
2. Japan Federation of the Blind (now Japan Federation of Organizations of the Visually Impaired), Results of a Questionnaire Survey on Falling Accidents," http://nichimou.org/wp-content/uploads/2014/02/1105tenrakujikoan.pdf, (2019-10-30)
3. Dirk Holz, Stefan Holzer, Radu Bogdan Rusu, Sven Behnke, Real-Time Plane Segmentation using RGB-D Cameras, *Robot Soccer World Cup XV*, pp. 306-317, 2012
4. A.Perez-Yus, D.Gutierrez-Gomez, G.Lopez-Nicolas, J.J.Guerrero, Stairs detection with odometry-aided traversal from a wearable RGB-D camera, Computer Vision and Image Understanding, Vol.154,pp.192-205,2017.
5. Shuihai Wang, Hangrong Pan, Chenyang Zhang, Yingli Tian, RGB-D image-based detection of stairs, pedestrian crosswalks and traffic signs, Journal of Visual Communication and Image Representation, 25(2), pp.263-272, 2014.
6. Kohei Kitagawa, Seiji Ishikawa, Joo Kooi Tan, Development of a Pedestrian Crossing Navigation System for a Visually Impaired Person Using MY VISION, Proceedings of International Conference on Artificial Life & Robotics (ICAROB2021), pp. 283-286, 2021.
7. Joo Kooi Tan, Tomoki Ishimine and Shohei Arimasu, Walk Environment Analysis Using MY VISION: Toward a Navigation System Providing Visual Assistance, International Journal of Innovative Computing, Information and Control ,Vol 15,No. 3, PP.1-11, 2019.
8. Rafael Grompone von Gioi, Je ́re ́mie Jakubowicz, Jean-Michel Morel, Gregory Randall, LSD: a Line Segment Detector, Image Processing On Line , Vol.2,pp.35-55.
9. Pedro F. Felzenszwalb, Daniel P. Huttenlocher, Efficient Graph-Based Image Segmentation, International Journal of Computer Vision, Vol.59, pp.167-181, 2004.

Authors Introduction

Mr. Yuki Kawaguchi

He received his B.E. degree in Department of Control Engineering in 2020 from the Faculty of Engineering, Kyushu Institute of Technology in Japan. He is acquiring the M.E. in Kyushu Institute of Technology.

Emeritus Prof. Seiji Ishikawa

He graduated from Tokyo University and was awarded BE, ME and PhD there. He is now Emeritus Professor of Kyushu Institute of Technology. He was Visiting Researcher of the University of Sheffield, UK, and Visiting Professor of Utrecht University, NL. His research interests include visual sensing & 3-D shape/motion recovery. He was awarded The Best Paper Award in 2008, 2013 and 2015 from Biomedical & Fuzzy System Association, Japan. He is a member of IEEE.

Dr. Takashi Shinomiya

He graduated from Nihon University where he was awarded B.E. He obtained the Ph.D. degree in Kyushu Institute of Design (current Kyushu University). Dr. Shinomiya is Special Advisor of N&N Co. Ltd, Tokyo, and Invitation Researcher of Waseda University, Tokyo. His research interests include information technology of visual perception.

Prof. Dr. Joo Kooi Tan

She is currently with Department of Mechanical and Control Engineering, Kyushu Institute of Technology, as Professor Her current research interests include ego-motion, three-dimensional shape and motion recovery, human detection and its motion analysis from videos. She was awarded SICE Kyushu Branch Young Author's in 1999, the AROB Young Author's Award in 2004, Young Author's Award from IPSJ of Kyushu Branch in 2004 and 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.

Detecting a Pedestrian's Walk Direction Using MY VISION for Supporting Safe Walk of a Visually Impaired Person

Shinya Iizumi*
Graduate of Engineering, Kyushu Institute of Technology, 1-1 Sensui-cho, Tobata, Kitakyushu Fukuoka, 804-8550, Japan

Joo Kooi Tan*, Yuta Ono, Seiji Ishikawa, Masuhiro Nitta
Faculty of Engineering, Kyushu Institute of Technology,, 1-1 Sensui-cho, Tobata, Kitakyushu Fukuoka, 804-8550, Japan
E-mail:{iizumi-shinya, etheltan}@ss10.cntl.kyutech.ac.jp
https://www.kyutech.ac.jp/~etheltan/

Abstract

In this paper, we propose a method of recognizing multiple objects using MSC-HOG (Multiple-Scale-Cell Histograms of Oriented Gradients) features and intensity models of both pedestrians and bicyclists. We also propose a method of detecting approaching passersby using different discriminators without using time-series information such as Optical Flow. The effectiveness of the proposed method is verified by experiments.

Keywords: MSC-HOG, Passerby Detection, MY VISION, Visually impaired person, Walk support

1. Introduction

"White canes" and "guide dogs" are two main tools for helping a visually impaired person walk safely[1]. However, the former has a narrow range in recognizing surrounding objects, and the latter has the problem of insufficient numbers of dogs. Therefore, a system has been proposed called MY VISION[2] to recognize surrounding pedestrians and their approaching directions, and give an alarm to the user if pedestrian approaches him/her, thereby it is possible for the user avoiding collision. However, since the MSC-HOG (Multiple-Scale-Cell Histograms of Oriented Gradients) features [3, 6] used in the literature[2] consider only a single model image (a pedestrian model), there is a problem that only one type of object (e.g. a pedestrian) can be recognized. Another problem is that, since they use an optical flow technique to detect a pedestrian, the recognition rate of approaching directions is low.

In this paper, we propose a method of recognizing multiple objects using MSC-HOG features using the models of both pedestrians and bicyclists. In addition, we also propose a method of detecting approaching passersby who are more dangerous to a visually impaired person. We verify the effectiveness of the proposed method using the images of passersby (pedestrians and bicyclist) in our own database, and the pedestrian images in INRIA Person Dataset.

2. Outline of the Proposal Method

The proposed method consists of a learning part and a discrimination part. As shown in **Fig.1**, we define three directions by which the surrounding passersby approach a user i.e., a visually impaired person. They are the front left, the front and the front right direction, expressed by the 10 o'clock, 12 o'clock, and 2 o'clock direction, respectively. In the learning part, we first train a three-class

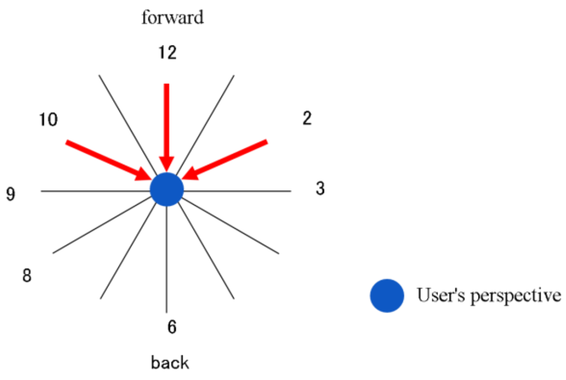


Fig.1 Defining the direction of approaching pedestrians perceived from a visually impaired person.

classifier of passersby approaching from the three directions mentioned above using MSC-HOG features. After known the direction, it is used as a risk index to give alarm to the user. For example, a pedestrian at the 2 o'clock direction walking towards the 10 o'clock direction is not a target of the recognition, whereas, a pedestrian at the 2 o'clock direction walking towards the 8 o'clock direction might be high risk to the user.

By recognizing only those pedestrians approaching from the three specific directions, we can recognize both the pedestrian and the approaching direction at the same time. In the proposed method, we create an average pedestrian model and an average bicyclist model, and then integrating the two models to obtain a model (hereafter referred to as passerby model) according to the three directions. We employ MSC-HOG feature to extract the features from a fed image using the three directional passerby model and train a Random Forest discriminator using the features. Having obtained an effective discriminator of each passerby direction, we apply it to the recognition of a passersby approaching from the three directions.

3. Proposed Method

In this section, some techniques contained in the proposed method are described.

3.1. *MSC-HOG features*

We use the MSC-HOG feature, which is an improved method from the conventional HOG (Histograms of Oriented Gradients). features[4]. In feature extraction, the position and the size of the rectangular cells of HOG are fixed, in contrast, they are variable with the MSC-HOG features.

3.1.1. *Creating a passerby model*

In MSC-HOG feature extraction, it is necessary to create a model of the recognition target in advance for feature extraction. In this paper, we propose a passerby model for recognizing both pedestrians and bicyclists. The MSC-HOG features mainly require the edge information at the cells placed on the target. So we apply Sobel Filter to grayscale images to obtain the edge images. Then, the average of each image is calculated by Eqs.(1)and(2), and a passerby model is created.

$$\overline{I_p}(x,y) = \frac{1}{N}\sum_{i=0}^{N} I_p(x,y) \quad (1)$$

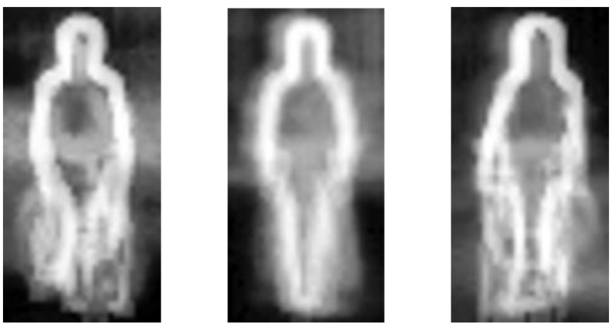
$$\overline{I_r}(x,y) = \frac{1}{N}\sum_{i=0}^{N} I_r(x,y) \quad (2)$$

Here $\overline{I_p}(x,y)$ is the pedestrian model and $I_p(x,y)$ is the differential image of the pedestrian, $\overline{I_r}(x,y)$ is the bicyclist model and $I_r(x,y)$ is the differential image of the bicyclist, and N is the number of images.

The average of each image calculated by Eqs.(1) and(2) is used to make a passerby model. Three passerby models are created, i.e., a model from 2 o'clock, a model from 12 o'clock, and a model from 10 o'clock directions. **Figure 2** shows the created models.

3.1.2. *Cell placement*

Based on the passerby model, we place cells for feature extraction. In order to obtain the edge of the target in more precisely, cells are placed densely at the locations where the edge strength of the passerby model is strong, and cells are placed sparsely at the locations where the edge strength is weak. The set of cells used for feature extraction is represented as *CELL* by Eqs(3)-(5).



(a) (b) (c)

Fig.2 Passerby Model. Passersby approaching from(a) 10 o'clock, (b) 12 o'clock, and(c) 2 o'clock direction.

$$CELL = \left\{ \begin{matrix} (cx, cy, w, h | cell(cx, cy, w, h) \subset window(W, H)) \\ w = a + ci, h = b + dj, \forall i \forall j (i, j \geq 0) \end{matrix} \right\} \quad (3)$$

$$cell(cx, cy, w, h) = \left\{ \begin{matrix} (x, y) | cy - \frac{h-1}{2} \leq y < cy + \frac{h+1}{2} \\ cx - \frac{w-1}{2} \leq x < cx + \frac{w+1}{2} \end{matrix} \right\} \quad (4)$$

$$window(w, h) = \{(x, y) | 0 \leq y < h, 0 \leq x < w\} \quad (5)$$

Here, cx, cy, w, h are the *x*- and *y*-coordinates of the center of the cell, the horizontal and vertical sizes of the cell, respectively: W, H a, b are the horizontal and vertical sizes of the detection window, the minimum values of the horizontal and vertical sizes of the cell, respectively: $c_{i,}$ and c_j are the horizontal and the vertical sizes of the cell, respectively. $c_{i,}$and c_j vary depending on the luminance of the passerby model.

3.4. *Computation of Feature Values*

In MSC-HOG features, the gradient direction and gradient intensity are calculated in the same way as in the conventional method of HOG[4] features, and a histogram is created. However, normalization is done in each cell. For the gradient vector, the gradient direction from 0° to 180° is divided into 9 directions with 20° each, and a histogram of 9 bins is created.

4. Experimental Results

We conducted two types of experiments. Experiment 1: Confirming the effectiveness of the passerby model, Experiment 2: Confirming the effectiveness of the direction of passerby.

In experiment 1, we use three direction passerby images as a positive dataset, and also use the images that do not included passersby as a negative dataset. In experiment 2, since our purpose is to recognize only the passersby approaching from the specified direction to a user, we choose one of the specified passersby directional images as a positive dataset, and we put the other two directional images as negative dataset, so that we can obtain a specified direction discriminator. In both experiments, Random Forest was used as the discriminator, and a total of three different approaching direction discriminators were created for recognizing passerby and their approaching directions.

For the passerby and directional learning images, we used 3000 images of a passerby approaching from the 10 o'clock direction, 1752 images of a passerby approaching from the 12 o'clock direction, 3000 images of a passerby approaching from the 2 o'clock direction, and 6000 negative images. For the learning images, we used our own dataset and the INRIA Person Dataset[5].

For the test images, 500 pedestrians in the 10 o'clock (referred to as group A in the experiment 1), 12 o'clock (referred to as group B in the experiment 1) and 2 o'clock (referred to as group C in the experiment 1) directions and 500 bicyclists in each direction were collected and used as a positive test data, and 1000 negative images from the INRIA Person Dataset were used as negative test data.

The results of the both experiments are shown in **Table.1** and **Table.2,** respectively. Recall, Precision, and F-values are used to evaluate the proposed method.

Table.1 Accuracy of the passerby detection (Experiment 1)

	group A	group B	group C
$Recall$	0.961	0.962	0.933
$Precision$	0.980	0.999	0.987
F	0.970	0.980	0.959

Table.2 Accuracy of the passerby approaching direction (Experiment2)

	Directions	10 o'clock	12 o'clock	2 o'clock
Pedestrian	$Recall$	0.788	0.670	0.776
	$Precision$	0.999	1.000	0.979
	F	0.881	0.802	0.866
Bicyclists	$Recall$	0.636	0.922	0.774
	$Precision$	0.998	1.000	0.979
	F	0.777	0.959	0.865
Average	$Recall$	0.712	0.796	0.775
	$Precision$	0.999	1.000	0.979
	F	0.831	0.886	0.865

5. Discussion

In this study, the goal is to help visually impaired people walk safely. It is therefore, important to avoid failing of recognition to approaching passersby. For this reason, among the three evaluation indexes shown in Table1&2, the Recall value is the most important index that needs to

be noticed. From Table.1, the minimum Recall value was 0.93 with the respect to the data group A, B, and C. This indicates the effectiveness of the proposed MSC-HOG feature and the passerby model.
From Table.2 the precision results are high, whereas, the value of Recall of a passerby is 0.76 in average.

The recall of the bicyclist was 92.2[%] in the 12 o'clock direction, but 63.6 [%] and 77.4 [%] in the 10 o'clock and 2 o'clock directions, respectively. One of the reasons of this lower accuracy in direction 10 and 2 o'clock may be the symmetric figure of a bicycle. On the other hand, the recall of pedestrians approaching directions were 67[%], 78.8[%], and77.6[%] in 12,10, and 2 o'clock, direction, respectively. Adversely, the 12 o'clock direction obtained a lower accuracy compared to the two other directions. Since the direction changes along time, the recall rate of pedestrian and bicyclist directions detection recall rate can be improved by adding another characteristics of the direction of movement such as Optical Flow.

6. Conclusion

In this paper, we have proposed a method for recognizing pedestrians that including bicyclist and their approaching directions using the MY VISION system. Unlike conventional method[2], we considered and detected not only pedestrians but also bicyclists and their approaching directions. For this purpose, MSC-HOG with a passerby model was introduced. We also proposed a method of recognizing the direction of approaching passersby using three different directions discriminators.

The effectiveness of the propose was confirmed by two experiments that focused on the recognition of passersby and also on the recognition of three main approaching directions. Further improvement on the recognition of passerby approaching directions, especially, 10 o'clock and 2 o'clock directions of bicyclists remains for future work.

References

1. "Social Welfare Corporation Japan Braille Library official website", https://www.nittento.or.jp/news/koekake.html
2. Joo Kooi Tan, Kenta Hori, Seiji Ishikawa, "Detecting pedestrians and their walk directions using a MY VISION system", Journal of Robotics, Networking and Artificial Life, vol.7-2, pp. 137-141, 2020.
3. Y. Ehara, J.K. Tan, H. Kim S. Ishikawa, "Human detection from an image employing reduced number of weak-classifiers", Proceedings of the IIEEJ Image Electronics and Visual Computing Workshop, IEVC, Kuching, Malaysia, 2012, pp. 1–5.
4. N. Dalal, B. Triggs, C. Schmid, "Human detection using oriented histograms of flow and appearance", Proceedings of European Conference on Computer Vision (ECCV), Lecture Notes in Computer Science, vol. 3952, Springer, Berlin, Heidelberg, 2006, pp. 428–441.
5. "INRIA Person Dataset", http://pascal.inrialpes.fr/data/human/.
6. Yuta Ono, Joo Kooi Tan, Akitoshi Sato, "Finding a rush-out human employing a human body direction detector", Proceedings of International workshop on Advance Image Technology (IWAIT 2019) , 2019, PP.1-6.

Authors Introduction

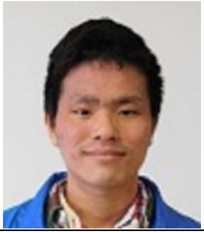
Mr. Shinya Iizumi

He received his B.E. degree in Department of Control Engineering in 2020 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, pattern recognition, and robotics.

Prof. Dr. Joo Kooi Tan

She is currently with Department of Mechanical and Control Engineering, Kyushu Institute of Technology, as Professor Her current research interests include ego-motion, three-dimensional shape and motion recovery, human detection and its motion analysis from videos. She was awarded SICE Kyushu Branch Young Author's in 1999, the AROB Young Author's Award in 2004, Young Author's Award from IPSJ of Kyushu Branch in 2004 and 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.

Mr. Yuta Ono



He received M.E. from Kyushu Institute of Technology. He is now in the Ph.D. course of the Graduate School of Engineering, Kyushu Institute of Technology. His research includes computer vision, machine learning, and rush-out pedestrian detection from videos.

Emeritus Prof. Seiji Ishikawa

He graduated from Tokyo University and was awarded BE, ME and PhD there. He is now Emeritus Professor of Kyushu Institute of Technology. He was Visiting Researcher of the University of Sheffield, UK, and Visiting Professor of Utrecht University, NL. His research interests include visual sensing & 3-D shape/motion recovery. He was awarded The Best Paper Award in 2008, 2013 and 2015 from Biomedical & Fuzzy System Association, Japan. He is a member of IEEE.

Dr. Masuhiro Nitta

He received his PhD from Nara Institute of Science and Technology. He is now an assistant professor in Faculty of Engineering at Kyushu Institute of Technology in Japan.

Fruits and Vegetables Detection using the Improved YOLOv3

Changhua Xu*
Department of Engineering, Kyushu Institute of Technology
Sensui1-1, Tobata,804-8550, Kitakyushu, Fukuoka, Japan

Ziyue Liu
Department of Engineering, Kyushu Institute of Technology
Sensui1-1, Tobata,804-8550, Kitakyushu, Fukuoka, Japan

Joo Kooi Tan
Faculty of Engineering, Kyushu Institute of Technology
Sensui1-1, Tobata,804-8550, Kitakyushu, Fukuoka, Japan
E-mail: {xu-changhua, liu-ziyue, etheltan }@ss10.cntl.kyutech.ac.jp

Abstract

As the global aging intensifies, it is more convenient for a robot to go for buying things like fruits and vegetables instead of elderly, and it is more human-like to select items according to a user's personal preferences such as maturity of fruits, sweetness, etc. However, Fruits and vegetables are generally displayed in a disorderly manner. Therefore, detection and recognition of fruits and vegetables is a difficult task for a robot. This paper proposes an improved YOLOv3 and also pre-training the networks to detect fruits and vegetables, we then using Bilinear-CNN to classifyfruit's maturity. The effectiveness of the proposed method is shown by experiments.

Keywords: Deep learning, Neural Network, YOLOv3, detection, maturity classification.

1. Introduction

With the development of robotics and the intensification of global aging, it has become a trend to use a robot to help elderly persons in shopping task. While doing shopping, a robot which can choose items based on personal preferences is thought to be more human-like. Under this approach, the detection, recognition and classification abilities have become an important task in the whole shopping process.

About research on target detection algorithms, it began with some classic algorithms which are mainly divided into three stages: (1) Finding a candidate area on an image, (2) extracting features from the candidate area, and (3) using the features for training and then creating a classifier for recognition or classification. Zhang & Wu[1] investigated different multi-class Kernel SVMs with appearance descriptors for fruit classification. It used a combination of color, texture and shape features as descriptors to build model and got an accuracy of 88.2% on their dataset. Rocha *et al*[2] proposed a unified approach that can combine many features and classifiers to achieve a 97% accuracy of classification to the supermarket fruit data set. The conventional algorithms, which is simple though, costs a lot of time for data processing. And, due to artificial feature, the model has a poor generalization

With the development of computer equipment, the calculation of an algorithm employing a huge image database has greatly improved image detection technology not only in an accuracy but also in the terms of process speed. In this way, deep learning was proposed. Series of R-CNN algorithms[3-5] are classic methods for detection, and Faster-RCNN[5] is the Upgraded version of R-CNN algorithm. YOLO[6] was presented in 2016, it improved the speed of detection to 45 f/s at the expense of accuracy. The employed method of combining classification and localization tasks is believed to have a profound impact on solving object detection problems. Owing to the shortcomings of YOLO, the authors[6] have made a series of improvements to the algorithm and provided a 76.8% of mean average precision (mAP) in YOLOv3[7,8]. For fruit classification tasks based on deep learning, Hossain *et al*[9]. proposed two models in which a light CNN model and a VGG-16 fine-tuned model have achieved better performance than the previous studies [2,10].

In this paper, we create a dataset of fruits and vegetables and modify a new network based on YOLOv3 for detection. After that, we use Bilinear-CNN for classifying the maturity of fruits and vegetables, so that robot can choose those maturated fruit or vegetables according to a user's preference.

2. Methods

2.1. Creating *datasets*

Currently, a public training dataset of fruit and vegetable is not available. In this paper, we created two kinds of dataset, *i.e*., fruit and vegetable dataset and their maturity dataset(banana) for detection and classification. As for detection task, we collected 15 kinds of fruits and vegetables pictures from Internet and got totally 1840 pictures. As for classification task, we classified 300 banana pictures into 5 different maturity levels. Examples of 15 classes of fruits and vegetables pictures are shown in Fig. 1.

In addition, in order to optimize the generalization performance of detection process and an anti-interference ability in the real scene, we use gamma correction to change the brightness and saturation, and also add White Gaussian Noise to the pictures (dataset) for data augmentation.

2.2. An *improved network of feature extraction*

In YOLOv3, Darknet-53 is used for the feature extraction. There are a total of 5 blocks in darknet-53, and each block from top to bottom of the network has 1, 2, 8, 8, and 4 units. Since there are certain differences in the detail of texture within and between fruits or vegetables categories, so we need more shallow feature information to improve the deep network for detection. Instead of 1- 2-8-8-4 unit's model in the original YOLOV3, the unit of each block of the proposed method is optimized to obtain a sequential unit number of 3-3-6-8-4-unit model. On the other hand, for each unit, we use two 3*3 convolutions and one 1*1 convolution to form a residual block. In this way, the proposed method not only increases the depth of the network layer to a certain extent, but also strengthens the expression of detailed features. In order to test the performance of the proposed unit model, we also modified the units to a 4-5-6-8-4-unit model for comparison. **Figure 2** shown the proposed improved network.

Fig. 1. Fifteen classes of fruits and vegetables image examples

2.3. *Maturity classification*

We use Bilinear-CNN to classify maturity of fruit. In the classification process, the maturity feature is extracted using two VGG-16 networks. After detection, in order to eliminate the background effect and to improve the classification accuracy, we trim the targets from the detected images and drop them to the maturity networks for classification. The process is shown in the **Fig.3.**

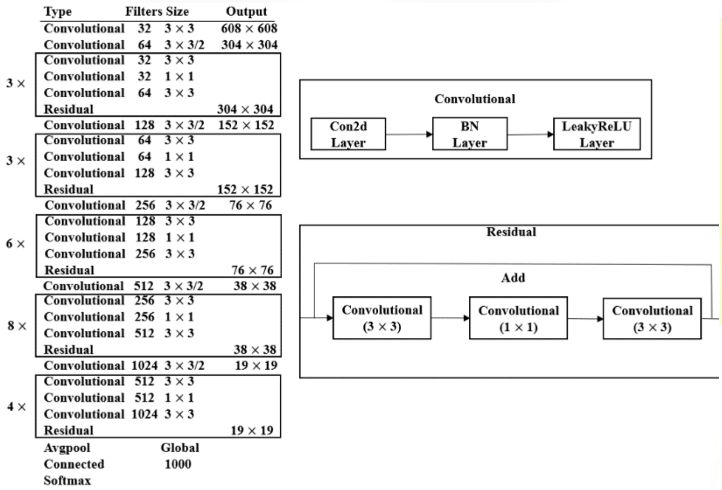


Fig. 2. The improved network of darknet-53

3. Experiments

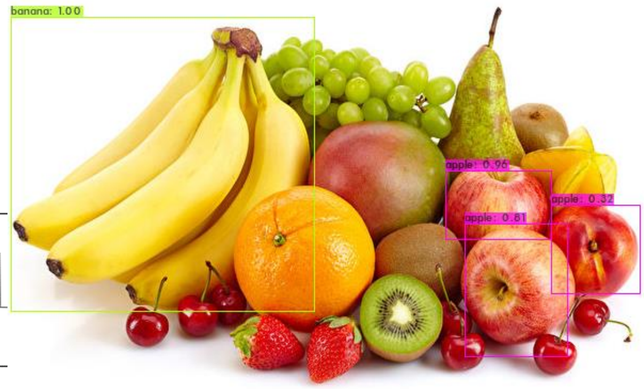
3.1. The performance of different units blocks model

In order to compare the performance of the proposed method with different kind of unit blocks, we used three kind of unit blocks network models in the experiment. Through comparative experiments, we evaluated the detection results of fruits and vegetables by the dataset mentioned in section 2.1. In the experiments, we used 1440 images for training and 400 images for test. For each class, we calculate the Average Precision (AP) to judge the accuracy of different unit model. The experimental results are shown in **Table 1.**

As shown in Table 1, the accuracy of the two proposed units has improved in many classes. especially among small targets such as garlic, potatoes, strawberries and cherry tomato. We also found that the result of 3-3-6-8-4 unit is superior then to 4-5-6-8-4 unit in the most classes.

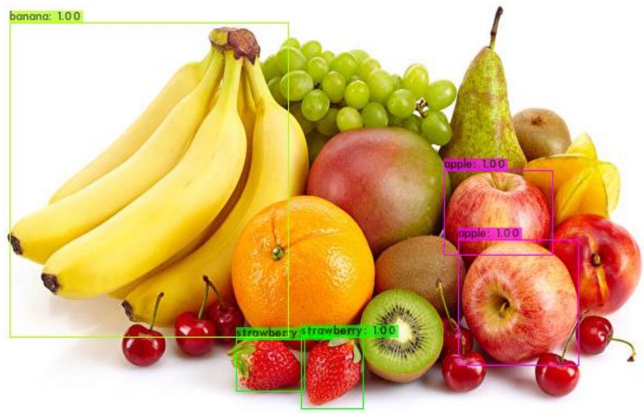
From the comparison of the results, we can get a more detailed effect difference, as shown in Fig.**4**, from which we can see that the network with a unit number of 4-5-6-8-4 eliminated false detections and was able to detect one strawberry. However, the unit number of 3-3-6-8-4 detected not only the banana and apple but also the two small strawberries correctly.

The mAP of the original 1-2-8-8-4unit model, 4-5-6-8-4-unit model come to be 79.04% and 81.41% respectively, the proposed 3-3-6-8-4-unit model gets a better result of 83.72%.



(a) Original 1-2-8-8-4

(b) 4-5-6-8-4



(c) 3-3-6-8-4

Fig. 4. The detection result of three kinds of unit model: (a)Original network, (b) network with a unit number of 4-5-6-8-4, (c) proposed network with a unit number of 3-3-6-8-4.

Table 1. The comparison of the detection result on our data set

Number	Classes	YOLOV3 1-2-8-8-4 [%]	YOLOV3 4-5-6-8-4 [%]	YOLOV3 3-3-6-8-4 [%]
01	apple	72.3	84.2	85.8
02	banana	76.3	79.4	83.1
03	corn	84.0	86.1	89.3
04	dragon fruit	96.2	96.3	98.5
05	eggplant	91.7	74.7	76.7
06	garlic	77.0	86.1	87.3
07	honeydew	93.0	96.6	96.6
08	onion	65.7	68.8	71.5
09	peach	78.6	65.5	79.3
10	pear	69.7	66.8	67.6
11	bell pepper	77.6	81.3	80.7
12	potato	62.5	74.2	79.2
13	strawberry	73.6	89.5	87.2
14	cherry tomato	71.5	84.0	85.7
15	watermelon	95.8	87.8	87.3

3.2. Result of maturity classification

Table 2. The result of banana maturity classification

Maturity label	Number of pictures	Correct number of pictures	Accuracy [%]
0	47	47	100.0
1	33	28	84.8
2	40	25	62.5
3	30	22	73.3
4	55	53	96.4

In the maturity classification, we only consider "banana maturity". Totally 205 pictures of banana are used in the experiments. **Table 2** shown the result of maturity classification.

We trimmed the targets from the detected images to eliminate the influence of the background and achieved an accuracy of 85.37% with our test dataset. However, maturity label 2 did not get a good result. This is because some trimmed images have lower brightness, resulting in some misclassification images in the maturity label 2.

4. Conclusion

In this paper, we adopted an effectively improved YOLO algorithm to detect fruits and vegetables mainly by considering the balance of the units in network, and by adding another 3*3 convolution based on YOLOv3 to the residual block to improve the detection accuracy. As the result, we achieved the mAP of 83.72% of detection. For the classification task, we divided the bananas into five categories according to their maturity for training, and trimmed the detection result for classification, finally we obtained the average accuracy of 85.4%.

Elimination of the background noises from the detected images needs to be done. Practical use of the proposed method in more complicated environments also remains for future works.

References

1. Y. Zhang and L. Wu, "Classification of fruits using computer vision and a multiclass support vector machine," Sensors, vol. 12, no. 9, pp. 12489- 12505, Sep. 2012.
2. A. Rocha, D.C. Hauagge, J. Wainer, S. Goldenstein, "Automatic fruit and vegetable classification from images," Computers and Electronics in Agriculture, vol. 70, no. 1, pp. 96-104, Jan. 2010.
3. Girshick R, Donahue J, Darrell T, et al: Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (2014)
4. Girshick R. Fast R-CNN[J]. Computer Science. (2015)
5. Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE Transactions on Pattern Analysis & Machine Intelligence. 39(6), 1137(2017)
6. Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi: You Only Look Once: Unified, Real-Time Object Detection. 779-788(2016)
7. Joseph Redmon, Ali Farhadi: YOLO9000. Better, Faster, Stronger. IEEE Conference on Computer Vision and Pattern Recognition (CVPR).6517-6525(2017).
8. Redmon J, Farhadi A: YOLOv3: An Incremental Improvement[J]. arXiv.org (2018)
9. M. Shamim Hossain, Muneer Al-Hammadi, and Ghulam Muhammad: Automatic Fruit Classification Using Deep Learning for Industrial Applications. IEEE Transactions on Industrial Informatics.15(2), pp. 1027 – 1034(2019).
10. S. R. Dubey and A. S. Jalal, "Robust approach for fruit and vegetable classification," Procedia Engineering, vol. 38, pp. 3449-3453, 2012.

Authors Introduction

Mr. Changhua Xu

He received his bachelor degree of Vehicle Engineering, China in 2018, now specializing in Control Engineering, Kyushu Institute of Technology, Japan. At present, his area of research is the development of human-machine collaboration system based on camera images.

Mr. Ziyue Liu

He received his 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 obtained the M.E. degree in Department of Control Engineering in 2020 from the Graduate School of Engineering, Kyushu Institute of Technology in Japan. He is currently a Ph.D. candidate in the same University. His research interests are image processing, pattern recognition, and robotics.

Prof. Dr. Joo Kooi Tan

She is currently with Department of Mechanical and Control Engineering, Kyushu Institute of Technology, as Professor Her current research interests include ego-motion, three-dimensional shape and motion recovery, human detection and its motion analysis from videos. She was awarded SICE Kyushu Branch Young Author's in 1999, the AROB Young Author's Award in 2004, Young Author's Award from IPSJ of Kyushu Branch in 2004 and 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.

Human-vehicle detection based on YOLOv5

Zhihui Chen[1], Xiaoning Yan[2], Shuangwu Zheng[2], Xiaoyan Chen*[1]

[1]Tianjin University of Science and Technology, China;

[2]Shenzhen Softsz Co. Ltd., China;

E-mail: 1594838831@qq.com

www.tust.edu.cn

Abstract

With the continuous improvement of social development level, traffic has become complicated. Therefore, the detection of people and vehicles becomes important. There are many application scenarios for human-vehicle detection, such as autonomous driving and transportation. This paper mainly introduces the research status of human-vehicle detection, analyzes the advantages and disadvantages of various current target detection algorithms, and focuses on YOLOv5 algorithm. Because the YOLOv5 model is much smaller than YOLOv4, and YOLOv5 also has strong detection ability. Finally, YOLOv5 is used to carry out human-vehicle detection experiments. The results the detection accuracy is improved slightly.

Keywords:human, vehicle, detection, YOLOv5

1. Introduction

At present, target detection has not only received a lot of research in academia, but also has been widely used in real life, such as video fire detection, unmanned driving, security monitoring, and drone scene analysis. Target detection algorithms are mainly divided into three categories: traditional target detection algorithms, classifier-based detection algorithms and regression-based detection algorithms.

In terms of pedestrian and vehicle detection. Initially, the pedestrian detection algorithm mainly used independent single features for feature selection in videos and pictures, and it was mainly aimed at pedestrians in social streets[1]. The most commonly used feature selection in China is the histogram of directional gradients (HOG). Its core point is that some appearances and features of the detected object can be better represented by the distribution of gradient and edge direction. Traditional foreign target detection methods mainly explore target classification, such as how to distinguish pedestrians from other objects in the street and how to use classifiers reasonably. In vehicle detection, the more commonly used method is for researchers to establish information and non-redundant derived values from the initial measurement data. The advantage of this method is to promote learning and improve the generalization ability of the model, and it can also bring better interpretability.

2. Target Detection Algorithms

Target detection algorithms can be divided into traditional target detection algorithms, classifier-based detection algorithms and regression-based detection algorithms.

2.1. *Traditional target detection algorithm*

The SIFT algorithm can extract its invariant characteristics from the features when the image proportion and rotation are unchanged.

AdaBoost is an iterative algorithm that adds a new weak classifier in each round until it reaches a predetermined and sufficiently small error rate.

2.2. *Detection algorithm based on classifier*

OverFeat is an early stage one-stage target detection method. This method mainly discusses multiscale and sliding window methods. It can be used in a feedforward network that includes convolution calculations and has a deep structure. This method can better locate the target by determining the boundary information of the target in advance.

R-CNN (Region with CNN features) algorithm is a type of recurrent neural network that takes sequence data as input, recursively in the evolution direction of the sequence, and all nodes (recurrent units) are connected in a chain.

Fast-RCNN algorithm is established on the basis of previous research. It uses a feedforward neural network that includes convolution calculations and has a deep structure, which can be used to effectively classify candidate targets. Compared with previous research, Fast-RCNN has many innovations, which improves the speed of training and testing while also improving the accuracy of detection.

2.3. *Regression-based detection algorithm*

YOLO (You Only Look Once) is an algorithm that can solve bounding boxes with accurate predictions while using convolutional sliding windows[2]. The core idea of YOLO is to transform target detection into a regression problem, using the entire image as the input of the network, and only going through a neural network to get the location of the bounding box and its category. The YOLO algorithm can realize real-time detection.

Compared with the YOLO algorithm, the SSD (Single Shot MultiBox Detector) algorithm directly uses CNN for direct detection. The algorithm uses convolutional feature maps of different scales for detection. Large-scale feature maps can be used to detect small things, and small-scale feature maps can be used to detect large objects, so that objects of different scales can be detected.

3. YOLOv5 and Improvements

3.1. *YOLOv5*

Compared with YOLOv3, YOLOv5 has innovated in four parts of the network structure.

3.1.1. Input

(1)Mosaic data enhancement

This method randomly zooms, cuts, and arranges four pictures randomly, turns them into a new picture, and then puts the new picture into the network for learning.

(2)Adaptive anchor frame

In the early stage of training, a predefined frame will determine the location of the target at a possible location. As the training progresses, the real frame will gradually shift based on the preset frame for construction. Calculate the best anchor frame value in different training sets adaptively.

3.1.2 Backbone

The focus layer is very similar to adjacent downsampling. Suppose there is a 4x4 picture that is concatenated with separated pixel values into four 2x2 pictures. This structure can avoid information loss, the number of channels has become 4 times the original, and the size is half of the original.

3.1.3 Neck

Inserting the Neck layer between the input layer and the

Backbone layer improves the ability to extract fusion features. Can better detect targets of different sizes.

3.1.4 Output

Bounding box is used in YOLOv5. Bounding box is to fine-tune the predicted box to make it close to the ground truth box.

3.2. *Improvement*

3.2.1 BN and RBN

Batch norm is batch normalization[3]. It is to complete the normalization by adding parameters during the training process to solve the problem of normalized learning features. When applying Batch Norm, it should be satisfied that the mutual characteristics of different instances completely obey the same distribution. However, the distribution characteristics of the test set instances and the training data cannot be consistent. Therefore, the inconsistency in the training and testing process will weaken the actual effect of the BN layer.

In order to solve the above problems, this article quotes the method of RBN(Representative Batch Norm). RBN mainly takes two main steps: centering calibration and scaling calibration.

(1)Centering Calibration

$$X_{cm(n,x,h,w)} = X_{(n,c,h,w)} + W_m \cdot K_m \qquad (1)$$

In Eq.(1), X is set as the input feature W_m is set as a learnable weight vector, and K_m is expressed as an example feature. . This article uses GMP(Global Max Pooling). The learnable variable Wm is set to (N, C, 1, 1). '·'stands for dot product operation, which mainly converts two features into the same shape and then performs dot product operation.

(2) Scaling Calibration

In the next step of Batch Norm, specifically, before the stretching adjustment, do the image scaling and alignment related operations.

In the experiment, the BN layer was replaced with RBN to test its training effect in the YOLOv5 network.

3.2.2 Ghost model network

The core idea of GhostNet is to use cheap operations to replace ordinary convolution operations to generate these redundant feature maps. The Ghost module divides the ordinary convolution into two parts. First, it performs an ordinary 1x1 convolution, which is a small amount of convolution. For example, a 32-channel convolution is normally used. Here, a 16-channel convolution is used. The effect of this 1x1 convolution Similar to feature integration, the feature enrichment of the input feature layer is generated. Then we perform deep separable convolution. This deep separable convolution is layer-by-layer convolution, which is the cheap operations we mentioned above[4]. It uses the feature enrichment obtained in the previous step to generate a Ghost feature map. Using the Ghost Net network can reduce the amount of model parameters and increase the execution speed of the model while ensuring a good detection effect.The Ghost model is shown in the Fig 1 below.

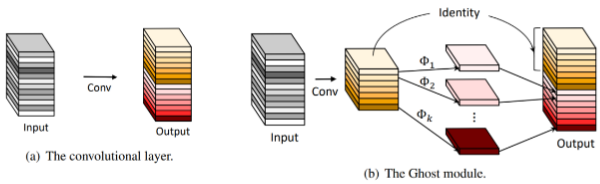


Fig.1. Ghost model schematic

4. Experiment and Result Analysis

4.1. *Experimental equipment*

The experimental equipment is shown in Table 1.

Table 1. Device-related configuration

Name	Version and Model
Ubunt version:	20.04.3 LTS
CUDA version:	11.2
Graphics:	GeForce GTX 1080 Ti
Frame:	Torch

4.2. *Data collection*

The preliminary data preparation mainly calibrates two categories, people and cars. It need to copy the .txt files with several preset classes to the relevant folder of the labelimg software. Through repeated screening and data supplementation, 30938 images of the data set and corresponding label files are provided, script codes are written, and the training set and the test set are allocated according to a 9:1 ratio. There are 27761 training sets and 3177 verification sets. Since the label file contains multiple categories(the experiment only recognizes people and cars, two categories), write script code to retain the specified label content (0-person, 1-car).

4.3. *Experimental setup*

Set lr=0.01, batch size to 16, and epoch to 300. The remaining parameters are default values for experimentation. Use the mAP (mean Average Precision) as the evaluation index.

Next, select about 500 pictures as the detect data and calculate the average inference speed. The experimental results are shown in Table 2.

Table 2 .Model size comparison

Model	mAP_0.5	mAP_0.5:0.95	Time(s)
YOLOv5s	0.9072	0.63213	0.02813
YOLOv5m	0.9278	0.67583	0.02879

It is known that the depth and width of the YOLOv5m network is greater than that of YOLOv5s. Combined with Table 2, it can be seen that increasing the network depth and width will increase the mAP value, but it will also increase its reasoning time. Therefore, this experiment uses the faster YOLOv5s for training and comparison.

4.4. *Experimental result*

The size of the convolution kernel restricts the amount of calculation of its model. Adjust the depth to separate the convolution kernel size, set g=1 and set the default kernel size to 5*5, respectively, adjust to 1*1 (point-by-point volume Product), 3*3 observe its model size and performance. The experimental results are shown in Table 3.

Table 3 .Network performance of different sizes of deep separable convolution kernels in Ghost mode

Convolution size	FLOPs	Time(H)	Parameters (M)	Model size	mAP@.5 mAP@.5:.95
1*1	14.5	14.7	6.283	27.8MB	0.900;0.626
3*3	14.5	16.6	6.287	27.8MB	0.901;0.629
5*5	14.6	19.4	6.295	30.8MB	0.905;0.632

As shown in Table 3, different convolution kernel sizes affect the amount of parameters and the size of the mAP. The larger the convolution kernel, the larger the mAP, and the amount of parameters will increase accordingly. The human-vehicle detection effect of the YOLOv5 network is shown in the Fig 2.

Fig.2. The result of recognition by model

5. Concluding

In terms of data, the universality can be improved by adding data types, for example, not only for learning and testing vehicles in a specific area, but also for learning and using vehicle data from all parts of the country or the world within the scope of the conditions.

In terms of algorithm selection, although the mainstream YOLOv5 network is relatively complete and has been widely used, some lightweight improvement ideas (pruning, distillation) can still provide better support for the use of this network on the mobile side, making the model size smaller. The complexity of calculation is reduced, the energy consumption of the mobile terminal is reduced, and the flexibility of deployment is improved[5]. A single picture containing a large number of people and cars can introduce the idea of clustering to group data points that are relatively close together. In addition, the transformer that shines in the field of NLP can now be applied to target detection. It provides a new idea for target detection to transform the detection problem into a set-based index problem, which is different from the traditional replacement of the backbone or the addition of a special FPN. If the attention mechanism can be properly combined with the YOLOv5 network, there may be better performance. This is also a research direction in the future.

References

1. Y. Hou, Y. Song, X. Hao et al., "Multispectral pedestrian detection based on deep convolutional neural networks," 2017 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), 2017, pp. 1-4, doi: 10.1109/ICSPCC.2017.824250
2. W. Li, "Infrared Image Pedestrian Detection via YOLO-V3," 2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), 2021, pp. 1052-1055, doi: 10.1109/IAEAC50856.2021.9390896.
3. W. Lan, J. Dang, Y. Wang et al., "Pedestrian Detection Based on YOLO Network Model," 2018 IEEE International Conference on Mechatronics and Automation (ICMA), 2018, pp. 1547-1551, doi: 10.1109/ICMA.2018.8484698.
4. Y. Huang, Y. Zhou, J. Lan, et al., "Ghost Feature Network for Super-Resolution," 2020 Cross Strait Radio Science & Wireless Technology Conference (CSRSWTC), 2020, pp. 1-3, doi: 10.1109/CSRSWTC50769.2020.9372549.
5. J. Tang, S. Liu, B. Zheng, et al., "Smoking Behavior Detection Based On Improved YOLOv5s Algorithm," 2021 9th International Symposium on Next Generation Electronics (ISNE), 2021, pp. 1-4, doi: 10.1109/ISNE48910.2021.9493637.

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.

Mr. Xiaoning Yan

He obtained a bachelor's degree from South China University of Technology. He obtained a master's degree in computer science from the University of Texas at Dallas from 2013 to 2015.

Mr. Shuangwu Zheng

He is a graduate student in control engineering of Tianjin University of Science and Technology. He is currently the deputy dean of Ansoft Algorithm Research Institute and an algorithm engineer of Ansoft 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).

Low light enhancement CNN Network based on attention mechanism

Xiwen Liang[1], Xiaoning Yan[2], Nenghua Xu[2], Xiaoyan Chen[1], Hao Feng[1]

[1]*Tianjin University of Science and Technology ,300222， China;*

[2]*Shenzhen Softsz Co.Ltd, Shenzhen, China;*

E-mail: 1540227453@qq.com

www.tust.edu.cn

Abstract

Low-light enhancement is a challenging task. With the image brightness increasing, the noises are amplified, and with the contrast and detail increasing, the false information is generated. In order to solve this problem, this paper proposes a novel end-to-end attention-guided method (A-MBLLEN) based on multi-branch convolutional neural network. The proposed network is composed with enhancement module (EM) and Convolutional Block Attention Module (CBAM). The attention module can make the CNN network structure gradually focus on the weak light area in the image, and the enhancement module can fully highlight the multi-branch feature graph under the guidance of attention. In this manner, image quality is improved from different aspects. Extensive experiments demonstrate that our method can produce high fidelity enhancement results for low-light images quantitatively and visually.

Keywords: Low-light image Enhancement, Deep learning, multi-branch Fusion, Convolutional Neural Network

1. Introduction

Due to unavoidable environments or technical limitations, many photographs are often taken under less than ideal lighting conditions. Poorly litght photos are not only bad for aesthetic quality, but also bad for messaging. The former affects the audience's experience, while the latter leads to misinformation being communicated. To solve these degradation and convert low-quality low light level images into normal light high-quality images, it is necessary to develop a good enhancement technology for low light image. In this paper, a deep neural network structure is proposed to improve the objective and subjective image quality. Extensive experiments demonstrate that our method can produce high fidelity enhancement results for low-light images quantitatively and visually. Our contributions are summarized as follows. 1) We introduce the convolution block attention module (CBAM) into the multi branch enhancement module, which can better highlight the low illumination area. 2) Our method is also effective in suppressing image noise and artifacts in low light region.

2. Related Work

Recently, deep learning has achieved great success in the field of low-level image processing. Powerful tools such as end-to-end networks and GANs[1] have been used in image enhancement. LLNet[2] ues the multilayer perception auto-encoder for low-light image enhancement and denoising. Retinex-Net[3] combines the Retinex theory with CNN to estimate the illumination map and enhance the low-light images by adjusting the illumination map. In order to integrate the advantages of CNN and GAN, Yang Etal[4] proposed a semi-supervised model of low-light image enhancement, which is enhanced in two stages. These methods are trained on paired data sets, and their enhancement performance largely depends on the data sets. Because the synthetic data can not fully describe the degradation in the real scene, and the real captured paired data contains a limited variety of scenes, the results of these methods are still imperfect, especially unable to deal with dense noise. Our model decomposes the complex image enhancement problem into sub-problem levels related to different features, which can be solved separately for multi-branch fusion, and the attention mechanism is embedded in the fusion process, which handling brightness/contrast enhancement and image denoising simultaneously.

3. Methodology

This paper proposes a novel end-to-end attention-guided method (A-MBLLEN) based on multi-branch convolutional neural network. The proposed network is composed with enhancement module (EM) and Convolutional Block Attention Module (CBAM). The attention module can make the CNN network structure gradually focus on the weak light area in the image, and the enhancement module can fully highlight the multi-branch feature graph under the guidance of attention. The overall network architecture and the data process flow is shown in Fig 1.

3.1. *Enhancement module (EM)*

EM contains multiple sub-nets, whose number equals to the number of the Branch numbers, and the output is a colorful image with the same size of the original low-light image. Each sub-net has a symmetric structure to

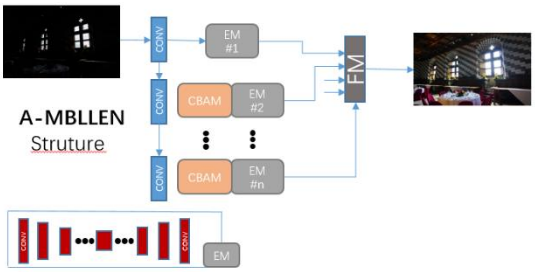


Fig 1. The workflow of the proposed approach for underexposed image restoration using A-MBLLEN. The proposed network with enhancement module (EM) and Convolutional Block Attention Module(CBAM) . The output image is produced via feature fusion.

apply convolutions and deconvolutions. The first convolutional layer uses 8 kernels of size 3×3, stride 1 and ReLU nonlinearity. Then, there are three convolutional layers and three deconvolutional layers, using kernel size 5×5, stride 1 and ReLU nonlinearity, with kernel numbers of 16, 16, 16, 16, 8 and 3 respectively. It should be noted that all subnets are trained at the same time, but are independent and do not share any learning parameters. The enhancement module can be considered to be composed of encoder and decoder. The encoder learns to extract the features of detection and weak light enhancement, and the decoder learns the accumulation from the feature space to the enhanced image. The encoder and decoder are connected skipply for detail reconstruction.

3.2. *CBAM*

Convolutional Block Attention Module (CBAM)[5] has two Attention submodules, CAM(Channel Attention Module) and SAM(Spatial Attention Module). CBAM details are shown below fig.2.
The CAM framework is shown in Fig3.
the input feature map F (H×W×C) is operated through global max pooling and global average pooling based on width and height, respectively, to obtain two 1×1×C feature maps which are separately feed into A two-layer multilayer perceptron (MLP). The number of neurons in

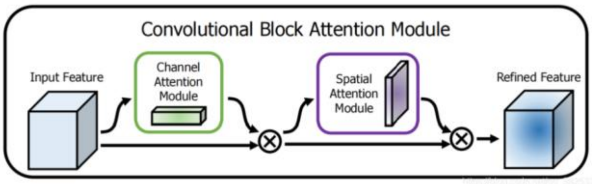


Fig 2. CAM(Channel Attention Module) and SAM(Spatial Attention Module). CAM is responsible for the attention weight on Channel, SAM is responsible for the attention weight on space (Height, Width).

the first layer is C/r (C is channel number, r is the

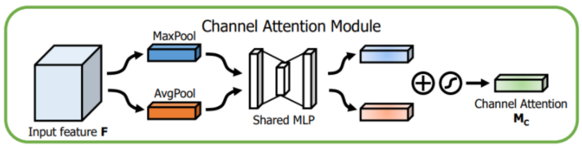


Fig 3. CAM struture

reduction rate), the activation function is ReLU, and the number of neurons in the second layer is C. The network is shared. Then, the MLP output features are subjected to an element-wise addition operation, and then the sigmoid activation operation is performed to generate the final channel attention feature, namely $M_C(F)$ in formula(1).

$$M_C(F) = \sigma\left(W_1\left(W_0(F^C{}_{\text{avg}})\right) + W_2(W_0(F^C{}_{max}))\right) \quad (1)$$

where $F^C{}_{\text{avg}}$ is the result of global average pooling, and $F^C{}_{max}$ is the result of global maximum pooling.

The SAM framework is shown in Fig4.

The feature map F output by the channel attention module as the input to this module. First, do a channel-based global maximum pooling and global average pooling to obtain two feature maps of size H×W×1, and then perform connection operations on these two feature

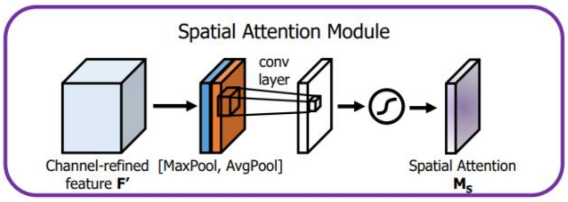


Fig 4 . SAM struture

maps based on channels. Then after a 7×7 convolution operation, the dimensionality is reduced to H×W×1, and then the spatial attention feature is generated through sigmoid, which is $M_s(F)$ in the formula (2).

$$M_s(F) = \sigma\left(f^{7\times7}[F^s{}_{\text{avg}}, F^s{}_{\text{max}}]\right) \quad (2)$$

where f represents the convolution operation and [,] represents the channel concatenation operation. Finally, $M_s(F)$ is multiplied by the input feature map of the module to obtain the final generated feature map.

3.3. *Loss Function*

The function of loss function is to make the enhanced image $E_{(x,y)}$ of the input image $I_{(x,y)}$ after the trainable CNN with parameters W enhancement as close as possible to the input reference image $R_{(x,y)}$. To improve the image quality both qualitatively and quantitatively, we propose a novel loss function. The MSE loss function to be minimized as:

$$\text{MSE}=\frac{1}{n}\sum_i^n \left\| E^i{}_{(x,y)} - R^i{}_{(x,y)} \right\|_2 \quad (3)$$

The structural loss function adopts DSSIM which is derived from structural similarity (SSIM)[6] and can be expressed as:

$$\text{DSSIM}=\frac{1}{n}\sum_i^n (1 - SSIM(R^i{}_{(x,y)} - E^i{}_{(x,y)})) \quad (4)$$

The Context loss can improve the visual quality. Because the VGG network [7] is shown to be well-structured and well-behaved, we choose the VGG network as the content extractor in our method. the context loss is defined as follows:

$$L_{VGG} = \frac{1}{C_{i,j}H_{i,j}W_{i,j}} \sum_{x=1}^{C_{i,j}} \sum_{y=1}^{H_{i,j}} \sum_{z=1}^{W_{i,j}} \left\| \varphi_{i,j}(E)_{x,y,z} - \varphi_{i,j}(R)_{x,y,z} \right\| \quad (5)$$

where E and G are the enhanced image and ground truth, and $w_{i,j}$, $H_{i,j}$ and $C_{i,j}$ describe the dimensions of the respective feature maps within the VGG network. Besides, $\varphi_{i,j}$ indicates the feature map obtained by j-th convolution layer in i-th block in the VGG-19 Network

Total Loss. The total loss can be expressed as:

$$L_{total} = MSE + DSSIM + L_{VGG} \quad (6)$$

4. Experimental Evaluations

Our implementation is done with Keras (Chollet et al. 2015)and Tensorflow (Abadi et al. 2016). The proposed network can be quickly converged after being trained for 20 epochs on a Titan-XGPU using the proposed dataset.

In the experiment, training is done using the ADAM optimizer [8] with a learning rate of $\alpha = 0.002$, $\beta_1 = 0.9$, $\beta_2 = 0.999$ and $\varepsilon = 10^{-8}$.

4.1. *Dataset and Metrcis*

Having a large dataset with diverse scenes and lighting conditions is significant for training a well-generalized model. We select 16925 images in the VOC dataset to synthesize the training set, 56 images for the validation set, and 144 images for the test set, The synthesis method is from reference [9]. Our synthetic low-light data sets have advantages over real low-light data sets, as they may be subject to device and environment constraints with a wide variety of light levels. By using the reference images provided in our dataset as enhancement groundtruth, we are able to quantitatively evaluate different methods in terms of PSNR and SSIM [10] indices. As no-referenced image quality assessment, We adopt Average Brightness (AB) [11] and Natural Image Quality Evaluator (NIQE) [12] which is a well-known no-reference image quality assessment for evaluating real image restoration without ground-truth, to provide quantitative comparisons. We compare our method with other methods on our synthetic dataset, Quantitative comparison is shown in Table 1.

Table 1 Comparison of different models.

Models	PSNR ↑	SSIM ↑	AB ↑	NIQE ↑
RetinexNet[3]	23.66	0.747	-4.14	30.57
ElightenGAN[14]	24.07	0.827	-3.13	27.76
MBLLEN[15]	25.95	0.885	0.008	29.47
OURS	**26.57**	**0.894**	**0.010**	27.62

5. Conclusions

In Table 1, our method achieves almost all the best results under all quality indicators. The only case where it ranks last is the brightness scaling result under the NIQE indicator. In general, our model has achieved good results from both qualitative and quantitative perspectives. We applied the model to actual noisy pictures, and found that A-MBLLEN can be adapted to real noisy low-light images to a certain extent, and can produce visually pleasing enhanced images. In the future, We will further explore better low-light adjustment methods to solve more complex realistic pictures.

References

1. A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta and A. A. Bharath, "Generative Adversarial Networks: An Overview," *IEEE Signal Processing Magazine, vol. 35, no. 1*, 2018, pp. 53-65.
2. Kin Gwn Lore, Adedotun Akintayo, Soumik Sarkar, LLNet: A deep autoencoder approach to natural low-light image enhancement, *Pattern Recognition*, 2017, 61:pp.650-662.
3. Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep Retinex Decomposition for Low-Light Enhancement. In BMVC, 2018. 2, 5, 6, 7, 8
4. W. Yang, S. Wang, Y. Fang, Y. Wang, and J. Liu, "From fidelity to perceptual quality: A semi-supervised approach for low-light image enhancement," *CVPR*, 2020, pp. 3063–3072.
5. Woo, Sanghyun, et al. "Cbam: Convolutional block attention module." Proceedings of the European conference on computer vision (ECCV). 2018.
6. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, , 2004, pp. 600–612.
7. Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.
8. Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014.
9. Lv F , Li Y , Lu F . Attention Guided Low-light Image Enhancement with a Large Scale Low-light Simulation Dataset[J]. 2019.
10. Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 2004,13(4) pp:600–612,.
11. ZhiYu Chen, Besma R Abidi, David L Page, and Mongi A Abidi. Gray-level grouping (glg): an automatic method for optimized image contrast enhancement-part i: the basic method. *IEEE transactions on image processing*, 2006, 15(8),pp :2290–2302,.
12. Anish Mittal, Rajiv Soundararajan, and Alan C Bovik. Making a" completely blind" image quality analyzer. *IEEE Signal Process*. 2013, Lett. 20(3) pp:209–212.
13. 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," 2019, arXiv arXiv:1906.06972.
14. F. Lv, F. Lu, J. Wu, and C. Lim, "MBLLEN: Low-light image/video enhancement using cnns," in BMVC, 2018.

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.

Mr. Xiaoning Yan

He graduated from Tianhua South University of Technology in 2012 with a bachelor of software engineering. In 2015, he received a master's degree in computer science from the University of Texas at Dallas. Since 2017, he has served as the technical director of Shenzhen Ansoft Vision Technology Co., Ltd.

Prof. 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.

Mr. Hao Feng

He received his bachelor's degree from the school of electronic information and automation of Tianjin University of science and technology in 2020. He is acquiring for his master's degree at Tianjin University of science and technology.

Fruit Recognition Based on YOLOX*

Keying Ren [1], Xiaoyan Chen[1]*, Zichen Wang[1], Xiaoning Yan[2], Dongyang Zhang[2]

[1] Tianjin University of Science and Technology, China;

[2] Shenzhen Softsz Co. Ltd., China

E-mail: renkeying@mail.tust.edu.cn
www.tust.edu.cn

Abstract

Pattern recognition is an urgent problem to be solved in the field of computer vision. In this paper, the network of fruit recognition based on YOLOX is studied. Due to the problem of slow training speed and low accuracy in the classical algorithms, the de-coupling detection head is optimized in YOLOX to overcome the above shortcomings. In terms of data enhancement, a new method combining Mosaic and MixUp is proposed. Through experimental verification, the method proposed in this paper has a great improvement over related algorithms such as YOLOv5, the accuracy is 98.6%, which is increased 5.2%.

Keywords: Pattern recognition, de-coupling detection head, YOLOX, Fruit recognition

1. Introduction

Pattern recognition is an urgent problem to be solved in the field of computer vision, aiming to detect objects of predefined categories. Accurate object detection impacts on various applications including image recognition and video surveillance. In recent years, with the development of convolutional neural network (CNN), object detection divided into one-stage and two-stage methods. A typical one-stage method is R-CNN algorithm. The main idea is to generate a series of bounding boxs through selective search or CNN network, These bounding boxes are classified and regressed[1]. The advantage of two-stage method is high accuracy. One-stage method, like You Only Look Once(YOLO) and Single Shot MultiBox Detector(SSD)[2,3].The main idea is to sample uniformly and concentratedly in different positions of the image. Different ratios can be used in sampling. CNN is used to extract features and then directly perform classification regression. The whole process only needs one step, so the advantage is fast speed. However, the disadvantage is that uniform and dense sampling is difficult to train. Due to the positive sample and the negative sample (background) unbalanced, which leads to the model accuracy decrease.

Both of them first tile a large number of preset anchors on the image, then predict the category and refine the coordinates of these anchors by one or several times, finally output these refined anchors as detection results. Currently, the leading algorithms of One-Stage are YOLO series. YOLO transforms the target detection task into regression problem, greatly speeding up the detection speed. At the same time, because the network used global information to predict each target window, the proportion of false positive was greatly reduced.

Recent academic attention has been geared toward anchor-free detectors due to the emergence of FPN[4] and Focal Loss[5]. Anchor-free detectors directly find objects without preset anchors in different ways. One is keypoint-based methods. Another is center-based method. These anchor-free detectors are able to eliminate those hyperparameters related to anchors and have achieved similar performance with anchor-based detectors, making them more potential in terms of generalization ability.

2. YOLOX

Megvii Technology proposed a new high-performance detector — YOLOX. YOLOX is an object detector which uses the feature of deep convolutional neural network. it makes some empirical improvements to the YOLO series, such as anchor-free, a decoupled head and the leading label assignment strategy SimOTA.

2.1. *Anchor-free*

Anchor-free detectors have developed rapidly in the past two years. These work have shown that the performance of anchor-free detectors can be on par with anchor-based detectors. Anchor-free mechanism significantly reduces the number of design parameters which need heuristic tuning and many tricks involved for good performance, making the detector, especially its training and decoding phase, considerably simpler.

Keypoint-based method. This type of anchor-free method first locates several pre-defined or self-learned keypoints, and then generates bounding boxes to detect objects. CornerNet [6] detects an object bounding box as a pair of keypoints (top-left corner and bottom-right corner). The R-CNN [7] locates objects via predicting grid points with the position sensitive merits of FCN and then determining the bounding box guided by the grid. ExtremeNet detects four extreme points (top-most, leftmost, bottom-most, right-most) and one center point to generate the object bounding box.

Center-based method. This kind of anchor-free method regards the center of object as foreground to define positives, and then predicts the distances from positives to the four sides of the object bounding box for detection. DenseBox uses a filled circle located in the center of the object to define positives and then predicts the four distances from positives to the bound of the object bounding box for location.

2.2. *Decoupled head*

In object detection, the conflict between classification and regression is a well-known problem. Thus the decoupled head for classification and localization is widely used in the detectors. However, as YOLO series' backbones and feature pyramids continuously evolving, their detection heads remain coupled. This structure is shown in Fig.1. The coupled detection head may harm the performance. Replacing YOLO's head with a decoupled one greatly improves theconverging speed. The decoupled head is essential to the end-to-end version of YOLO.

YOLOX replaced the YOLO head with a decoupled YOLO head, the convergence speed has been greatly improved. Concretely, it contains a 1 × 1 conv layer to reduce the channel dimension, followed by two parallel branches with two 3 × 3 conv layers respectively. This structure is shown in Fig.2.

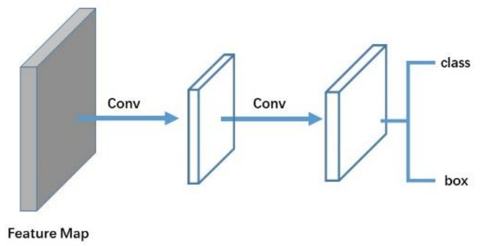


Fig.1. The original YOLO series coupling head structure, a convolution directly carried out classification and anchor regression.

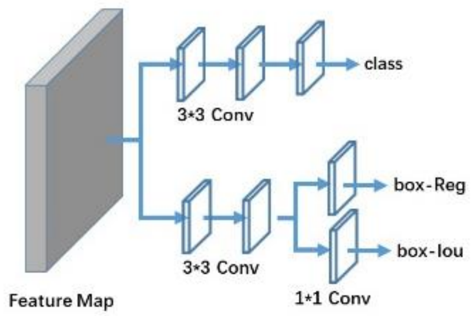


Fig.2. replace the YOLO head with a decoupled YOLO head, the convergence speed has been greatly improved.

Concretely, it contains a 1 × 1 conv layer to reduce the channel dimension, followed by two parallel branches with two 3 × 3 conv layers respectively.

3. Related work

3.1. *Images dataset*

Our data set consists of 30,000 photos of 60 fruits, labeled by VOC. Images were divided into training and test datasets with ratio of 9:1. The hardware configuration include one NVIDIA GeForce GTX1080ti SUPER graphic card, Intel(R) Xeon(R) CPU E5-2630 - Core Processor, and. The CSPDarknet framework and Python 3.6 were used.

3.2. *Image Preprocess*

Mosaic and MixUp were added to the enhancement strategy to improve YOLOX performance. Mosaic is an effective enhancement strategy proposed in Yolov3. and widely used in YOLOv4. MixUp is originally designed for image classification task but then modifie for object detection training. MixUp and Mosaic are uesd during training and hand off for the last 15 epochs.

Fig.3. First, the dataset was expanded using mirroring and rotation. Then, Mosaic method was used to merge images.

Fig.4. The MixUp method

4. Experiments and Results

Object detection performance is evaluated with Average Precision (AP) and mean Average Precision (mAP) between ground truth and predicted bounding box (IOU). To verify the feasibility of YOLOX in fruit recognition, two ex-periments are performed in this study. The experiment results are shown in Table 1.

Table 1. Comparison of mAP and IOU on YOLOv4 and SO-YOLO.

Method	**Average IOU%**	**mAP%**
YOLOv5-s	54.09	93.4
YOLOX-s	53.27	**98.6**

Experiment results reported in Table 1 is the average of multiple experiments. It can be seen that YOLOX has higher accuracy than YOLOv5-s. As shown in Table 1, mAP is increased from 93.4% to 98.%. Comparing to YOLOv5-s the proposed method has better performance and also well-balanced accuracy and processing time, experimental results are shown as Table 2.

Table 2. Comparison of network model parameters of each model.

Version	**parameters**	**Latency**
YOLOX-s	9.0 M	9.8 ms
YOLOv5-s	**7.3 M**	**8.7 ms**

Comparing with YOLOv5-s, the model parameters re-duced from 7.3M to 9.0M. Although the parameters and reasoning speed increased, it was acceptable for our project.

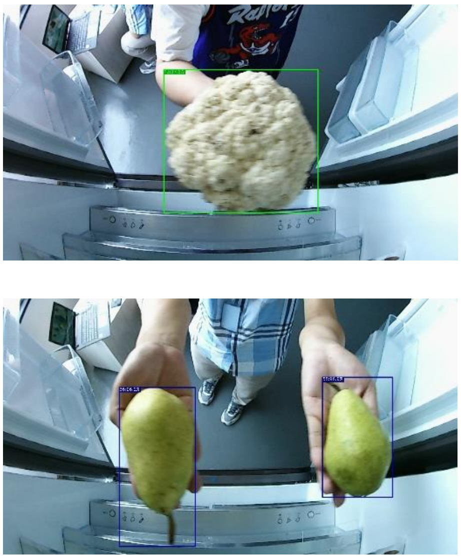


Fig.4. The result of recognition by YOLOX detector, it can be seen that the category confidence is very high, and box regression is very good.

5. Discussion

In order to improve the classification and detection of fruits, YOLOX method was applied and verified in this paper. We can see that YOLOX is good at fruit recognition tasks. YOLOX integrates the advantages of other target detection algorithms. Mosaic and MixUp were added to the enhancement strategy, which improved YOLOX-s performance. Replacing YOLO's head with a decoupled one greatly improves the converging speed. the performance of anchor-free detectors can be on par with anchor-based detectors. Anchor-free mechanism significantly reduces the number of design parameters.

References

1. Ren S , He K , Girshick R , et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017, 39(6):pp.1137-1149.
2. 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.
3. Liu W , Anguelov D , Erhan D , et al. SSD: Single Shot MultiBox Detector, *European Conference on Computer Vision. Springer, Cham*, 2016: pp.21-37.
4. Lin T Y , Dollar P , Girshick R , et al. Feature Pyramid Networks for Object Detection, *Proceedings of the IEEE conference on computer vision and pattern recognition. 2017* : pp.2117-2125.
5. Lin T Y , Goyal P , Girshick R , et al. Focal Loss for Dense Object Detection, *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017, PP(99):2999-3007.
6. Duan K , Bai S , Xie L , et al. CenterNet: Keypoint Triplets for Object Detection, *Proceedings of the IEEE/CVF International Conference on Computer Vision.* 2019: pp.6569-6578.
7. Lu X , Li B , Y Yue, et al. Grid R-CNN, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.* 2019: pp.7363-7372.

Authors Introduction

Mr. Keying Ren

Keying Ren 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

Xiaoyan Chen 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.

Mr. ZiChen Wang

ZiChen Wang was born in Tianjin, China in 1997. He received the B.E. degrees from Tianjin University of Science and Technology, Tianjin, China, in 2019. He is currently pursuing the M.S. degree in Information and Automatic College at Tianjin University of Science and Technology, Tianjin, China.

Mr. Xiaoning Yan

Xaoning Yan graduated from South China University of Technology with a bachelor's degree in software engineering. 2013.8-2015.5 studied for a master's degree in computer science at the University of Texas at Dallas.

Mr. Dongyang Zhang

Dongyang Zhang, a master's degree in control science and engineering from Tianjin University of Science and Technology. is currently an algorithm engineer of Ansoft Technology, engaged in deep learning and computer vision algorithm research, It mainly involves the research direction of target detection and pedestrian re-recognition.

An Improved Small Target Detection Method Based on YOLOv4*

Xia Miao[1], Xiaoyan Chen*[1], Keying Ren[1], Zichen Wang[1], Xiaoning Yan[2], Yue Sun[2]

[1]Tianjin University of Science and Technology, 300022, China;

[2] Shenzhen Softsz Co. Ltd., China
E-mail: miaoxia@mail.tust.edu.cn
www.tust.edu.cn

Abstract

In order to improve the efficiency and accuracy of small target detection in current traffic flow, this research proposes an improved YOLOv4 framework and applies it to small target detection task. A new small target-friendly 4-fold down-sampling residual is added between the second and third residual blocks of CSP Darknet-53 block to improve the detection accuracy of small target. The novel YOLOv4 model is optimized by above strategy. Compared with the original network, the modified framework can significantly improve the recall rate and average detection accuracy of small target.

Keywords: YOLOv4, small target detection, traffic flow

1. Introduction

With the continuous development of urbanization, traffic problems have gradually become an inevitable problem. Traffic congestion is a very important problem; it restricts people's travel. Traffic flow detection is an important part of solving the problem of traffic congestion. At present, target detection has been widely used in military and civilian fields. Among these target detections, small target detection, as an important target detection technology, has become a hotspot and focus of research. Long-range targets are usually small targets. Compared with large targets, small targets have the shortcomings of fewer pixels and unobvious features. Therefore, a series of low detection rate problems will occur in detection. Therefore, small target detection is still a topic worthy of research in target detection.

With the development of artificial intelligence in recent years, more and more researchers have begun to pay attention to vehicle detection algorithms based on deep learning. Among them, a variety of target detection and target tracking algorithms based on deep learning are proposed. Compared with traditional methods, methods based on deep learning can learn more target features. Currently widely used target detection algorithms based on deep learning can be divided into two categories: Two-step target detection algorithms, such as Fast R-CNN[1] (Regional Conventional Neural Network), R-CNN[2], etc. These algorithms divide target detection into two stages, that is, first use the regional candidate network (RPN) to extract candidate target information, and then use the detection network to complete the prediction and identification of the candidate target location and category; Single-step target detection algorithms, such as SSD (Single Shot Multi-box

Detector), YOLO[3] (You Only Look Once), YOLO 9000[4], YOLO v3[5], YOLOv4[6]. This algorithm does not need to use RPN, and directly generates target location and category information through the network. It is an end-to-end target detection algorithm. Therefore, the single-step target detection algorithm has the advantage of faster detection speed.

This research proposes an improved YOLOv4 framework and applies it to small target detection task. A new small target-friendly 4-fold down-sampling residual is added between the second and third residual blocks of CSP Darknet-53 block to improve the detection accuracy of small target. The novel YOLOv4 model is optimized by above strategy. Compared with the original network, the modified framework can significantly improve the recall rate and average detection accuracy of small target.

2. Improved YOLOv4 Network Structure

2.1. *YOLOv4 algorithm introduction*

YOLO is a single-step detection algorithm based on regression methods. This article uses YOLOv4 feature extraction network is CSP(Cross Stage Partial)Darknet53 + PANet-SSP structure. The backbone feature extraction network CSPDarknet53 strengthens the feature extraction networks SPPNet[7] and PANet, and finally uses the YOLO header to convert the extracted features into prediction results. CSPDarknet53 consists of a convolution block and five Resblock_body, which contains a convolution, normalization and Mish activation function.

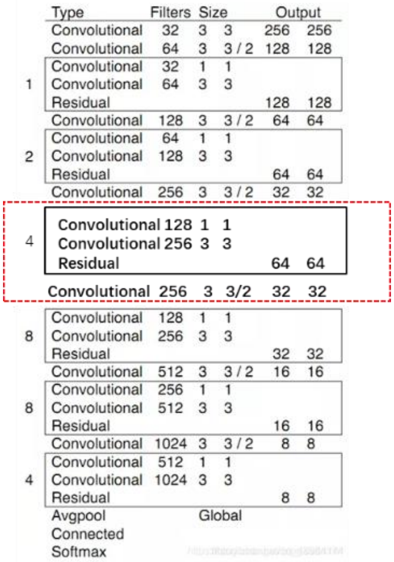
CSPDarknet reduces the amount of data transmission and calculation in the network, and reduces the amount of calculation and memory consumption required by the CNN network without losing accuracy and light weight. The input feature map is divided into two paths in the channel dimension, one is passed directly backwards, the other is passed backwards through multiple residual blocks, and finally merged with the CSP end. This cross-stage splitting and merging effectively reduces the possibility of gradient replication, increases the diversity of gradients, and reduces the amount of data transmission and calculation in the network. In YOLOv4, each CPSX contains 3+2*X convolutional layers, so the backbone of the entire backbone network contains 72 convolutional layers. In this paper, a new 4 times down-sampling residual block is added between the second and third residual blocks of CSPDarknet-53 block to improve the detection accuracy of small targets.

2.2. *Data set anchor box based on cluster analysis*

The anchor frame mechanism of the output layer of YOLOv4 is the same as that of YOLO v3. The main improvement is the loss function CIOU_ Loss during training (Complete Intersection Over Union_ Loss), which uses the anchor frame idea in Fast R-CNN, and the initial choice of anchor frame The frame size is a set of fixed frames, the selection of which directly affects the accuracy and efficiency of target detection. YOLOv4 uses K-means clustering algorithm to cluster the width and height of the internal target frame of coco data set, and takes the average overlap degree as the required measure for the target clustering analysis.

2.3. *Improved YOLOv4 network*

CSPDarknet-53 is the backbone network structure adopted by YOLOv4, in which the original intention of the CSP structure is to reduce the amount of calculation and enhance the performance of the gradient. The characteristics of the CSP structure are Strengthen the learning ability of CNN; Reduce computing bottlenecks; Reduce memory consumption. Darknet-53 down-sampled the input image 5 times, and predicted the target in the last 3 down-sampling. The last three down-sampling includes the feature map of three-scale target detection. Small feature maps provide in-depth semantic information, and large feature maps provide target location information. The sampled small feature map and the large feature map are fused, so that the model can detect both large targets and small targets. Darknet-53 draws on the idea of residual network-work (RN), which consists of five residual blocks. Each residual block is composed of multiple residual units (R), in which targets of different sizes correspond to different residual modules. In order to give full play to the advantages of Yolov4 network structure detection, a large amount of small target feature information can be obtained to improve detection efficiency and Accuracy, we can improve the detection of small targets on the basis of the original network structure. Add a residual module to the residual module 2 and residual module 3 of CSPDarknet53. The added residual module is a down-sampled target detection layer 4 times. Since the 4x down-sampling feature map contains a lot of small target position information, the 4x up-sampling feature map output by Yolov4 can be realized. The obtained feature map is mapped to the 4x down output of the second residual block of Darknet-53 Sampling feature mapping connection, building a 4x down-sampling feature fusion target detection layer, applied to small target detection. The structure diagram is shown in Fig. 1.



	Type	Filters	Size		Output	
	Convolutional	32	3	3	256	256
	Convolutional	64	3	3 / 2	128	128
1	Convolutional	32	1	1		
	Convolutional	64	3	3		
	Residual				128	128
	Convolutional	128	3	3 / 2	64	64
2	Convolutional	64	1	1		
	Convolutional	128	3	3		
	Residual				64	64
	Convolutional	256	3	3 / 2	32	32
4	Convolutional	128	1	1		
	Convolutional	256	3	3		
	Residual				64	64
	Convolutional	256	3	3/2	32	32
8	Convolutional	128	1	1		
	Convolutional	256	3	3		
	Residual				32	32
	Convolutional	512	3	3 / 2	16	16
8	Convolutional	256	1	1		
	Convolutional	512	3	3		
	Residual				16	16
	Convolutional	1024	3	3 / 2	8	8
4	Convolutional	512	1	1		
	Convolutional	1024	3	3		
	Residual				8	8
	Avgpool		Global			
	Connected					
	Softmax					

Fig. 1.   Figure 1. New backbone network structure diagram

3. Experimental Results and Analysis

In the current detection field, YOLOv4 is one of the representative algorithms. It has the characteristics of detecting large, medium and small targets, and has good detection performance in small target detection. Therefore, the improved YOLOv4 detection algorithm is compared with the original YOLOv4 algorithm. By using the VEDAI data set, the implementations of the two algorithms are compared. Divide the original satellite image into 1024*1024 images, including vehicles and background. Each image in the VEDAI dataset involves 5.5 vehicles, and the target only accounts or 0.7% of the total image pixels, which belongs to the small target detection dataset. There are nine types of goals, as shown in Table 1.

Table 1. Number of Targets in the Dataset

Class name	Boat	Camping	Car	Others	Pick up	Tractors	Truck	Airplane
Total number	160	400	1430	200	850	200	350	53

This paper uses the following method to verify the small target detection ability of the improved algorithm: the image data is set to a resolution of 512*512, the smallest three categories of data sets (cars, trucks, and trucks) are divided into one category, and 80% of them are randomly selected The sample of the data set is used as the training set, and the remaining data set is used as the test set. Then use the YOLOv4 algorithm and the improved YOLOv4 algorithm for training and testing respectively. After repeated iterations, all parameters tend to stabilize, and the final loss value decreases. The performance comparison between the original YOLOv4 algorithm and the improved YOLOv4 algorithm is shown in Table 2.

Table 2.   Performance comparison table of original algorithm and improved Yolov4.

	Algorithm Performance		
	Recall	Precision	mAP
YOLOv4	86.5%	89.3%	56.87%
NEW-YOLOv4	87.6%	92.1%	63.25%

The performance of various algorithms is calculated by the following formula.

$$Recall = \frac{TP}{TP + FN} \cdot \tag{1}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

$$mAP = \frac{\sum AP}{NC} \tag{3}$$

In the above formula, recall is the recall rate, which indicates the proportion of the real target identified in the algorithm return result to the total target; precision is the accuracy rate, which indicates the proportion of the real target in the algorithm return result; TP represents that the detection is a positive sample, which is actually a positive sample; FN represents that the detection is a negative sample, which is actually a positive sample; FP represents that the detection is a positive sample Samples are actually negative samples; mAP represents multi class average precision.

4. Conclusion

This paper proposes an improved YOLOv4 algorithm and applies it to small target detection. Mainly optimize and improve the network structure and data set. In the optimization of the network structure, in the second and third residual modules of the original network, a 4-fold down-sampling residual module is added to detect small targets, and then merged with the original output module. In the data set, all samples are clustered and a cluster center is found. Through the above optimization and improvement, the final experimental results show that the improved algorithm has a significant improvement in accuracy, recall and average accuracy compared to the original algorithm.

References

1. Girshick, R. . "Fast R-CNN." arXiv e-prints (2015).
2. Girshick, R. , et al. "rich feature hierarchies for accurate object detection and semantic segmentation tech report (v5)." (2017).
3. Redmon, J. , et al. "You Only Look Once: Unified, Real-Time Object Detection." IEEE (2016).
4. Redmon, J. , and A. Farhadi . "YOLO9000: Better, Faster, Stronger." IEEE Conference on Computer Vision & Pattern Recognition IEEE, 2017:6517-6525.
5. Redmon, J. , and A. Farhadi . "YOLOv3: An Incremental Improvement." arXiv e-prints (2018).
6. Bochkovskiy, A. , C. Y. Wang , and H. Liao . "YOLOv4: Optimal Speed and Accuracy of Object Detection." (2020).
7. He, K., et al. "Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition." IEEE Transactions on Pattern Analysis & Machine Intelligence 37.9(2014):1904-16.

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.

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.

Mr. ZiChen Wang

He was born in Tianjin, China in 1997. He received the B.E. degrees from Tianjin University of Science and Technology, Tianjin, China, in 2019. He is currently pursuing the M.S. degree in Information and Automatic College at Tianjin University of Science and Technology, Tianjin, China.

Mr. Yue Sun

He is a master's degree in computer application technology from Tianjin University of Science and Technology, is currently an algorithm engineer of Ansoft Technology, engaged in deep learning and computer vision algorithm research, mainly related to the research direction of target detection.

Mr. Xiaoning Yan

He graduated from South China University of Technology with a bachelor's degree in software engineering. 2013.8-2015.5 studied for a master's degree in computer science at the University of Texas at Dallas.

Yolov5-DP: A New Method for Detecting Pedestrian Aggregation

Kunzhi Yang[1]**, Xiaoyan Chen**[1*]**, Xiaoning Yan**[2]**, Dashuo Wu**[2]

[1]*Tianjin University of Science and Technology, China*[*]

[2]*Shenzhen Softsz Co. Ltd., China*

E-mail: 1663491981@qq.com

www.tust.edu.cn

Abstract

In this paper, a novel network Yolov5-DP (Yolov5-DBSCAN-P) is proposed. Deep separable convolution and ACON-C activation function are added into Yolov5 network to improve the detection accuracy of pedestrians. Firstly, DBSCAN-P is used as the clustering detector to detect pedestrians in the area. Secondly, the depth-separable convolution is used to replace the common convolution in Yolov5. Finally, the loss function Swish is improved to ACON to increase the model speed and reduce the model size. The Yolov5-DP network is tested on the public dataset MOT20Det. The experimental results show that good detection results and accurate aggregation detection results are obtained.

Keywords: Target detection, YOLOv5, clustering, regional population analysis

1. Introduction

Both at home and abroad, there are tragedies caused by crowds gathering every year. However, in daily life, schools, factories and other places often show the phenomenon of crowd gathering; During rallies, crowds will inevitably gather in squares and streets. In the face of this crowd load, every extreme behavior is very easy to cause accidents. In this context, in order to reduce the occurrence of this kind of tragedy, this paper aims to study the pedestrian detection task in (gathering) area.

In recent years, most research on pedestrian detection at home and abroad is based on object detection in computer vision. The development of target detection has roughly experienced two periods: traditional target detection (before 2014) and target detection based on deep learning (after 2014)[1]. The traditional target detector has sliding window (VJ), HOG, DPM, etc. With the manual characteristic performance becoming saturated, it is replaced by deep learning. Nowadays, target detectors based on deep learning have become common, which can be divided into Two categories: two-stage Detection based on candidate regions and one-stage Detection based on regression classification. Two-stage target detector is represented by RCNN[2] series, while one-stage target detector is represented by SSD[3], RetinaNet[4], YOLO[5] series.

The main research direction of crowd is crowd counting, and the key technology is to integrate the generated thermal map to get crowd counting. There are also crowd-based abnormal (clustering) behavior research, which has many methods, including traditional methods based on machine learning, big data analysis, and methods combined with deep learning and machine learning, with few data.

2. Yolov5-DP Network

YOLOv5 model is a recently proposed neural network model, which performs very well in target detection tasks. Excellent results have been achieved in both COCO and VOC datasets. YOLOv5 network structure is shown in Fig.1.

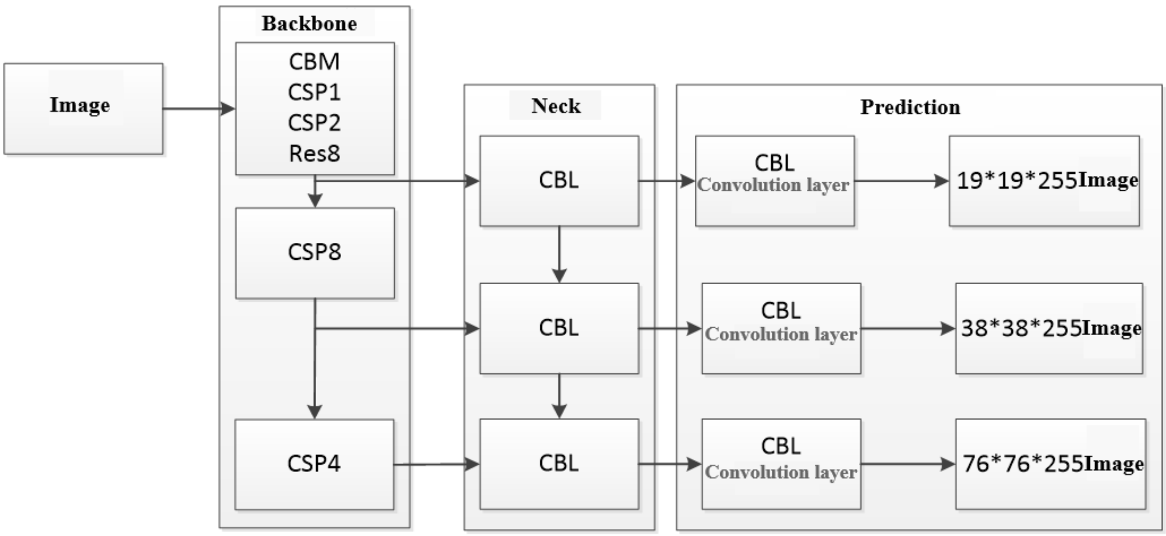


Fig. 1. YOLOv5 network structure

1.1. *Depthwise separable convolution*

Depthwise Separable Convolution has two steps: Depthwise Convolution and Pointwise Convolution. First, each feature graph at each layer corresponds to a convolution with a convolution kernel size of 3×3 for convolution operation, and then carries out the second convolution operation with a convolution kernel size of 1×1 for each channel to obtain the output feature graph with the set size.

Let the size of depthwise convolution kernel be $D_K * D_K * 1$, the number of input channels C_{in}, the number of output channels C_{out}, W and H are the width and height of the input feature graph respectively. The formula for calculating the proportion of parameter number (param) and computation amount (FLOPs) between depth-separable convolution and ordinary convolution is as follows:

$$\eta_{param} = \frac{D_K * D_K * C_{in} + C_{in} * C_{out}}{D_K * D_K * C_{in} * C_{out}} = \frac{1}{N} + \frac{1}{D_K^2} \quad (1)$$

$$\eta_{FLOPs} = \frac{D_K * D_K * C_{in} * W * H + C_{in} * C_{out} * W * H}{D_K * D_K * C_{in} * C_{out} * W * H} = \frac{1}{N} + \frac{1}{D_K^2} \quad (2)$$

In this paper, deep separable convolution is used to replace ordinary convolution in YOLOv5 networks. The advantage of using depth separable convolution is that the number of parameters and computation are smaller than ordinary convolution. It can be seen from the formula that the number of parameters and calculation quantity are $\frac{1}{N} + \frac{1}{D_K^2}$ of ordinary convolution, and the value is between $1/8$ and $1/9$ when the convolution kernel size is 3×3.

1.2. *ACON-c activation function*

Ma N et al proposed a new activation function ACON (Activate or not)[6]. ACON has several variants, and

ACON-C is one of the most widespread forms. ACON-c simply uses hyperparameters to scale the feature graph, and the formula is:

$$f_{ACON-C}(x) = (p_1 - p_2)x \cdot \sigma[\beta(p_1 - p_2)x] + p_2 x \qquad (3)$$

Unlike the first derivative of Swish's activation function[7], which has fixed upper and lower bounds, the first derivative of ACON-C has learnable upper and lower bounds.

2. Algorithm Testing and Experimental Analysis

2.1. *Selection and processing of experimental data*

In the pedestrian detection algorithm based on YOLOv5 network, the selection of data sets has a great influence on the detection effect, and there are certain requirements on the total amount of data images or the number of single objects used for training model. MOT20Det open source dataset was selected for the dataset. The visualization of the MOT20Det dataset for different scenarios is shown in Fig.2.

The amount of data used for train and VAL (verification) in the training were 8000 groups and 931 groups respectively, and the grouping was randomly selected. There were 4479 sets of data used for tests.

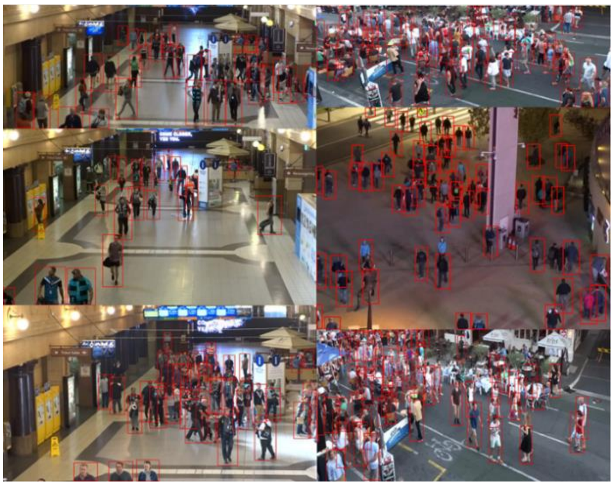


Fig. 2. Visualizations of different scenarios from the MOT20Det dataset

2.2. *Pedestrian detection experiment and test*

3.2.1.*Model training*

As the task only needs to detect pedestrians, the smallest model YOLOv5s is selected to ensure high accuracy and high performance pedestrian detector with high detection efficiency.

The convolutional substitution depth detachable convolution experiment in YOLOv5 model, this paper designed three detachable convolution structures for comparative experiments, as shown in Fig.3, Fig.4 and Fig.5.

Fig. 3. Structure of DP 1

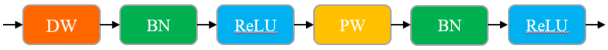


Fig. 4. Structure of DP 2

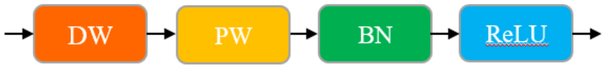


Fig. 5. Structure of DP 3

3.2.2.*Experimental results and analysis*

In Table 1, Conv represents the original YOLOv5 model, DP_3_S is to replace the ReLU activation function in DP_3 with SiLU, DP_3_ac is the replacement of ACON-C activation function, and DP_3_F is the depth-separable convolution replacement except the Focus layer. The activation function of YOLOv5 model from input layer to SPP layer is replaced by SiLU with ACON-C, that is, ACON-C.

YOLOv5 has 1×1 convolution and 3×3 convolution. According to the experiment, compared with only replacing 3×3 convolution, the model accuracy of replacing all convolution remained unchanged, and the number of parameters and training time increased a little. Therefore, the subsequent deeply separable convolution replacement experiment only targeted at 3×3 convolution.

It can be seen from Table 1 that the single ACON-C activation function replacement has a small increase in accuracy at the cost of a larger model. DP_3_S structure has the best performance in the deeply-detachable convolution substitution experiment. The combination of ACON-C function and deeply-detachable convolution can accelerate the network training speed on the premise of keeping the accuracy unchanged. Compared with the original YOLOv5, the size of the DP_3_ac model is reduced by half while the accuracy is basically the same.

Another conclusion is that the training speed is not proportional to the size of the model.

Table 1. Comparison of YOLOv5 model adjustment results

Method	mAP@0.5	mAP@0.5:0.95	param	GFLOPs	epochs	duration/h
Conv	0.9673	0.7904	7063542	16.4	100	5.932
ACON-C	0.968	0.7922	7408118	18.1	100	6.296
DP_1	0.9559	0.7359	3194746	6.7	100	8.205
DP_2	0.9582	0.745	3194746	6.7	100	8.17
DP_3	0.96	0.7558	3190114	6.6	100	12.387
DP_2_S	0.9641	0.7721	3194746	6.7	100	7.772
DP_3_S	0.9649	0.7785	3190114	6.6	100	7.191
DP_3_F_S	0.9645	0.7767	3193078	7.2	100	7.801
DP_3_ac	0.9648	0.7793	3534690	8.3	100	6.418

2.3. *Pedestrian Cluster Detection Test and Testing*

3.3.1. *Algorithm implementation details*

In the experiment of clustering detection algorithm, there are data processing and parameter selection. Each part is introduced as follows:

(1) Firstly, the normalized coordinate data of pedestrian detection frame is extracted, and the pixel coordinates of the center point are generated into array data.

(2) K-means algorithm ADAPTS to elbow method to get the optimal clustering result; The MinPts parameter of DBSCAN algorithm is set to 3, and eps value is obtained by the adaptive algorithm. The MinPts parameter in the OPTICS algorithm is 5 (default); HDBSCAN algorithm has no manual parameter input. The MinPts parameter of DBSCAN -p is set to 3, and the proportionality coefficient n is set to 1.25.

3.3.2. *Performance evaluation index*

Since there is no performance evaluation index for clustering detection, this paper carries out clustering detection based on clustering algorithm, and the performance measurement of clustering is selected as one of the performance evaluation indexes for clustering detection. Real-time detection is also particularly important, so the time complexity of the algorithm is also used as the comparison item of the aggregation detection experiment.

The performance measures of clustering are divided into internal indicators and external indicators, both of which are effective performance indicators of clustering. The external indicators need to be evaluated according to the reference model. The MOT20Det data set used in the experiment has no aggregation data annotation. Therefore, this paper only compares the internal indicators to judge the rationality of clustering within the cluster, and then makes a comparative analysis on the clustering effect, that is, the visual images of clustering detection. Silhouette Coefficient, a common internal indicator, is adopted in this paper [8]

3.3.3. *Experimental results and analysis*

The 139 groups of data used in this experiment were randomly extracted from all 13410 groups of data in MOT20Det dataset at a ratio of 1%, including 8 groups of lens framing images of training set and detection set. The effect map in the experiment only visualized the aggregation target with rectangular boxes of different colors. Experimental data of aggregation detection algorithm comparison are shown in Table 2, where the effective data amount is the reference sub-data of contour coefficient s(i).

Table 2. Performance comparison of clustering detection algorithms

methods	Silhouette Coefficient/s(i)	Valid data quantity/sheet	Time complexity/s
K-Means	0.4301	139	0.13519
DBSCAN	0.3221	129	0.00135
OPTICS	0.5814	124	0.04315
HDBSCAN	0.4502	103	0.00257
DBSCAN-P	0.6269	63	0.05815

Different from k-means algorithm, the clustering result must be greater than or equal to 2 clusters, DBSCAN series algorithms will produce clustering results with the number of clusters equal to or less than 1. However, both contour coefficient method and DI index and DBI index, both common internal indicators, need to be used in the case of more than one cluster, so the experiment only carries out statistics on valid data. The contour coefficient s(i) value in Table 2 is the average sum of s(i) of each group of effective data, and the time complexity is the average sum of operation time of each algorithm on 139 groups of data.

3. Algorithm testing and experimental analysis

In this paper, the method of target detection is used to analyze and study pedestrian detection in the region. For pedestrian detection, this paper adjusted the YOLOv5 network and conducted a comparative experiment on the original model and the model with depth detachable volume and ACon-C activation function added to each structure combination, specifically for the comparison of training and test results. Finally, the optimized pedestrian detector is obtained. For clustering detection, four classical clustering algorithms including K-means, DBSCAN, OPTICS and HDBSCAN and the dbSCAN-P clustering algorithm adjusted based on DBSCAN in this paper were compared and tested to obtain clustering detector. The optimized YOLOv5 was used as the pedestrian target detector, and dbSCAN-P algorithm was used for clustering detection. Finally, the two were verified on the MOT20Det test dataset to complete the study of pedestrian detection in the area based on YOLOv5.

References

1. Zou, Zhengxia, et al. "Object Detection in 20 Years: A Survey." ArXiv Preprint ArXiv:1905.05055, 2019.
2. Girshick, Ross, et al. "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation." CVPR '14 Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 580–587.
3. Liu, Wei, et al. "SSD: Single Shot MultiBox Detector." 14th European Conference on Computer Vision, ECCV 2016, 2016, pp. 21–37.
4. Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollar, P. (2020). Focal Loss for Dense Object Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 42(2), 318–327.
5. Redmon, Joseph, et al. "You Only Look Once: Unified, Real-Time Object Detection." 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779–788.
6. Ma N , X Zhang, Sun J . Activate or Not: Learning Customized Activation. 2020.
7. Ramachandran, Prajit, et al. "Searching for Activation Functions." ArXiv: Neural and Evolutionary Computing, 2017.
8. Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics, 20, 53–65.

Authors Introduction

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. Dr. Xiaoyan Chen

She is currently a professor and doctoral supervisor at the School of Electronic Information and Automation, Tianjin University of Science and Technology. Her research interests are pattern recognition and deep learning.

Mr. Xiaoning Yan

He received his Master's degree in engineering from the University of Texas at Dallas in 2015. He is currently the Technical director of Shenzhen Ansoft Huishi Technology Co., LTD.

Mr. Dashuo Wu

He received his master's degree in engineering from Tianjin University of Science and Technology in 2019. He is currently an algorithm engineer of Ansoft Technology, engaged in deep learning and computer vision algorithm research, mainly involving target detection research direction.

Flame Recognition based on Yolov5 Algorithm

Kunzhi Yang[1], **Xiaoyan Chen**[1*], **Xiaoning Yan**[2], **Dashuo Wu**[2]

[1]*Tianjin University of Science and Technology, China*[*]

[2]*Shenzhen Softsz Co. Ltd., China*

E-mail: 1663491981@qq.com

www.tust.edu.cn

Abstract

To address the problem of low accuracy and speed of flame detection, this paper proposes an improved YOLOv5 for flame detection. The new network is based on YOLOv5 by changing the loss function to DIoU(Distance Intersection over Union). Through introducing a large number of training data sets, it is hoped to improve the object detection accuracy. The experimental results show that the proposed YOLOv5 algorithm is effective with higher accuracy and faster detection for different flames.

Keywords: Flame Recognition, YOLOv5, deep learning, computer vision

1. Introduction

Since the flame was used by human beings, it brings people not only convenience, and convenience followed by all kinds of disasters caused by flames. If the fire did not develop into a serious fire detection and extinguishment in time, can effectively prevent the occurrence of fire, reduce the loss of fire to people. Therefore, the early detection of fire is very important.

In recent years, computer vision has become a very hot research and learning direction, and many fields have begun to use computer vision to help improve work efficiency. When using computer vision for flame detection, many defects of traditional instruments can be reduced. It has the advantages of fast response, wide detection range and less environmental pollution. Compared with traditional fire detection methods, it has wider prospects. However, these new methods are still immature and need to improve their detection accuracy. Different algorithms need to find different characteristics of the flame, such as color, motion, shape, frequency, texture and other aspects to judge whether the result is correct, and then combine with the existing deep learning algorithm, finally become a set of completed flame recognition technology.

Object detection algorithms based on deep learning are mostly divided into Two ideas. One is two-stage method, in which a candidate frame is formed and objects in the candidate frame are detected by algorithm. The other is the one-stage method, which integrates the whole detection process together and directly presents the detection results, such as YOLO (You Only Look Once) series[1-3] and SSD(Single Shot MultiBox Detector) series[4]. YOLO series detection algorithms have been developed since 2016, and the detection accuracy has been significantly improved.

2. YOLOv5 Network

YOLOv5 model is a recently proposed neural network model, which performs very well in target detection tasks. Excellent results have been achieved in both COCO and VOC datasets. YOLOv5 network structure is shown in Fig. 1.

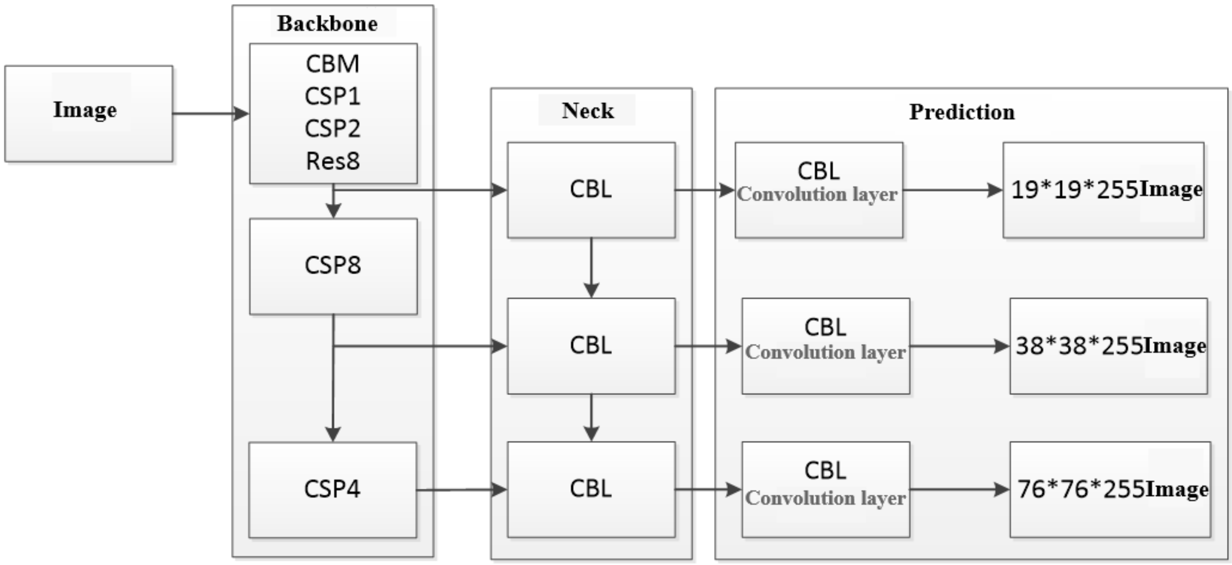


Fig. 1. YOLOv5 network structure

2.1. *Improved regression loss function*

DIoU loss is to reintroduce a penalty term $R(B^p, B^{gt})$ on the basis of IoU loss, so the calculation formula of DIoU loss is:

$$L_{DIoU} = 1 - IoU + R(B^p, B^{gt}) \quad (1)$$

Where $R(B^p, B^{gt})$ is the penalty term:

$$R(B^p, B^{gt}) = \frac{\rho^2(b^p, b^{gt})}{c^2} \quad (2)$$

$\rho(\cdot)$ is the Euclidean distance, b^p and b^{gt} represent the center point of the prediction frame B^p and the real frame B^{gt} respectively, and c is the diagonal length of the minimum area covering the prediction frame B^p and the real frame B^{gt}.

As shown in Fig.2, the green box is the target box, the black box is the prediction box, and the gray box is the minimum area covering the two boxes. d is the distance between the center point of the prediction box and the real box, and c is the diagonal length of the gray box, namely the value of variable c in Formula 2.

Like GIoU loss, the loss value of DIoU loss is still not related to size. Furthermore, DIoU losses can be further optimized in cases where the prediction box does not overlap with the real box. Different from GIoU loss, DIoU can minimize the distance between the center point of the prediction box and the real box, which greatly speeds up the convergence of the loss.

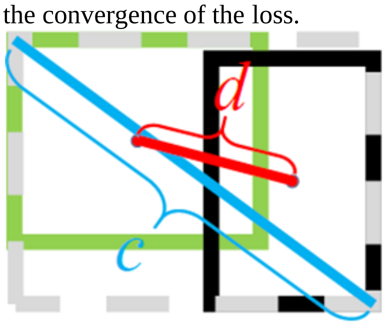


Fig. 2. Schematic diagram of DIoU

3. Construction of Flame Detection Model

In this experiment, a large number of data are used to train the improved YOLOv5.

3.1. *Experiment preparation*

LabeImg was selected as the image annotation tool. The annotation principles are as follows: label bright flame edge; The outline of smoke that can block the object is marked, and the other cases are not marked. A total of 14,649 training sets and 3,767 verification sets were used in the training of this model, which were labeled as fire. The flow chart of YOLOV5 model training is shown in Fig.2.

3.2. *Tests and results*

After the successful training of the model, a model can be obtained to detect random flame images, which can output the position information and confidence of the image to be detected in the form of a label box on the image. The specific results of the target detection task are shown in the figure. The rectangular box represents the position of the flame, and the text above represents the classification and confidence of the target. Fig. 3 is the output of several test images:

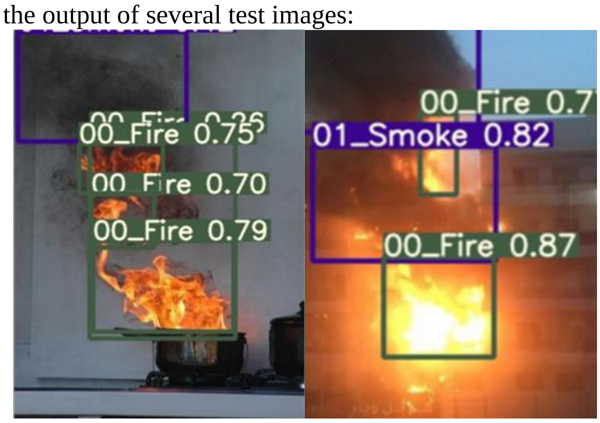


Fig. 3. Flame detection result

Take the picture on the left as an example, the green rectangular box represents the detected target position, 00_Fire in the picture represents the target in the rectangular box is flame, and 0.70, 0.79 and other numbers represent confidence.

4. Conclusion

In this paper, computer vision is used to realize the identification of flames in various Spaces, and flames in video clips can be detected in real time. Specifically, after the video image data captured by each camera is collected, the image is labeled first, and then sent into the neural network for model training. Finally, the flame image can be labeled on the random input to output flame information.

References

1. Redmon J, Divvala S, Girshick R, et al. You only look once : Unified , real-time object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016 : 779-788.
2. Redmon J, Farhadi A. YOLO9000 : better, faster, stronger. arXiv preprint, 2017.
3. Redmon J, Farhadi A. YOLOv3 : An incremental improvement. arXiv preprint arXiv : 1804. 02767, 2018.
4. Liu W, Anguelov D, Erhan D, et al. Ssd : Single shot multibox detector. European Conference on Computer Vision. Springer, Cham, 2016 : 21-37.

Authors Introduction

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. Dr. Xiaoyan Chen

She is currently a professor and doctoral supervisor at the School of Electronic Information and Automation, Tianjin University of Science and Technology. Her research interests are pattern recognition and deep learning.

Mr. Xiaoning Yan

He received his Master's degree in engineering from the University of Texas at Dallas in 2015. He is currently the Technical director of Shenzhen Ansoft Huishi Technology Co., LTD.

Mr. Dongyang Zhang

He received his master's degree in engineering from Tianjin University of Science and Technology in 2020. He is currently an algorithm engineer of Ansoft Technology, engaged in deep learning and computer vision algorithm research, mainly involving target detection and pedestrian re-recognition research direction.

Research on Face Detection Algorithm Based on Improved YOLOv5

Zhen Mao[1], Xiaoyan Chen[1*], Xiaoning Yan[2], Yuwei Zhao[2]

[1] Tianjin University of Science and Technology, China;

[2]Shenzhen Softsz Co. Ltd., China
E-mail: 1037331281@qq.com
www.tust.edu.cn

Abstract

Face detection technology is one of the research hotspots in the field of deep learning in recent years. Aiming at the problems of slow detection speed and low accuracy of various target detection algorithms, this paper proposes an improved target detection algorithm based on YOLOv5. By introducing lightweight network, changing the depth and width of YOLOv5 network structure and reducing the number of model parameters, the network reasoning speed can be greatly accelerated. At the same time, the method uses Acon adaptive activation function to further improve the accuracy of face detection. Experimental results show that the improved algorithm has faster detection speed and higher detection accuracy than the traditional algorithms.

Keywords: Deep learning, Face Detection, Yolov5, Acon ctivation function

1. Introduction

With the rapid development of deep learning, face detection has become a hot research direction in the field of artificial intelligence. In recent years, with the continuous progress of target detection algorithm, the research on face detection is also effective, and even can exceed people's resolution level in some scenes. Face detection technology is widely used in video surveillance, identity recognition, and human-computer interaction and so on, which makes people's life more convenient and safe. However, with the continuous improvement of CPU processing speed and the increasing ability of graphics card image processing, people have a further demand for the detection accuracy and speed of face detection. How to achieve higher accuracy and speed detection based on the existing algorithms has become the research direction of this paper[1].

The research period of face detection algorithm can be divided into three stages. The first stage is the semi-mechanical recognition stage, the second stage is man-machine interactive recognition stage, and the third stage is face detection method based on deep learning. The function and purpose of face detection is to accurately find all faces in a digital picture and mark their positions and information. The traditional detection methods have low detection accuracy due to the lack of classification ability and feature extraction ability of the classifier. With the development and progress of convolutional neural network (CNN)[2], the detection method of deep learning based on neural network gradually replaces the traditional face detection method, and has made great improvement in accuracy, recall and location. In 2015, R. Joseph et al. proposed a new target detection algorithm Yolo. In this paper, the latest yolov5 algorithm is used to study face detection[3]. On the basis of the original model, the depth and width of the network structure are changed to reduce the parameters of the model. At the same time, ACON adaptive activation function is used to improve the speed and accuracy of detection.

2. Yolov5s Model Structure

In this paper, Yolov5 detection algorithm is used for face detection. When Yolov5 network is used for detection, there are deficiencies in detection speed and accuracy. A method to reduce the depth and width of the network is proposed, which reduces the amount of parameters and operations, improves the network speed, and adopts the newly proposed Acon activation function, It makes up for the decline of network performance after reducing network depth and width.

Yolov5 model contains four target detection versions. Among them, Yolov5s has the smallest network scale and the fastest speed, but at the same time, the value of AP, that is, the accuracy, is also the lowest. It is suitable for application scenarios that focus on large target detection and pursue high detection rate[4]. The other three models are based on the Yolov5s model to improve the ability of feature extraction and feature fusion by continuously deepening the depth of the network and broadening the width of the feature map. Yolov5s model consists of input, backbone, neck and prediction. The input adopts adaptive image scaling and mosaic data amplification methods, including random scaling, clipping and layout splicing, adaptive anchor box calculation to adapt to different data sets and adaptive image scaling to improve the reasoning speed. Backbone mainly includes slice structure (focus), convolution module (conv), bottleneck layer (C3) and spatial pyramid pooling (SPP)[5]. The focus structure is unique to Yolov5 model, which is different from Yolov3 and Yolov4. It is mainly used to slice images. The neck network part is composed of csp2_ Composed of X structure and FPN +PAN structure, the feature fusion ability of the network is strengthened and richer feature information is retained. Output layer: bounding box loss using GIOU_ Loss replaced IOU_ Loss, as a loss function, predicts targets of different sizes on characteristic maps of different sizes[6]. The structure diagram of the model is shown in Figure 1.

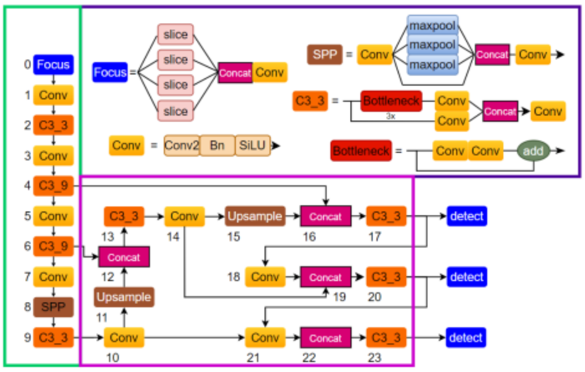


Fig. 1. The structure diagram of YOLOv5

3. Improved Yolov5 Model

Due to the task of face detection is relatively simple, single category, and has a sufficient number of data sets, so the use of native Yolov5s network still exists the phenomenon of network performance surplus. Therefore, the Yolov5 network was modified in this paper, and the depth and width of the network were reduced. As a result, the number of convolution kernels for sampling operations during the call of residual modules in the network and feature extraction were reduced, which greatly reduced the parameters and computation amount in the network operation. At the same time, Acon activation function is used to realize automatic adaptive automatic learning to improve network performance.

3.1. *Adjustment of network parameters*

In the Yolov5s network structure, GW = width_ Multiple is 0.5, while in the standard focus, C2 = 64, which is substituted into the code to calculate the result 32, that is, the number of convolution layers in the first convolution down sampling operation is 32, while in the second convolution down sampling operation, C2 = 128, GW = 0.5, the calculated result is 64, that is, the number of convolution cores in the second convolution down sampling operation is 64. The parameters for training with yolov5s model are: model summary: 283 layers, 7063542 parameters, 7063542 gradients, 16.4 gflops. We can see that there are 7063542 parameters, 283 convolution layers and 16.4g flops. After final adjustment, the depth of the network is 0.2 and the width is 0.3. The parameters of the final model are: model summary: 265 layers, 2621006 parameters, 2621006 gradients, 6.3 gflops.

3.2. *Acon activation function*

In the original Yolov5 model, the activation function in the backbone network is Silu type activation function, which is replaced by Acon activation function in this paper. Acon function learns the parameter switching between nonlinear (active) and linear (inactive) by introducing a switching factor. The activation function can significantly improve the depth model in the tasks of image classification, target detection and semantic segmentation.

4. Analysis of Experimental Results

4.1. *Experimental data set*

The data set in this paper is mainly the face and mask data set. The data set source is the public data set. After sorting, the data set photo is 11396, including 11042 face frames with masks and 17444 label frames with single face. It is divided into 9409 training sets (9483 face frames with masks and 15808 single face frames), 983 test sets (626 face frames with masks and 1229 single face frames), and 1004 verification sets (933 face frames with masks and 407 single face frames).

4.2. *Evaluating indicator*

In this paper, the accuracy P (precision), model size, recall, AP (average precision) and mAP (mean average precision) are used to evaluate the performance of the model for small target detection. Accuracy rate is used to measure the accuracy of model detection, that is, precision rate. Recall rate is used to evaluate the comprehensiveness of model detection, that is, recall rate. The formula used is as follows.

$$R = \frac{TP}{TP+FP} \times 100\% \tag{1}$$

$$P = \frac{TP}{TP+FN} \times 100\% \tag{2}$$

$$AP = \int_0^1 P(\mathrm{r})dr \tag{3}$$

$$mAP = \frac{\sum_{i=1}^{k} AP_i}{k} \tag{4}$$

In the formula, TP represents the number of positive samples predicted, FP represents the number of negative samples predicted, FN represents the number of negative samples predicted,AP represents the area covered under P(r) curve with recall as X-axis and Precision as Y-axis, which measures the recognition accuracy of a certain category. mAP is the average value of AP of each category. It measures the average quality of AP in all categories.

4.3. *Analysis of test results*

Firstly, based on yolov5s network, change the width and depth of the network, and compare the depth and width coefficient of the original network. The adjusted network is 0.2 in depth and 0.3 in width. The comparison of the three performance indicators is shown in Figure 2. Table 1 shows the training results of adjusting network parameters.

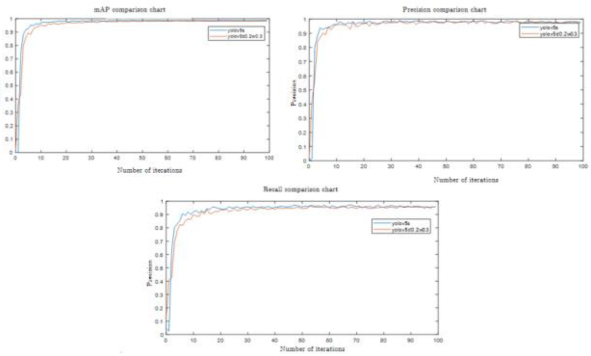


Fig. 2. Yolov5s and Yolov5_ d0. 2_ w0. 3 performance comparison chart

Table 1. Compare the performance of the model.

model	AP	Precision	Recall
Yolov5s	0.9880	0.9800	0.9574
Yolov5_d0.2_w0.3	0.9825	0.9720	0.9566

It can be seen that the AP value of the original network is 0.9880, while the AP value of the changed network is 0.9825, a decrease of 0.0055; The precision value of the original network is 0.9800, while the precision value of the changed network is 0.972, a decrease of 0.008; The recall value of the original network is 0.9574, while the recall value of the changed network is 0.9566, a decrease of 0.0008. While keeping the accuracy basically unchanged, our parameters are less. After changing the network, the parameters of the network are reduced from more than 7 million to more than 2 million, the amount of computation is changed from 16.4 g flops to 6.3 g flops, and the model reasoning speed is faster.

After the activation function is changed in this paper, the performance of the model is further improved. The comparison of the changed indicators is shown in Figure 3. Table 2 shows the training results using Acon activation function.

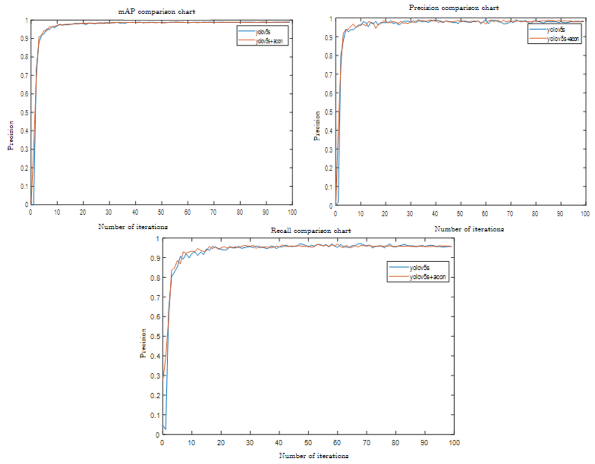


Fig. 3. Yolov5s and Yolov5s+Acon performance comparison chart

Table 2. Compare the performance of the model.

model	AP	Precision	Recall
Yolov5s	0.9880	0.980	0.957
Yolov5s+Acon	0.9883	0.985	0.956

According to the AP, precision and recall comparison data of yolov5s and yolov5s + Acon activation functions in the above chart, the map value of yolov5 + Acon is 0.9883, while the map value of the original yolov5s is 0.988. Because its ap value is very close to 1, it is very difficult to improve all, but Acon activation function finally increases the AP value by 0.0003. In the comparison diagram of precision, the precision of the original network is 0.98, while the precision after using Acon activation function is 0.985, an increase of 0.005. In the recall comparison chart, the original network is 0.957, yolov5s + Acon is 0.956, a decrease of 0.001. In general, adding Acon activation function can improve AP and precision to a certain extent. The performance of the whole network is improved.

The comparison between adjusting model parameters and changing Acon function and yolov5 model indicators is shown in Figure 4 below. Table 3 shows the training results of the comprehensive improved network

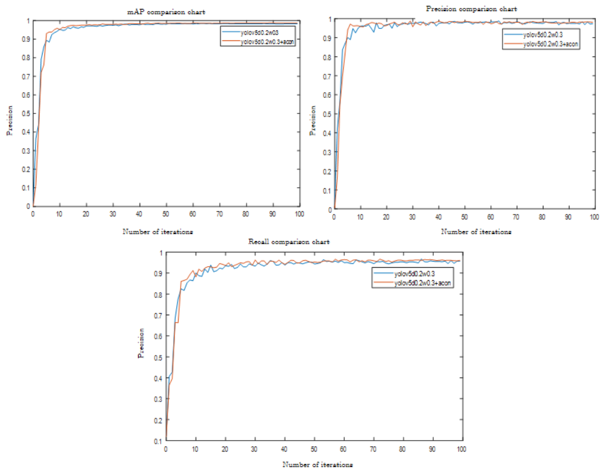


Fig. 4. Yolov5s and Yolov5_ d0. 2_ w0. 3+Acon performance comparison chart

Table 3. Compare the performance of the model.

model	AP	Precision	Recall
Yolov5s	0.988	0.9801	0.957
Yolov5_d0.2_w_0.3 +Acon	0.987	0.9802	0.96

It can be seen from Figure 4 and table 3 that the AP, precision and recall values of yolov5s network are 0.988, 0.9801 and 0.957 respectively, while the AP, precision and recall values of the network finally changed in this paper are 0.987, 0.9802 and 0.96 respectively. Although the final network differs from the original yolov5s network by 0.001 in AP value, the precision and recall are improved by 0.0001 and 0.003 respectively. In general, the performance of the two networks is comparable. However, because the depth and width of the network in this paper are smaller than that of yolov5s network, the amount of parameters and computation are small, especially in terms of computation, the comprehensive performance of the network in this paper is better, Reasoning is faster.

5. Discussion and Conclusion

This paper changes the depth and width of yolov5 network structure, significantly reduces the amount of parameters, and greatly speeds up the reasoning speed of the network. At the same time, the Acon activation function is used to further improve the performance of the network. The yolov5 network with depth of 0.2 and width of 0.3 is used with Acon activation function, so that the network can not only maintain rapidity but also not lose accuracy.

6. References

1. Zou Z, Shi Z, Guo Y, et al. Object detection in 20 years: A survey. arXiv preprint arXiv:1905.05055, 2019.
2. R. Alex Krizhevsky, Ilya Sutskever, Geoffrey Hinton. ImageNet Classification with Deep Convolutional Neural Networks. Advances in neural information processing systems, 2012, 25: 1097-1105.
3. Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi. You Only Look Once: Unified, Real-Time Object Detection. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA, 2016.
4. J GAO M H,YANG C.Traffic target detection method based on improved convolutional neural network .Journal of Jilin University(Engineering Edition),2021,45(5):1-9(in Chinese).
5. HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition.IEEE transactions on pattern analysis and machine intelligence, IEEE, 2015, 37(9): 1904–1916.
6. 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: 8759–8768.

Authors Introduction

Mr. Zhen Mao

He obtained his bachelor's degree from Nanjing Institute of Technology in China in 2020. He is currently studying for a master's degree at Tianjin University of Science and Technology. The research direction is the edge computing field of deep learning, engaged in the research of model deployment.

Prof. Dr. Xiaoyan Chen

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.2012y.

Mr. Xiaoning Yan

He obtained a bachelor's degree in software engineering from South China University of technology in 2012 and a master's degree in computer science from the University of Texas Dallas in 2015 . Now he is the technical director of Shenzhen Ansoft Huishi Technology Co., Ltd

Dr. Hidekazu Suzuki

He Graduated from Tianjin University of science and technology, majoring in control science and engineering. Now he is an algorithm engineer of Ansoft technology. He is engaged in edge computing and artificial intelligence algorithm application research, mainly involving the research direction of embedded edge computing.

Visibility Analysis Based on Deep Learning

Zhen Mao[1]，Xiaoyan Chen[1*]，Xiaoning Yan[2], Shuangwu Zheng[2]
[1]Tianjin University of Science and Technology, China;

[2]Shenzhen Softsz Co. Ltd., China
E-mail:1037331281@qq.com
www.tust.edu.cn

Abstract

In recent years, visibility analysis through deep learning processing and analysis of video images for different places has become a hot research topic that attracts people’s attention. A new deep learning model (A-VGGNet) is proposed to evaluate the visibility of real scenes. The model is constructed on the basis of the VGG classification model, and the classification accuracy of the deep learning model is improved by adding an attention mechanism. The experimental results show that the training success rate is 97.62%, the verification set test accuracy rate is 75.05%, and the test set classification accuracy rate is 85.05%. The proposed model has a good effect on the accuracy evaluation and classification of visibility.

Keywords: Deep learning, Regression analysis, VGG network, Attention mechanism

1. Introduction

Visibility is an indicator of atmospheric transparency in meteorology. In daily life, it indicates the maximum distance that people with normal vision can see and recognize objects. The unit is usually meter. Atmospheric visibility is closely related to people's daily life. Visibility analysis has important research significance in transportation, navigation, aviation and national defense military activities. Traffic accidents caused by low visibility are 2.5 times higher than other disasters, which will have a great impact on road transportation, aircraft flight and maritime navigation[1].The main factors affecting visibility are fog and haze. Generally speaking, the diameter of fine powder floating particles in haze is less than 0.01 μ m, so they can directly enter the human body through breathing, causing harm to people's health. Therefore, the accurate prediction and prediction of atmospheric visibility plays an important role and significance in urban air pollution control, ensuring public transport safety and maintaining people's life and property safety[2].

At present, the methods of visibility measurement mainly include human eye estimation, equipment detection and video image analysis. Among them, the traditional human eye estimation has great subjectivity, with large precision error and no specific quantitative standard. Although the instrument detection has high detection precision, it cannot be used on a large scale due to its high cost, difficult operation and troublesome carrying [3].With the development of computer vision technology and artificial intelligence science, video image visibility detection based on deep learning has gradually attracted people's attention. Using video to analyze visibility has the advantages of high speed, low cost and easy access to observed image data[4]. Through the analysis and processing of video images, this paper establishes the relationship between video images and real scenes, and then analyzes and classifies the meteorological elements affecting visibility according to the image characteristics.

AAAAAAAAA

2. Research Methods and Technical Routes

In this paper, regression analysis is used to study the relationship between visibility and temperature, humidity and wind speed in fog. The variables are fitted and analyzed, and the relationship between visibility and the three is described according to the visibility index. Aiming at the problem of visibility accuracy evaluation, this paper designs a new deep learning model A-VGG net to improve the model on the basis of VGG classification model. The model improves the classification accuracy of deep learning model by adding attention mechanism, and realizes the accurate classification of visibility.

2.1. *Regression analysis*

This paper studies the relationship between visibility and temperature, humidity, wind speed and other different weather conditions. In order to find the relationship between visibility and different meteorological parameters, the meteorological elements affecting visibility are analyzed by fitting MOR (Meteorological Optical Range) with temperature, humidity and wind speed, and with the help of regression idea.

Regression analysis is a predictive modeling technology, which studies the relationship between dependent variable (target) and independent variable (predictor). This technology is usually used for prediction analysis, time series model and discovering the causal relationship between variables. This study uses polynomial regression fitting. The regression function of polynomial regression is a regression in which the regression variable is polynomial, and the regression model is a kind of linear regression model. At this time, the regression function is linear with respect to the regression coefficient.

2.2. *A-VGG Net*

The model adopted for this visibility estimation is named A-VGG net improves the VGG (Visual Geometry Group) classification model and improves the classification accuracy of the deep learning model by adding an attention mechanism. The network structure is shown in Figure 1.

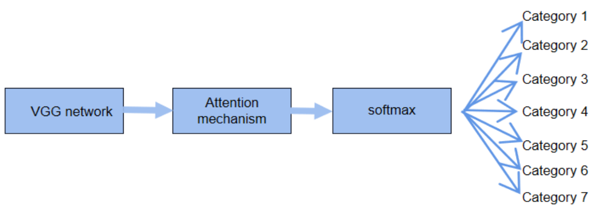


Fig. 1.A-VGG network structure

The VGG network[5] was proposed by the famous research group VGG of Oxford University in 2014. VGG is composed of 5-layer convolution layer, 3-layer full connection layer and softmax output layer. Max pooling is used to separate layers. The activation units of all hidden layers adopt Relu function. VGG network replaces large-scale convolution kernel VGG for feature extraction by stacking multiple 3x3 convolution kernels. In this way, multi-layer neural network can extract more features under a certain receptive field, and the training speed is relatively fast due to the small amount of parameters[6]. VGG16 classical convolutional neural network model is used in this model, which can be well suitable for classification and location tasks.

Attention mechanism is a data processing method in machine learning. It is widely used in natural language learning, image processing, speech recognition and other fields. Its core goal is to select information more critical to the current task goal from a large number of information. The network structure is shown in Figure 2 below.

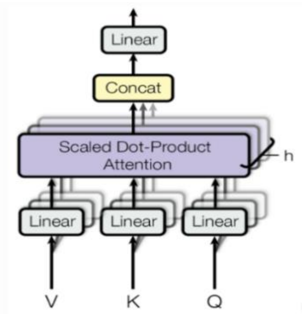


Fig. 2.Structure of attention mechanism

Neural attention mechanism can make neural network have the ability to focus on its input subset, and can select specific input. Attention can be applied to any type of input, regardless of its shape. In the case of limited computing power, attention mechanism is a resource allocation scheme that is the main means to solve the problem of information overload. Computing resources are allocated to more important tasks. Attention is generally divided into two types: one is top-down conscious attention, which is called focused attention. Focused attention refers to the attention that has a predetermined purpose, depends on tasks, and actively and consciously focuses on an object. The other is the bottom-up unconscious attention, which is called significance based attention[7].

3. Experiments and Results

3.1. *Data processing*

Firstly, the detected video is frame extracted. The extracted videos are selected and classified. The training sets are divided into 7 categories, each training set contains 600 pictures. The verification set and test set are also divided into 7 categories, each category contains 100 pictures.

In this paper, the training data set is expanded and enhanced through image translation, rotation, mirror image and Gaussian noise, 2400 images of each type are obtained, and the training set is expanded to 15800.The enhancement of the model training data can not only improves the generalization ability and robustness of the model, but also basically overcomes the imbalance of positive and negative samples. And label the selected and classified photos. Resize the picture to 224 × 224 size pictures and then sent to the designed network for training.

3.2. *Analysis of factors affecting visibility*

In this paper, the meteorological optical range is used to fit the relationship between visibility and temperature, humidity and wind speed. After sorting out the data, 1439 input samples are counted. A single sample is composed of temperature, humidity and wind speed. The temperature, humidity and wind speed are fitted respectively. The fitting results are as follows. Figure 3 shows the fitting results of MOR and temperature, figure 4 shows the fitting results of MOR and humidity, and figure 5 shows the fitting results of MOR and wind speed.

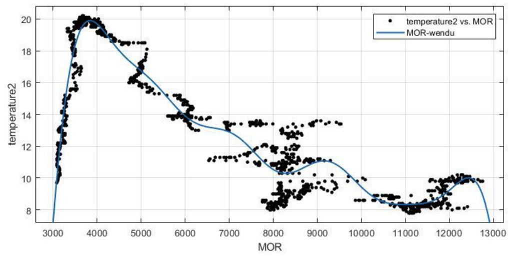


Fig. 3. Fitting curve of MOR and temperature

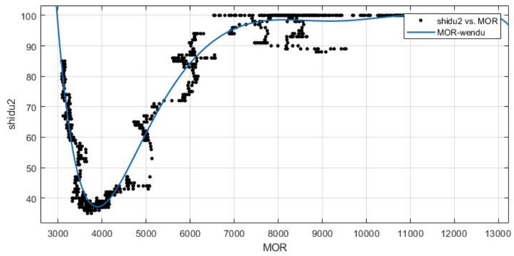


Fig. 4. Fitting curve of MOR and humidity

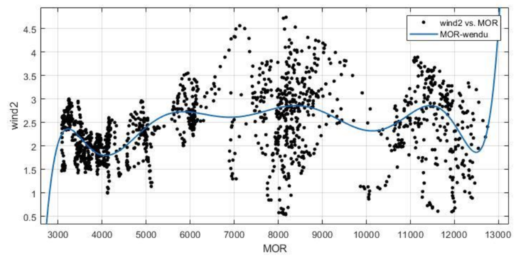


Fig. 5. Fitting curve of MOR and wind speed

From the fitting results, the fitting effect is good for the two parameters of temperature and humidity, but for the case of discrete data points such as wind speed, it cannot achieve good results, and other methods may be needed to analyze it separately. Multiple linear regression fitting combined with several main influencing factors shows that visibility is related to meteorological factors, directly proportional to temperature and wind speed and inversely proportional to humidity.

3.3. *Visibility classification*

In the process of evaluating the accuracy, the visibility is 7 classified and processed, and the improved A-VGG net is used to classify the visibility in video data. During the experiment, the training environment adopts Ubuntu

18.04 operating system, the CPU is Intel xeone5-2360, and the graphics card is two NVIDIA geforce gtx1080ti 11gb. During the experiments, the batch size is set to 16, each epoch takes 16 samples for training in the training set until all samples are trained once, that is, one epoch. 100 epochs were trained in the training process. The test environment of this paper adopts windows 10 operating system with i5 CPU- 9400F@2.9HZ The graphics card is NVIDIA geforce gtx1660 6GB. Through the training of the model, the accuracy of the training set and the accuracy of the verification set are obtained.

In this paper, the depth learning model is constructed. Based on the intercepted visibility data set, the experiment is carried out with ten-fold cross validation. Train_acc (TrainingAccuracy) and val_acc(Validation Accuracy) curves of the deep learning model were obtained according to the experiments. At the same time, train_loss and val_loss curves were obtained. The resulting curve is shown in Figure 6.

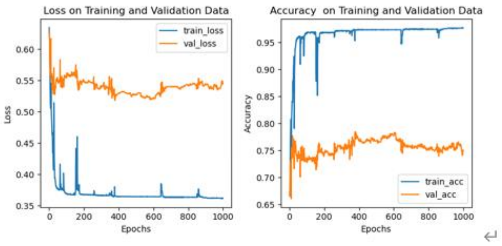


Fig. 6. Model training results

The results show that the loss curve of this algorithm decreases rapidly with the number of iterative steps, approaches 0 rapidly, and fluctuates slightly near the loss value of 0. Meanwhile, as the number of iteration steps increases, at 50 epoch, the train_ acc and val_ acc curve rises in a straight line and quickly approaches 1.0, that is, 100%. After a small shock of the validation data, the curve is basically stable at 100 epoch, the training loss value of the model is 0.3620, and the training success rate is train_ acc: 97.62%, where the validation set loss value val_ loss: 0.5445, verify the integrated power val_ acc: 75.05%, achieving an ideal training success rate, verifying the effectiveness and feasibility of the proposed network model, and maintaining a good fitting state in the process of learning the data set. The model uses one sixth of the data set for classification test, and the test score of algorithm test is 85.05%.

4. Discussion and conclusion

In this paper, the curve between visibility and temperature, humidity and wind speed is fitted through regression analysis. From the fitting results affecting the visibility mor value, it can be concluded that visibility is related to meteorological factors, directly proportional to temperature and wind speed, and inversely proportional to humidity. At the same time, the improved A-VGG network based on the VGG combines the attention mechanism to produce a good classification effect on the accuracy evaluation and classification of visibility. It can be seen that the deep learning network model designed in this paper has a good classification effect on the visibility data. By adding the deep convolution layer and adding the main force mechanism module, the distinguishability and characteristics of the data set are enhanced, which verifies the effectiveness and feasibility of the model proposed in this paper.

5. References

1. S. K. Nayar, S. G. Narasimhan, Vision in bad weather, ICCV’99.
2. R. T. Tan, Visibility in bad weather from a single image, CVPR, 2008.
3. Jinjia G, Zhaobin S, Zhishen L. Comparison of Visibility Measurements over Horizontal Path by Micro-pulsed Lidar and Visibility Meter. Journal of Ocean University of China, 2007, 6(3):315-318.
4. Shaoyun Hu, Zewei Jiang, Qingan Meng, et al. The development of a new-type atmosphere visibility meter. International Society for Optics and Photonics, 2014.
5. Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deepconvolutional neural networks. Advances in neural information processing systems, 2012, 25: 1097-1105.
6. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. ar Xiv preprint ar Xiv:1409.1556, 2014.
7. Zagoruyko S, Komodakis N. Paying more attention to attention: Improving the nperformance of convolutional neural networks via attention transfer. ar Xiv preprint ar Xiv:1612.03928, 2016.

Authors Introduction

Mr. Zhen Mao

He obtained his bachelor's degree from Nanjing Institute of Technology in China in 2020. He is currently studying for a master's degree at Tianjin University of Science and Technology.

Prof. Dr. Xiaoyan Chen

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.2012y.

Mr. Xiaoning Yan

He obtained a bachelor's degree in software engineering from South China University of technology in 2012 and a master's degree in computer science from the University of Texas Dallas in 2015 . Now he is the technical director of Shenzhen Ansoft Huishi Technology Co., Ltd

Mr. Xiaoning Yan

He is a graduate student majoring in control engineering of Tianjin University of science and technology. He is now the vice president of Ansoft Algorithm Research Institute and Ansoft algorithm engineer. He is mainly engaged in the application research of deep learning, computer vision and artificial intelligence algorithms.

*Price Prediction of Diamonds

Xiran Wen, Qiqi Xu, Zirui Su, Jiayi Fang
The Chinese University of Hong Kong, Shenzhen

Email: 119010333@link.cuhk.edu.cn

Abstract

The experiment aimed at price prediction based on diamond dataset which contains 53940 rows of information. The model is constructed based on linear regression model with the lowest estimated test error among all methods including tree and nonlinear models. The experimental results show that the mean square error for the training dataset and validation dataset are 592182.6 and 603833.2 respectively, and the R^2 reached 98%. The test MSE is 631947. The proposed model can well predict diamond prices.

Keywords: Price prediction, Linear regression, Mean square error, Cross Validation

1. Introduction

The classic ***Diamonds*** dataset contains the prices and 10 attributes of 53940 diamonds. It's a typical dataset of linear regression to analyze and visualize data.[1]

The price of a commodity is determined by its value, so as diamonds. Precisely estimating the price of diamonds could help merchants and customers make transactions properly. As the price of diamonds is significantly affected by the data of diamonds' attributes, we train a model on price of diamonds. The internal relationship of the dataset will be converted into the model with predictors. When we input the predictors, the model will automatically output a predicted price. And the goal of the model is to precisely predict the price according to their weights, color, quality of cut, measurement of clarity, width, length, depth, total depth percentage and width of top of diamond relative to widest point.

2. Research Material and Methods

2.1. *Data description*

The dataset contains 53940 rows and 10 columns. The price column is treated as the output y. Other variables and their meanings are as follows: (1)price: price in US dollars; (2)carat: weight of the diamond; (3)cut: quality of the cut (Fair, Good, Very Good, Premium, Ideal); (4)color: diamond color, from J (worst) to D (best); (5)clarity: a measurement of how clear the diamond is (I1 (worst), SI2, SI1, VS2, VS1, VVS2, VVS1, IF (best)); (6)x: length in mm; (7)y: width in mm; (8)z: depth in mm; (9)depth: total depth percentage = z / mean(x, y) = 2 * z / (x + y); (10)table: width of top of diamond relative to widest point.

2.2. *Methods*

In this paper, several regression analysis methods are used to determine the main factor that influence the price and study the relationship between diamonds' characteristics and prices. Linear regression model is used when the relationship between input and output is linear.[2]

*

Several non-linear regression models including polynomial, regression spline, natural spline regression and tree, are used to fit the data. Each model is estimated in two ways.[2] First, for the validation set approach, calculate training MSE based on training data set and estimated test MSE based on validation set. Second, for 10-fold cross validation approach, the estimated test MSE is calculated based on the training dataset only.[2]

Since, the training error of the model is approximately equivalent to its estimated test error, according to the Variance-Bias Trade off, which implies that the model has not yet over-fitted the data. At the same time, training error of the model is small enough when most of data is well explained which is the model with high R^2. However, smallest training error will lead to poor performance on test data that is over-fit. Therefore, the model with the high R^2 and the training MSE similar to estimated test MSE is chosen to be the final combination of predictors.

The final model is fitted based on training and validation dataset and test MSE is obtained by test dataset.

3. Experiments and Results

3.1. *Data processing*

The dataset is checked whether there is any invalid data. Rows contain NA are removed. The data is randomly split into three parts, 50% for training data, 20% for validation data and 30% for test data.3 category variables, cut, color and clarity, are transformed into 18 dummy variables using R.

3.2. *Data visualization*

Weight is generally considered as the critical factor of diamond price. So let x axis be the carat and y axis be price and randomly draw 500 data to plot.

Diamonds in different colors, cuts and clarity indicate linear relationship between price and carat as well. (Fig.1) Therefore, our first assumption of the model is a linear model with all 10 predictors.

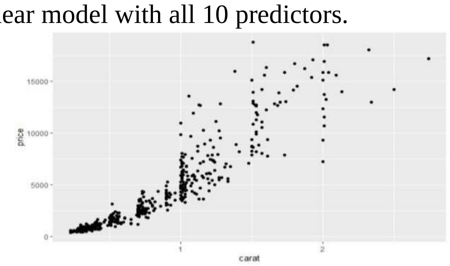


Fig.1 The relationship between the price and carat

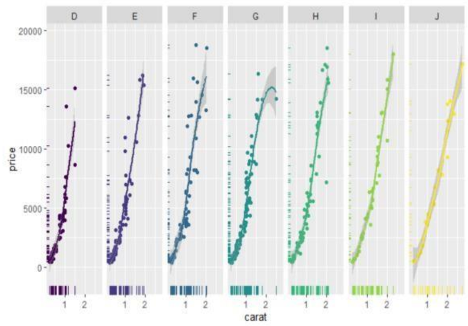


Fig.2 The price vs carat relationship based on diamond colors

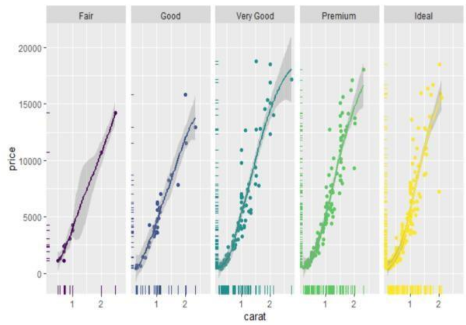


Fig.3 The price vs carat relationship based on cut quality

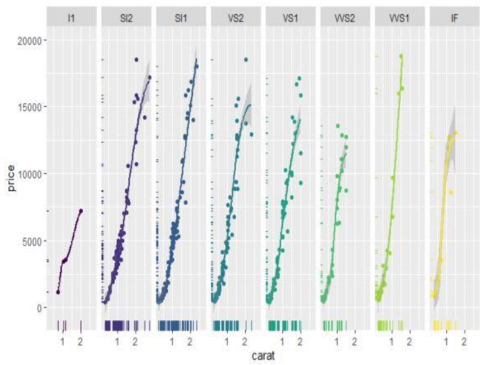


Fig.4 The price vs carat relationship based on clarity

All figures above show a nearly linear relationship. Then assumption is made that it is a linear model.

3.3. *Linear models*

From the data visualization, linear model is a good choice. The initial model is composed by all 10 predictors with 1301920 training MSE and estimated test MSE 1219288. Nearly 92% training data can be explained by this model which is not good enough.

After checking GVIF^(1/(2*Df)), it is found that the dimension of diamonds and their weight have value over 2 which indicates collinearity between them.[3] Through correlation map, it can also be seen that predictor selection is needed. Different methods are used to select significant variables.

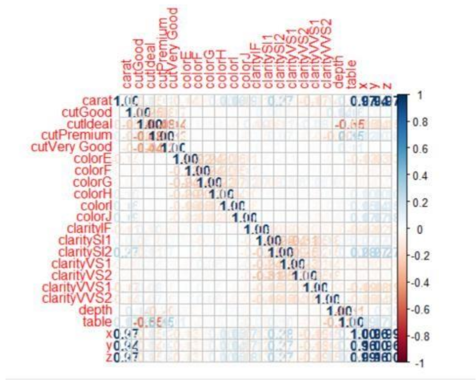
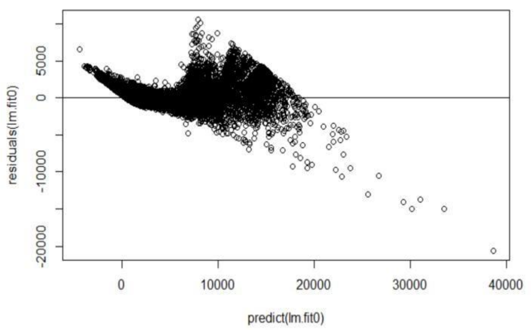


Fig.5 Correlation between single predictors in initial model

The residual plot of training data shows non-linearity in data as Fig.6 shown.



Shrinkage methods, Lasso and Ridge regression does not work well since no parameter shrink small enough to ignore the corresponding predictors.[4] Subset methods, best subset selection, forward and backward stepwise selection, are used with selection criteria of CP, BIC and R^2 to choose the best fit subset of predictors to remain in the model. And under the inspiration of data visualization plot, intercepts are removed since the line nearly cross the original point. It leads to the model without y and z predictors and intercept which gives 1302139 of training MSE and 1216096 of estimated test MSE. However, the R^2 increases from 92% to 95.85% which improves dramatically. This model is called Model21.

Since MSEs are not improved, interaction terms are added separately into the model of carat or x with other predictors. The p-value of the added interaction terms and the R^2 for the models are shown as follows to fix the non-linear problem shown in residual plot. Interaction terms with above 96% and a p-value smaller than 0.05 are added to the current model.

Table 1 R2 above 96% and p-value for interaction term

Interaction term added in the current model	R^2	p-value
carat:cut	0.9611	< 2e-16
carat:color	0.9639	Some levels < 2e-16
carat:clarity	0.9717	< 2e-16
x:cut	0.9605	< 2e-16
x:color	0.9633	Some levels < 2e-16
x:clarity	0.9702	< 2e-16

Since carat and x has a high correlation, this interaction term is also added to the Model21 as well as above terms. This model is called Model17. It decreases the MSEs by half that is 595552.7 for training data, 593042.4 for validation data and 610405.1 for C.V. estimated test MSE. R^2 reaches 98.1%. Since the training MSE and estimated test MSE are still similar to each other, it is not overfit so that qualifies the model selection criteria.

3.4. *Nonlinear models*

The residual plot of Model17 is shown as follow.

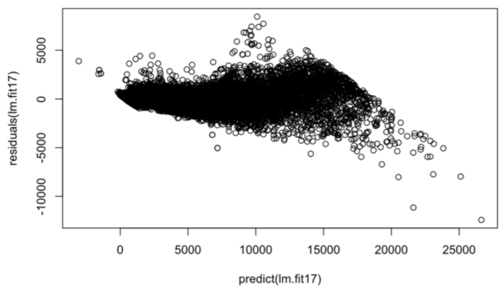


Fig.7 Residual plot of Model17

It still shows some outliers and non-linear relationship since the points do not form a horizontal band around zero. Therefore, nonlinear models are constructed.

The pruned tree with 6 terminal nodes has the training MSE of 1948964 and estimated test MSE of 1948964.[5]

Polynomial regression is applied with degrees from 2 to 5 based on Model21. The training MSE and the estimated test MSE is shown as Table 2.

Table 2 Polynomial regression with different degree

Degree	Training MSE	Estimated test MSE

2	1162500	1221956
3	1087814	1174913
4	1063071	1134305
5	1072774	1550505

According to the results, the lowest estimated test MSE is still higher than the linear model. Thus, the polynomial regression is not a good fit.

Spline regressions are applied with degree of freedom from 4 to 10 based on Model21 and spline regression with degree 10 gives the lowest training MSE of 1062536 and estimated test MSE of 1154814.[6]

Natural spline regressions with degree of freedom from 2 to 9 based on Model 21 are applied and the one with 9 degree gives the lowest training MSE of 1049669 and estimated test MSE of 1116726.

To wrap up, the nonlinear models are not suitable for the dataset as the estimated test MSE is too high, implying the low predictive accuracy of the nonlinear models.

4. Result and Discussion

The final model is Model17 we get in the linear part. Then we use tr_va data to fit this model and calculate the true test MSE on the test data. The result is encouraging as the R^2 on the test data achieves 0.9811 and difference between test MSE of 603833.2 and training MSE of 592182.6 is acceptable.

The final model is Price = carat + cut + color + clarity + depth + table + x + carat:cut + carat:color + carat:clarity +x:cut +x:color +x:clarity + carat:x with parameters shown in the table.

The model indicates that the price is affected by carat, cut, color, clarity, depth, table and x. Considering both the single term and the interaction term for each attribute, larger carat, better cut, better clarity, better color, smaller depth, smaller table and smaller x can lead to higher price, which conforms to the common sense.

However, from data visualization plot, as the weight increase, the data points are more scattered which indicates a quadratic formation. But polynomial regression did not give a satisfying result. The reason might be that only a small part of the dataset shows the scatter pattern compared with the whole dataset of more than fifty thousand diamonds. Most diamonds are not so heavy and therefore the data will mostly distributed on range of smaller carat. The scatter plot on the small-carat section indicates an obviously linear relationship. Thus, the linear model fits better than the nonlinear model on the whole dataset.

Table 3 Parameters of the final model

Predictor	**Para.**	**Predictor**	**Para.**
carat	-4039.014	cutFair	6804.717
cutGood	6487.471	cutIdeal	7428.431
Cutpremiun	8080.426	cutVery Good	8239.953
colorE	1263.565	colorF	671.327
colorG	575.086	colorH	1249.292
colorI	1417.310	colorJ	2592.538
clarityIF	-1899.250	claritySI1	5131.024
claritySI2	3302.777	clarityVS1	3124.533
clarityVS2	3831.450	clarityVVS1	3034.499
clarityVVS2	2599.145	depth	-66.094
table	-27.084	x	-27.281
carat:cutGood	391.478	carat:cutIdeal	1647.206
carat: cutpremiun	1378.577	carat: cutVery Good	1895.888
carat: colorE	596.943	carat: colorF	163.011
carat: colorG	-751.679	carat: colorH	-1428.383
carat: colorI	-2090.482	carat: colorJ	-3028.556
carat: clarityIF	9573.599	carat: claritySI1	9376.079
carat: claritySI2	6665.108	carat: clarityVS1	10051.921
carat: clarityVS2	9771.941	carat: clarityVVS1	12065.551
carat: clarityVVS2	11029.471	cutGood:x	30.490
cutIdeal:x	-252.080	cutpremiun:x	-351.244
cutVery Good:x	-457.012	colorE:x	-334.640
colorF:x	-193.257	colorG:x	-106.895
colorH:x	-200.700	colorI:x	-213.227
colorJ:x	-352.775	clarityIF:x	-338.959
claritySI1:x	-1981.432	claritySI2:x	-1393.376
clarityVS1:x	-1571.735	clarityVS2:x	-1711.464
clarityVVS1:x	-1654.501	clarityVVS2:x	-1495.811
carat:x	1001.630		

5. References

1. Shivam, A. (2017). Kaggle, "Diamonds".
2. James, G., Written, D., Hastie, T. & Tibshirani, R. (2021). An introduction to statistical learning (2^{nd} ed.), New York, NY: Springer.
3. VIF in car package of R language, "Variance Inflation Factors".

4. Glmnet of R language, "Cross-validation for glmnet".
5. Tree of R language, "cv.tree", "prune.tree".
6. Splines2 of R language, "Regression spline functions and classes".

Authors Introduction

Miss. Xiran Wen

She is an undergraduate student majored in Science Data in The Chinese University of Hong Kong (Shenzhen). She is interested in data analysis based on statistics.

Mr. Zirui Su

He is an undergraduate student majored in Financial Engineering in The Chinese University of Hong Kong (Shenzhen). He is interested in financial analysis.

Miss. Qiqi Xu

She is an undergraduate student majored in Science Data in The Chinese University of Hong Kong (Shenzhen). She is interested in data analysis based on statistics.

Miss. Jiayi Fang

She is an undergraduate student majored in Financial Statistics in The Chinese University of Hong Kong (Shenzhen). She is interested in Financial analysis.

The Effect of Preprocessing with Gabor Filters on Image Classification Using CNNs

Akito Morita
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 image classification tasks, preprocessing of input images is one of the promising approaches for improving the performance. In this study, we investigated the effect of neuro-inspired preprocessing, such as Gabor filtering. We compared the averaged classification accuracy of multiple CNNs with the following three types of preprocessing: no preprocessing, Gabor filtering, and calculation of the difference between two Gabor filtered signals in the opposite color channels. The results showed that Gabor filtering increased the classification accuracy.

Keywords: convolutional neural network, Gabor filter, image classification, bio-inspired

1. Introduction

Convolutional neural networks (CNNs) are one of the most useful components in the field of image classification. In the CNNs that receive RGB images, features required for classification are extracted in the hidden layers. For example, Krizhevsky et al. showed that some of the kernels of the first layer in their CNN were selective to a particular image feature, such as spatial frequency, orientation, and color.[1] Using an input image in which features are extracted in advance could improve the accuracy and shorten the time required for learning. Actually, several models that combine CNNs and feature extraction with Gabor filters have been proposed.[2,3]

On the other hand, in the mammalian visual nervous systems, by which CNNs were inspired, the spatial characteristics of a type of neurons found in the primary visual cortex are modeled as Gabor filters.[4] Applying the processing in the early visual cortex to preprocessing of CNNs is expected to improve the performance of CNNs because the sophisticated visual functions in the brain, such as classification, are achieved by using the signals preprocessed by the primary visual cortex, which is the front end of the visual signal processing in the brain.

The purpose of this study is to investigate the effects of neuro-inspired preprocessing, such as Gabor filtering and calculation of the difference between opposite color signals, on CNNs. We compared the classification performance of CNNs with the following three types preprocessing: no preprocessing, Gabor filtering, and calculation of the difference between two Gabor filtered signals in the opposite color channels. In addition, in order to remove the influence of the topology of a particular CNN, we evaluated the performance in terms of the average value of the results of 50 CNNs whose parameters were determined randomly.

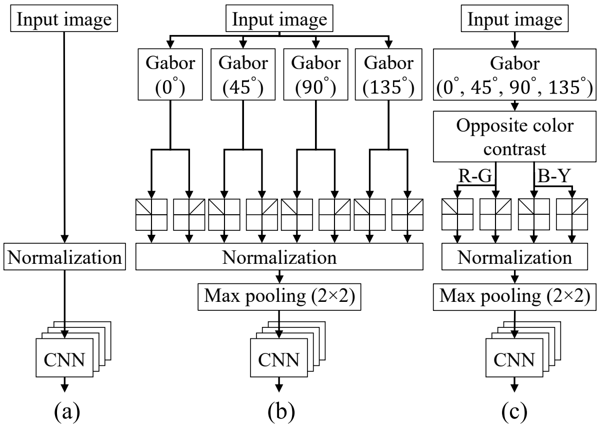


Fig. 1. Processing flow of classification. (a) Processing flow without preprocessing. (b) Processing flow with Gabor filtering. Four types of Gabor filters that enhance the following orientations were used: 0, 45, 90, and 135 degrees. The rectangles above the normalization represents rectification. (c) Processing flow with the calculation of the difference between two Gabor filtered signals in the opposite color channels. The processing to achieve the opposite color contrast is applied to the Gabor filtered signals of each orientation separately.

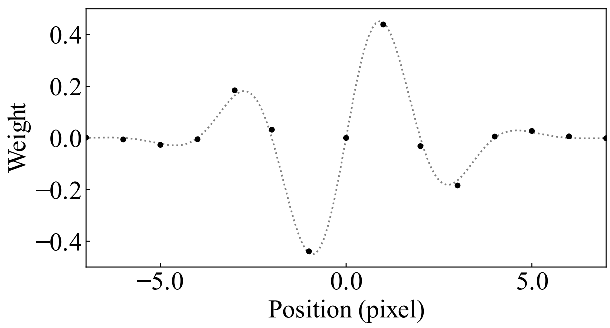


Fig. 2. Weights of the Gabor filter used in this study at $y = 0$. The black dots represent the weights of the Gabor filter, and the dotted line plots the fitted curve of the Gabor function.

2. Image classification algorithm

2.1. *Processing flow of classification*

Fig.1 shows the processing flow diagrams with three different preprocessing: no preprocessing (Fig. 1(a)), Gabor filtering (Fig.1 (b)), and calculation of the difference between two Gabor filtered signals in the opposite color channels (Fig.1 (c)). The input images are composed of RGB three channels.

The Gabor filters enhance edges with 0, 45, 90, and 135 degree orientations of each color channel. The component labeled "opposite color contrast" calculates the difference between the red and the green channels and the difference between the blue and the yellow (the average of red and green) channels, and this processing is applied to the Gabor filtered signals of each orientation separately. In addition, the outputs of Gabor filters and the outputs of the opposite color contrast processing are rectified as shown in Fig.1(b) and (c).

The average of outputs from 50 CNNs whose parameters were chosen randomly were used for evaluation; the parameters are the number of layers and neurons, and the size of kernels for convolution.

2.2. *Preprocessing*

2.2.1. *Gabor filtering*

A Gabor filter is a two dimensional spatial filter that enhances edges with a particular orientation, and was used to simulate the spatial characteristics of a simple cell, which is a well-studied neuron in the primary visual cortex.[4] In this study, we used Gabor filters that enhance edges with 0, 45, 90, and 135 degree orientations. The kernel of a Gabor filter that enhances edges with θ degree orientation is expressed as:

$$G(x,y) = A \cdot \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \times \cos\left(\frac{2\pi}{\lambda}(x\cos\theta + y\sin\theta) + \phi\right) \tag{1}$$

where x and y represent the coordinates in the kernel. The parameters were set as follows: $A = 0.506$, $\sigma = 2.0$, $\lambda = 4.0$, $\phi = \pi/2$. Fig. 2 shows the spatial characteristics at $y = 0$ of the filter whose $\theta = 0$.

The Gabor filtered images of RGB channels are hereinafter referred to as (G_R, G_G, G_B), respectively.

2.2.2. *Opposite color contrast*

Before calculating the difference between opposite color signals, the yellow channel is generated by

$$G_Y(x,y) = \frac{G_R(x,y) + G_G(x,y)}{2}, \tag{2}$$

where x and y represent the coordinates on the image.

The output signals of the opposite color contrast processing are given by $G_{R-G} = G_R - G_G$ and $G_{B-Y} = G_B - G_Y$. This processing is applied separately to each output image of the Gabor filter with four orientations.

2.3. *Convolutional neural network*

CNNs are neural networks (NNs) mainly composed of convolutional layers, fully connected layers, and pooling layers.

Fig.3 shows the architecture of the CNNs used in this study. The average of outputs from 50 CNNs whose parameters were chosen randomly were evaluated to remove the dependency on the topology of CNNs, and to investigate the effect of preprocessing alone. The following parameters were randomly selected: the number of layers and neurons, and the kernel size for convolution. Fig. 3(a) shows the entire network, whose components labeled “Randomly selected layers” were selected probabilistically from Fig. 3(b) and Fig. 3(c). The kernel size was randomly selected from {3, 5, 7}. The initial number of filters for convolutional layers was 32, and each time a layer was added, the number was doubled with a probability of 0.3. The number of neurons of the component labeled as “Fully connected layer” was chosen probabilistically from 2^n ($n \in \{5, 6, \ldots, 11\}$).

The rectified linear unit (ReLU) function[5] was used for the activation function of all layers except for the output layer, whose activation function was the softmax function.

The backpropagation algorithm and Adam[6] were used to train all weights. Adam is one of the most widely used optimization algorithms for NNs that changes the update amplitude of weights depending on the past updates of weights.

3. Experiments and results

3.1. *Experimental environment*

We implemented the methods described in Chapter 2 using python. The STL-10 dataset[8] was used for the evaluation. The STL-10 dataset is an image dataset consisting of 10 classes of images whose resolution is 96 × 96 pixels; each class has 500 training images and 800 test images. In order to perform the opposite color contrast processing, we removed grayscale images from the STL-10 dataset in this experiment.

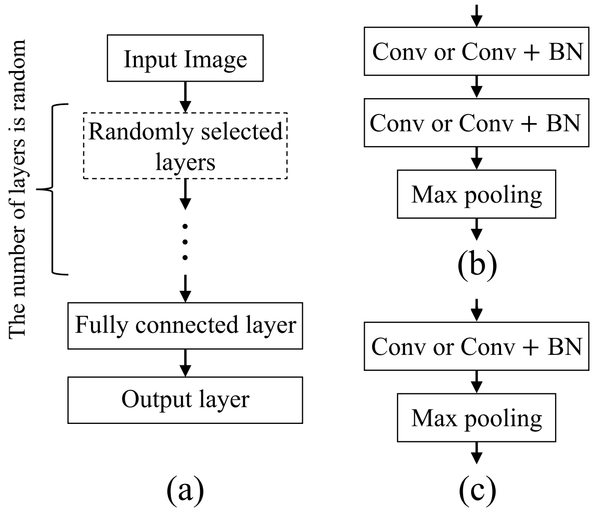


Fig. 3. Network architecture whose parameters are randomly selected. (a) Entire network architecture. Layers (b) or (c) are randomly embedded in the rectangle labeled “Randomly selected layers”. The number of neurons of the component labeled “Fully connected layer” is random. (b), (c) Candidate layers to be embedded in the network. Conv represents the convolutional layer and BN represents the batch normalization layer[7]. In the component labeled "Conv or Conv + BN", only the convolutional layer or the combination of the convolutional layer and the batch normalization layer is selected probabilistically. The parameters in the layers are also random (See body text for detail).

Table. 1. Averaged accuracy (%) of the three methods for the test data.

no preprocessing	Gabor filtering	Gabor filtering and opposite color contrast
51.687	57.879	51.867

The results were evaluated in terms of averaged accuracy obtained from the 50 CNNs described in Section 2.3. The number of epochs for learning was 50, and the size of the mini-batch was 128.

3.2. *Results*

Table.1 shows the accuracy of the three methods with different preprocessing described in Section 2.1. The method using Gabor filtering obtained the highest accuracy, whereas no significant difference in accuracy was found between the method without preprocessing and that with both Gabor filtering and opposite color contrast processing.

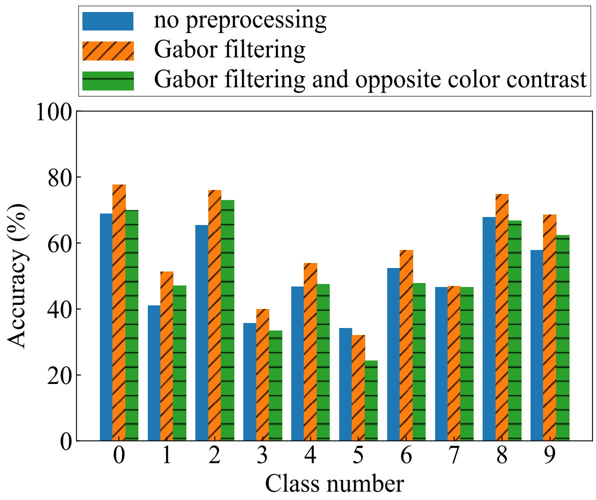


Fig. 4. Averaged accuracy of the three methods for each class.

Fig. 4 shows the accuracy for each class of the test data. As can be seen from Fig. 4, the method using Gabor filtering has the highest accuracy for all classes except for class 5.

4. Conclusion

In this study, we compared the classification accuracy of the average of multiple CNNs with the three types preprocessing to investigate the effect of neuro-inspired preprocessing. As a result, the highest accuracy was obtained by the method using Gabor filtering as a preprocessing, and no significant difference in accuracy was found between the method without preprocessing and that with both Gabor filtering and opposite color contrast processing. The results suggested that Gabor filtering used as a preprocessing for CNNs is an effective way to improve classification accuracy.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 19K12916.

References

1. A. Krizhevsky, I. Sutskever, G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Communications of the ACM, Vol.60, No.6, pp. 84-90, 2017.
2. M. M. Taghi Zadeh, M. Imani, B. Majidi, "Fast Facial emotion recognition Using Convolutional Neural Networks and Gabor Filters," 2019 5th Conference on Knowledge Based Engineering and Innovation, pp. 577-581, 2019.
3. S. Luan, C. Chen, B. Zhang, J. Han, J. Liu, "Gabor Convolutional Networks," IEEE Transactions on Image Processing, Vol. 27, No. 9, pp. 4357-4366, 2018.
4. S. Marcelja, "Mathematical description of the responses of simple cortical cells," Journal of the Optical Society of America, Vol. 70, No. 11 pp. 1297-1300, 1980.
5. V. Nair, G. E. Hinton, "Rectified Linear Units Improve Restricted Boltzmann Machines," Proceedings of the 27th International Conference on Machine Learning, pp. 807-814, 2010.
6. D. Kingma, J. Ba, "Adam: A Method for Stochastic Optimization," Proceedings of the 3rd International Conference for Learning Representations, 2015.
7. S. Ioffe, C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," Proceedings of the 32nd International Conference on Machine Learning, pp. 448-456, 2015.
8. A. Coates, H. Lee, A. Y. Ng, "An Analysis of Single-Layer Networks in Unsupervised Feature Learning," Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, pp. 215-223, 2011.

Authors Introduction

Mr. Akito Morita

He received his B.S. degree from the Department of Information Science and Technology, Osaka Institute of Technology, Japan in 2020. 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.

An Image Coding Algorithm with Color Constancy Using the Retinex Theory and the Naka-Rushton Equation

Shota Hisamitsu
Graduate School of Information 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 proposed an image-coding algorithm with color constancy based on the center / surround (C/S) retinex model and the Naka-Rushton (N-R) equation; this equation was used in place of the logarithmic function used in the original C/S retinex model in order to encode the intensity information efficiently. Using images acquired at different periods of a day, we compared the output of our algorithm with the original C/S retinex algorithm. The results suggested that the N-R equation can improve the color constancy performance.

Keywords: color constancy, retinex, image sensor, bio-inspired

1. Introduction

Color is a useful cue for image understanding, and is used in a wide range of image processing systems. However, estimating the intrinsic color of an object from an image can be a difficult task because the apparent color changes depending on the color of illumination. Therefore, methods for reducing the influence of illumination changes are required to use color information.

We humans have ability to perceive the appropriate color under a wide range of colored illumination, and the ability is called color constancy. The retinex model is a widely known computational model for achieving color constancy.[1] A lot of modified models have been proposed based on the retinex model.[2] Among such models, one of the most practical models is the center / surround (C/S) retinex model.[3] The model showed that applying logarithmic compression and a C/S antagonistic spatial filter, such as difference of Gaussians (DoG) filter, to three color channels (RGB) of images is a promising way to achieve color constancy. Hardware implementation of a part of the retinex model was also proposed.[4,5]

In the previous study, we have developed a compact image sensor system with color constancy based on the retinex model.[6] The system was composed of an image sensor and a field-programmable gate array (FPGA) with the following functions required by the retinex model: logarithmic compression and DoG filtering. The system was compared with the gray world (GW) algorithm, which is a widely used white-balancing algorithm,[7] in terms of the coded color information and showed a better performance.

However, the color representation of the system was less stable in a particular environment including outdoors. One of the possible reasons for this is the characteristics of the logarithmic compression, whose compression ratio is extremely high in the bright region whereas the ratio is relatively low in the dark region. When a large number of gray levels is assigned to dark regions, the gray levels for representing the target object can be insufficient. Efficient compression method is required.

In the present study, we proposed an image coding algorithm with color constancy using the retinex model and the Naka-Rushton (N-R) equation.[8] Employing the N-R equation in place of the logarithmic function is a possible solution for the problem of the compression ratio described above because the balance of the compression ratio is controllable by changing a particular parameter in the N-R equation. We examined the effect of the parameter on the represented color and compared the proposed algorithm with the original C/S retinex model in terms of the coded color information.

2. Color constancy algorithm

2.1. *Center / Surround retinex model*

Fig. 1 (a) shows the flow of the color constancy algorithm based on the C/S retinex model. This algorithm consists mainly of the following three steps: logarithmic transformation, Gaussian filtering, and subtraction of two Gaussian filtered images. These procedures are applied separately to the RGB three channels.

First, the light intensity I, which is the product of the illumination intensity E and the object reflectance ρ, is compressed logarithmically. This process separates the product of the terms into a sum form as follows:

$$\begin{aligned} L(x,y) &= \log I(x,y) \\ &= \log E(x,y) + \log \rho(x,y)\,. \end{aligned} \tag{1}$$

Here, (x,y) represents the coordinates of the image.

Next, Gaussian filters $g(x,y;\sigma)$ with two different standard deviations (σ_1, σ_2 ($\sigma_1 < \sigma_2$)) are applied to the logarithmic image.

Finally, a DoG image is given by the subtraction of the two Gaussian filtered images:

$$\begin{aligned} X(x,y) &= g(\sigma_1) * L(x,y) - g(\sigma_2) * L(x,y) \\ &= g(\sigma_1) * \log E(x,y) - g(\sigma_2) * \log E(x,y) \\ &+ g(\sigma_1) * \log \rho(x,y) - g(\sigma_2) * \log \rho(x,y)\,. \end{aligned} \tag{2}$$

Here, $*$ represents convolution, and the Gaussian kernel $g(x,y;\sigma)$ is denoted as $g(\sigma)$. In this study, we designed filters so that the integrals of the weights of these two filters equal, and therefore, the uniform change of illumination light should be canceled out.

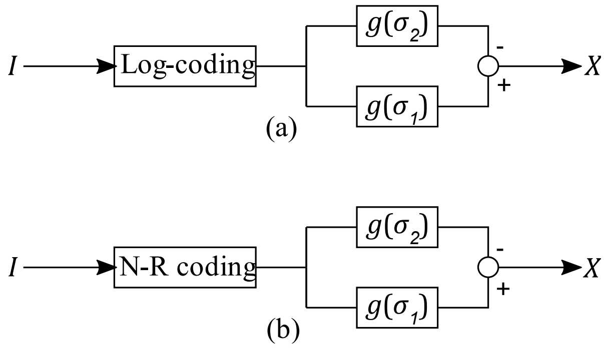


Fig. 1. Processing flow of the C/S retinex model. (a) Original C/S retinex model. (b) Proposed C/S retinex model with N-R equation.

The above procedures are applied separately to the RGB channels, yielding images X_R, X_G, and X_B, respectively, in which the influence of illumination is reduced.

In this study, we proposed an image coding algorithm with color constancy based on the C/S retinex model by replacing the logarithmic transformation with the N-R equation (Fig. 1 (b)).

2.2. *Naka-Rushton equation*

It is known that the response of photoreceptor cells of the biological retina to light stimuli follows the N-R equation[9]:

$$V = V_m \frac{I^n}{I^n + I_h^n}, \tag{3}$$

where V is the response of photoreceptor cells, V_m is the maximum response value, I is the light intensity, I_h is the parameter that changes depending on the adaptation state. n is a fitting parameter whose values for actual animals are between 0.7 and 1. In this study, we set $n = 1$.

Fig. 2 shows the input-output relationship of the logarithm and the N-R equation. Compared to the logarithm, the N-R equation spends less values to express the low-luminance information, and more values to express the high-luminance information. The problem of the logarithmic transformation, in which too much values are assigned to low-luminance information, can be alleviated by the above property of the N-R equation. The

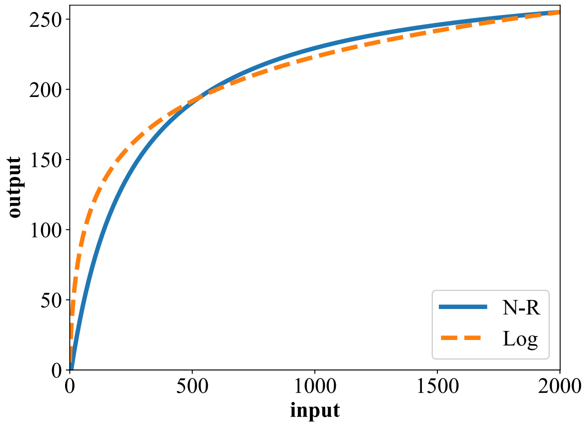


Fig. 2. Input-output relationship of the logarithmic function and the Naka-Rushton equation. The dashed and solid lines plot the logarithmic function and the Naka-Rushton equation, respectively.

N-R equation also has the advantage that the response characteristics can be easily controlled by changing I_h.

Also, the N-R equation with $n = 1$ is approximated by logarithmic function at the neighborhood of $I = I_h$, and therefore, the cancelation of illumination described in section 2.1 can be also applied to the proposed model for a particular range of light intensity.

2.3. *Parameters for Naka-Rushton equation*

The algorithm proposed in this study uses the N-R equation to compress the input image I. Eq. (4) is the implemented form of the equation, and is obtained by adding parameter I_{min} to Eq. (3) and by setting $n = 1$.

$$V = \begin{cases} V_m \dfrac{(I - I_{min})}{(I - I_{min}) + I_h} & (I_{min} \leq I \leq I_{max}) \\ 0 & (I < I_{min}) \\ 255 & (I > I_{max}) \end{cases} \quad (4)$$

where I_{min} and I_{max} represent the minimum and maximum pixel values to be converted. Values for I_{min} and I_{max} were determined by the cumulative histogram $H_c(k)$ of the input image, where k is the pixel value, instead of using the actual minimum and maximum pixel values for I_{min} and I_{max} directly because these values can fluctuate greatly for each frame. For convenience, the minimum values of k that satisfy inequality (5) and (6) were used for I_{min} and I_{max}, respectively.

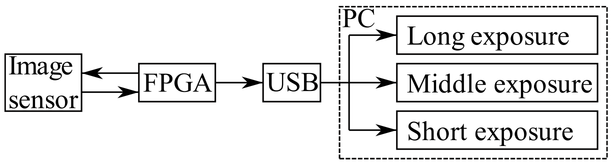


Fig. 3. System structure. The FPGA sends images acquired with three different exposure times to the PC, and the three images are recorded separately in the PC.

$$H_c(k) > \frac{N_T}{256} \quad (5)$$

$$H_c(k) > N_T - \frac{N_T}{256} \quad (6)$$

Here, N_T represents the total number of pixels. In this study, $N_T = 19200$.

V_m was determined so that the maximum value of V became 255 and is given by:

$$V_m = \frac{255(I_{max} - I_{min} + I_h)}{I_{max} - I_{min}}$$

I_h is a parameter that affects the characteristics of the converted image. In this study, we investigated how the performance of color constancy was influenced by the value of I_h. The minimum value of k that satisfies inequality (7) was used for I_h:

$$H_c(k) > N_r. \quad (7)$$

The value of N_r was selected from 1000 to 19000 in the present study. The performance of color constancy was evaluated by changing N_r.

3. Experiments and Results

3.1. *Environment for experiment*

Fig. 3 shows the system structure used in this study. The system mainly consists of a CMOS image sensor (OmniVision, OV5642) and an FPGA (Xilinx Artix-7 XC7A100T). Three images with eight-bit data of 160 × 120 pixels acquired by the image sensor with three different exposure times are sent to a PC via the FPGA and a USB interface, and are recorded in the PC. These

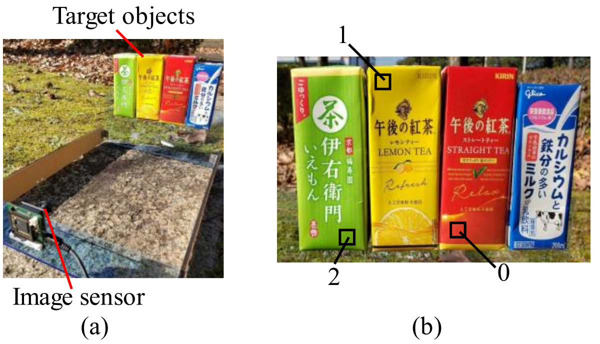


Fig. 4. (a) Outdoor environment for the experiment. (b) Target objects (green, yellow, red, and blue). Three regions (0, 1, and 2) indicated in the figure was used in the experiment.

three images are integrated into one image by replacing overexposure pixels, which was defined as pixels whose values exceeded 240, with pixels acquired with a shorter exposure time. We applied the C/S retinex algorithm with the N-R equation, hereinafter referred to as N-R C/S retinex, and the original C/S retinex algorithm to the image data. We evaluated the results in accordance with the method explained in Section 3.2.

Fig. 4 (a) shows the outdoor environment for the experiment, in which images were acquired at different periods of a day (i.e., at the daytime and the evening). Beverage cartons (Fig. 4 (b)) of four colors (green, yellow, red, and blue) were used as subjects.

3.2. *Data analysis*

In this study, we investigated the effect of the value of I_h in the N-R equation on the color constancy performance by changing N_r, and compared the color constancy performance of the N-R C/S retinex and the original retinex. The color constancy performance was evaluated in terms of the hue value in the HSV color space of the image processed by the retinex model.

Here, the HSV color space is a color space expressed by hue, saturation, and brightness. Hue expresses the type of color in the range of 0 to 360. Saturation expresses color vividness in the range of 0 to 100 %. The brightness expresses the brightness of the color in the range of 0 to 100 %.

3.3. *Results*

Fig. 5 shows the relationship between N_r and the hue in regions 0, 1, and 2, which are shown in Fig. 4, of the

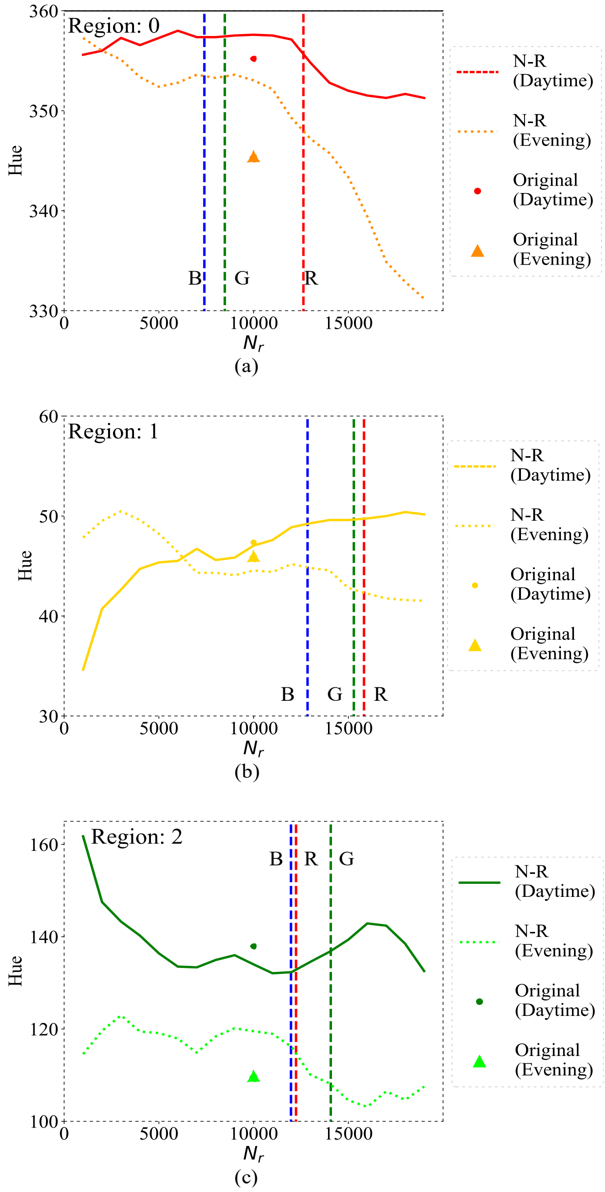


Fig. 5. Relationship between $\boldsymbol{N_r}$ and the hue of the output image. (a)(b) and (c) show the relationship in regions 0, 1, and 2 (shown in Fig. 4), respectively. The vertical dashed lines represent the position in the cumulative histogram of the target region. The color of the lines indicates the corresponding color channel.

output image. The difference between the hue in the daytime and that in the evening of the N-R C/S retinex was small when N_r was around the half of the total number of pixels, i.e., $N_r = 9600$, and the difference was smaller than that of the original C/S retinex depending on the value of I_h. On the other hand, when the value of N_r

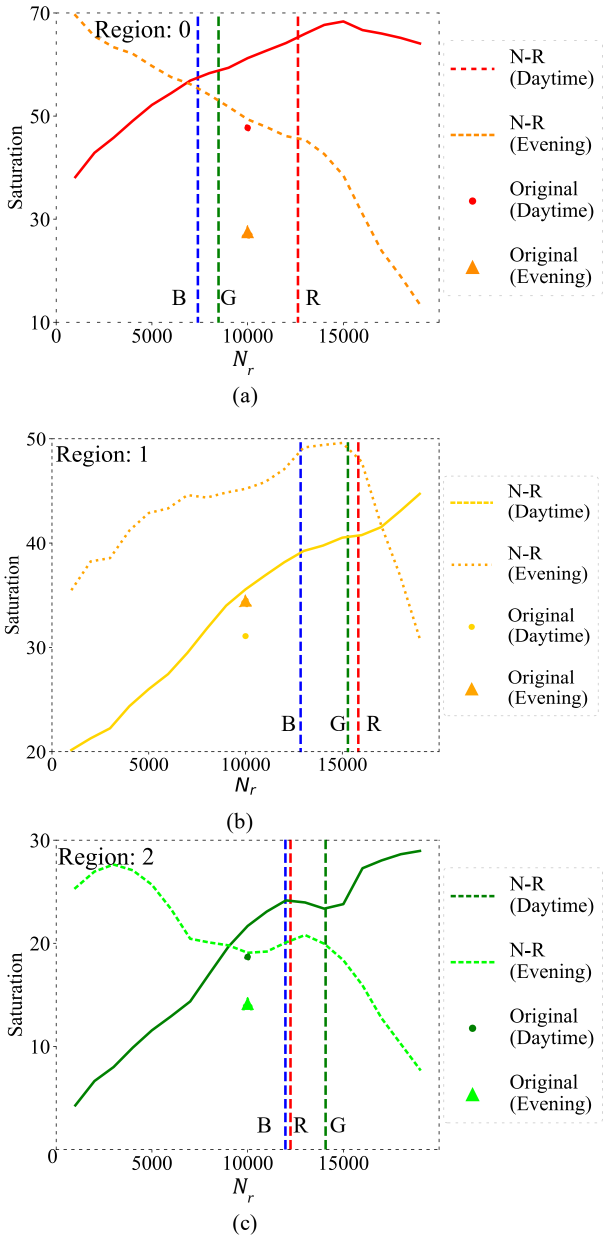


Fig. 6. Relationship between $\boldsymbol{N_r}$ and the saturation of the output image. (a)(b) and (c) show the relationship in regions 0, 1, and 2 (shown in Fig. 4), respectively. The vertical dashed lines represent the position in the cumulative histogram of the target region. The color of the lines indicates the corresponding color channel.

was deviated significantly from this value, the difference could be large.

Fig. 6 shows the relationship between N_r and the saturation. The saturation of the N-R C/S retinex tended to be higher than that of the original retinex, depending on the value of I_h. A lower value of saturation could lead to instability of the hue.

4. Conclusion

In this study, we proposed a color constancy algorithm by combining the N-R equation and the C/S retinex model. We evaluated the effects of a parameter in the N-R equation on the color constancy performance. The results suggested that the color constancy performance can be improved by setting I_h to an appropriate value. Further study is required for finding the method for setting an appropriate value of I_h.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 19K12916.

References

1. E. H. Land, J. J. McCann, “Lightness and retinex theory,” the Optical Society of America, vol.1, No.1, pp.1-11, 1971.
2. E. H. Land, “An alternative technique for the computation of the designator in the retinex theory of color vision,” Proceedings of the National Academy of Sciences of the United States of America, vol. 83, No. 10, pp. 3078-3080, 1986.
3. D. J. Jobson, Z. Rahman, G. A. Woodell, “Properties and performance of a center/surround retinex,” IEEE Transactions on Image Processing, vol. 6, No. 3, pp. 451-462, 1997.
4. A. Moore, J. Allman, R. M. Goodman, “A real-time neural system for color constancy,” IEEE Transactions on Neural Networks, vol. 2, no. 2, pp. 237-247, 1991.
5. K. Shimonomura, S. Kameda, A. Iwata, T. Yagi, “Wide-dynamic-range APS-based silicon retina with brightness constancy,” IEEE Transactions on Neural Networks, vol. 22, No. 9, pp. 1482-1493, 2011.
6. Kento Misaka, Hirotsugu Okuno, “FPGA implementation of an algorithm that enables color constancy,” Proceedings of the IEEE/SICE International Symposium on System Integration, pp. 991-995, 2020.
7. G. Buchsbaum, “A spatial processor model for object colour perception,” Journal of the Franklin Institute, vol. 310, no. 1, pp. 1–26, 1980.
8. K. I. Naka, W. A. H. Rushton, “S-potentials from luminosity units in the retina of fish (Cyprinidae),” J.Physiol. vol. 185, pp. 536-555, 1966.

Authors Introduction

Mr. Shota Hisamitsu

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.

A Fast Image Sensor System with an Efficient Multi-Scale Gaussian Filtering Circuit

Yuki Yamaji
Graduate School of Information Science and Technology, Osaka Institute of Technology,
1-79-1 Kitayama Hirakata, Osaka 573-0196, Japan

Akito Morita
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 designed an efficient multi-scale Gaussian filtering circuit whose coefficients of the standard deviation are selectable from any multiple of the square root of two. We also developed an image sensor system composed of a CMOS image sensor and a field-programmable gate array that contains the proposed filtering circuit. The system provided eight images whose resolution is 160 × 120 filtered by different scales of Gaussian filters at 156 frames / second.

Keywords: multi-scale, Gaussian filter, FPGA, image sensor, bio-inspired

1. Introduction

Efficient extraction of multi-scale visual features is required for a wide range of image processing systems. In particular, many useful bio-inspired visual processing algorithms, such as saliency-based visual attention[1] and the scale-invariant feature transform[2], rely on multi-scale Gaussian and/or Gabor filters because the early stages of visual nervous system are modeled as such filters[3,4] based on physiological studies (Refs. 5 and 6 for examples).

In the visual nervous system, the spatial property modeled as the Gaussian and the difference of Gaussians (DoG) with multiple scales is implemented in the retina[3], and all the visual functions are achieved using the output signal of the retina. Therefore, multi-scale Gaussian filtering is an essential component for implementing a wide range of vision-based tasks. Although the computation itself of spatial filtering is simple, the computational cost required for spatial filters with a large kernel is very high because the computation contains a large number of multiply-accumulate (MAC) operations. One of the most widely used solutions for this problem is separating the two-dimensional (2D) filter into two one-dimensional (1D) filters when the filter kernel is separable. Several studies on hardware implementation of 2D filters also use the technique to achieve efficient filtering architecture (Refs. 7 and 8 for examples).

The separable filter technique is applicable to multi-scale Gaussian filtering. In addition to the technique, the recursive algorithm proposed by Burt can further reduce the computational cost of filtering.[9] In the case of hardware implementation, the algorithm is an effective way to save the hardware resources required.

In the present study, we designed a multi-scale Gaussian filtering circuit based on the recursive filtering algorithm[9]. In order to save hardware resources, we used fixed-point computation, and investigated the adequate bit width with which errors are negligible even after repetitive recursive filtering. We also developed an image sensor system composed of a CMOS image sensor and a field-programmable gate array (FPGA) that contains the proposed filtering circuit.

2. Filtering algorithm

2.1. *Fast recursive filtering algorithm*

The fast recursive filtering algorithm implemented in this study is based on the algorithm described in Ref. 9. The algorithm applies separated 1D filters whose kernel is very simple to the input image. A set of kernels that differ in width but not in weights are convolved recursively. Fig. 1(a) shows how the kernel is convolved with the input data for the first two scales. The basic size of the kernel is five, and therefore, only five MAC operations are needed for 1D filtering in each scale. The kernel with the doubled spacing between two weights (d in Fig. 1(a)) is applied to the output image of the previous filter to increase the scale by one. One scale increment doubles the standard deviation of the Gaussian filter.

2.2. *Algorithm implemented in the circuit*

In order to enable configuration of the standard deviation of the Gaussian with a smaller step, we modified the recursive filtering algorithm described in section 2.1. Fig. 1(b) shows the kernels of the first three scales of the modified algorithm. Different from the original algorithm, the basic size of the kernel of the modified algorithm is three, and one-scale increment in the modified algorithm corresponds to multiplying the standard deviation by the square root of two. The spacing (d in Fig. 1(b)) is expressed as:

$$d = 2^m \tag{1}$$

$$m = floor\left(\frac{n-1}{2}\right) \tag{2}$$

where n represents the scale. The following two sets of weight values are employed depending on the scale:

$$(w_0, w_1) = \begin{cases}(a_1, b_1)(n \text{ is odd}) \\ (a_2, b_2)(n \text{ is even})\end{cases} \tag{3}$$

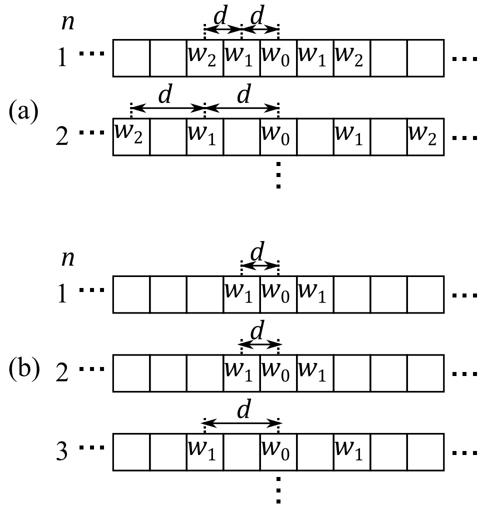


Fig. 1. Relationship between scale $\boldsymbol{n}$ and a set of kernels. (a) Original algorithm. (b) Modified algorithm.

The following constraints should be met for normalization:

$$a_1 + 2b_1 = 1 \tag{4}$$

$$a_2 + 2b_2 = 1 \tag{5}$$

The following set of parameters was used in this study:

$$(a_1, b_1) = (0.625, 0.1875) \tag{6}$$

$$(a_2, b_2) = (0.5, 0.25) \tag{7}$$

In the circuit implemented in this study, the filters are applied using fixed point computation to save hardware resources. An appropriate bit width was examined using the simulation described in section 3.

3. Simulation

3.1. *Filter kennel*

We examined the impulse response of the modified Gaussian filtering algorithm by using Python. Fig. 2 shows the impulse response of the filter whose scales are 1, 2, 3, and 4. The dashed lines plot fitted Gaussian functions. Gaussian functions are well-approximated by the weights achieved by the algorithm.

Table 1 shows the standard deviations of the fitted Gaussian functions. The ratio of the standard deviation of the adjacent scales is approximated by the square root of two, for a larger scale in particular.

Table. 1 Standard deviations of the fitted Gaussian functions.

scale n	1	2	3	4	5
σ_n	0.639	1.057	1.482	2.201	3.128
σ_n/σ_{n-1}	-	1.654	1.402	1.485	1.421

scale n	6	7	8	9	10
σ_n	4.467	6.322	8.971	12.694	17.964
σ_n/σ_{n-1}	1.428	1.415	1.419	1.415	1.415

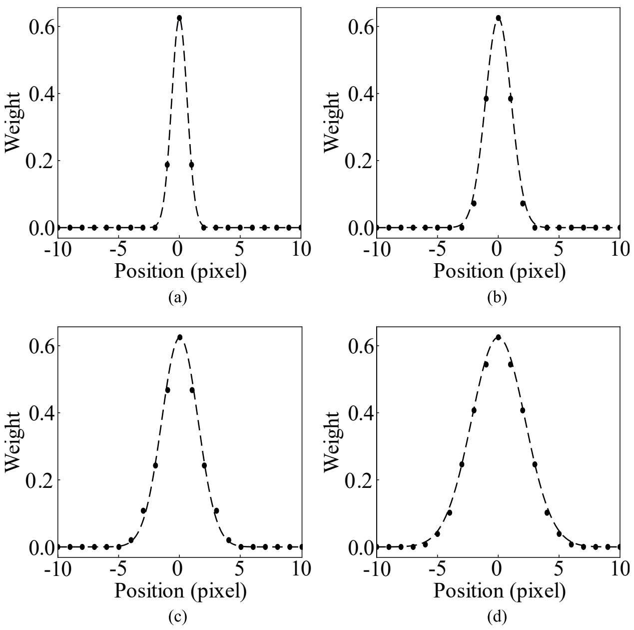


Fig. 2. Impulse responses of the Gaussian filters. Solid black circles represent the weight and the dashed line plots the fitted curve of the Gaussian function. (a)(b)(c) and (d) shows the weights of the filters whose scales are 1, 2, 3, and 4, respectively.

3.2. *Bit width*

We examined an appropriate bit width for implementing the modified Gaussian filtering algorithm by comparing the error between the filtered image computed with double-precision floating point and that computed with a fixed point. In this experiment, the bit width of the input and output images was eight, and the bit width with which the accumulation of the error was negligible was investigated. Here, the round-to-even method was employed. We chose several images from SIDBA[10] as an input image in this experiment. Fig. 3(a) shows an example of the relationship between the scale and the maximum error in the filtered image. Although Fig. 3(a) shows the relationship for a single image shown in Fig. 3(b), the results for other input images were essentially the same. When eight and nine bits were used, the error exceeded 0.5; this magnitude of the error can affect the least significant bit (LSB). On the other hand, the error never exceeded 0.5 when ten or more bits were used.

Taking these results into account, we adopted ten-bit computing for hardware implementation.

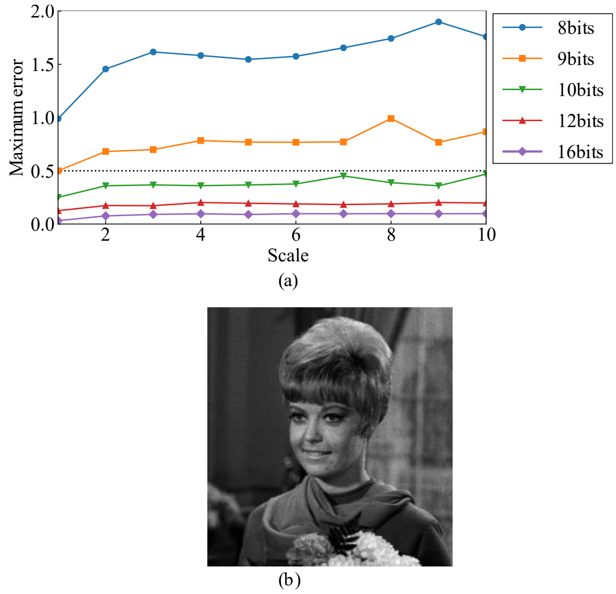


Fig. 3. (a) Relationship between the filter scale and the maximum error in the filtered image. (b) Input image[10] used in this experiment.

4. Circuit and system implementation

4.1. *System structure*

We implemented the multi-scale Gaussian filtering algorithm described in section 2.2 into an FPGA (Xilinx Artix-7 XC7A100T), and developed an image sensor system composed of a CMOS image sensor (Omni Vision OV5642), the FPGA, and a USB interface. Fig. 4 and 5 show the structure and the appearance of the image sensor system, respectively.

The CMOS image sensor in the system acquires image with 160 ×120 pixels at a maximum rate of 156 frames / second. The multi-scale Gaussian filtering circuit in the FPGA provides up to eight Gaussian filtered images with different standard deviations; the filtered images are sent to PC via the USB interface.

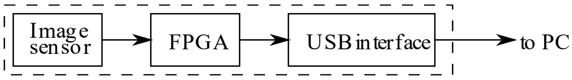


Fig. 4. System structure of the image sensor system with the multi-scale Gaussian filtering circuit

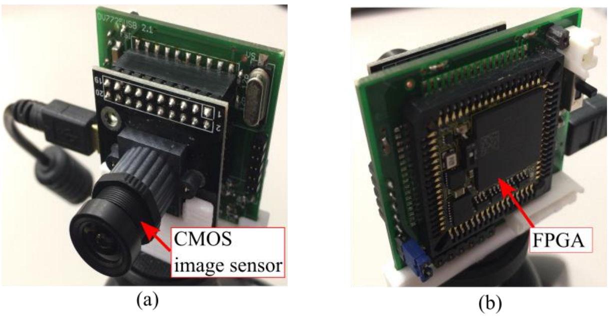


Fig. 5. Appearance of the image sensor system. (a) Front side. (b) Back side.

4.2. *Circuit implementation*

Fig. 6(a) shows the structure of the multi-scale Gaussian filtering circuit implemented in the FPGA.

First, the address counter circuit generates a read-out address signal, and an image is read out from the raw image buffer (not shown in the figure). Here, the image is read out along the X axis and is sent to the component labeled "cascade of 1D filters" through the selector, which is the trapezoid labeled "S". The 1D filter circuit applies the filter along the X axis, and the filtered image is recorded in the temporal RAM.

Next, the address counter circuit generates a read-out address signal again, and the image is read out from the temporal RAM. Here, the image is read out along the Y axis and is sent to the cascade of 1D filters through the selector. The 1D filter circuit applies the filter along the Y axis, and the filtered image is recorded in the temporal RAM. The repetition of filtering along the X and Y axes provides Gaussian filtered images with a large scale.

The scale counter outputs n, which is the scale of the current filter applied to the image. The RAM selector compares the current scale n with the target scales, and records the filtered image if the current scale equals one of the target scales. Components in the later stage (not shown) can read the filtered images by sending signals to the address ports (ADDR1 and 2 in the figure).

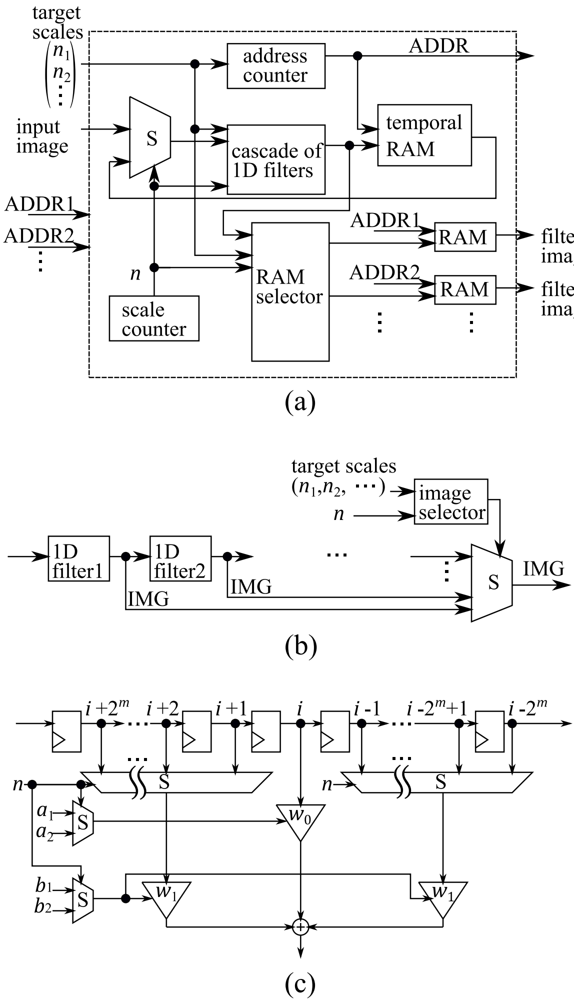


Fig. 6. (a) Structure of the entire multi-scale Gaussian filtering circuit. (b) Structure of the component labeled "cascaded 1D filters". (c) Structure of the 1D filter circuit. The weights and the width of the filter are determined by the current scale $\boldsymbol{n}$.

Fig. 6(b) shows the structure of the component labeled "cascade of 1D filters". This circuit is composed of a cascade of 1D filter circuits, each of which performs a single scale of Gaussian filtering. The selector compares the current scale n with the target scales, and outputs the filtered image with the target scale. The number of the cascaded filter implemented in this study was two.

Fig. 7 shows the time chart of a cascade of two 1D filter circuits. This circuit performs two scales of

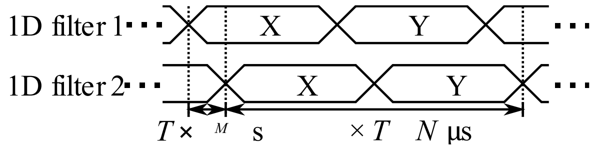


Fig. 7. Time chart of a cascade of 1D filters. ***T*** and ***N*** represent the clock cycle and the total number of pixels, respectively. ***M*** represents the maximum number of ***m*** in equation (2).

Gaussian filtering at $T \times 2^M + 2 \times T \times N$ µs, where T represents the clock cycle, M represents the maximum number of m in equation (2), and N represents the total number of pixels. In the present study, $= 0.02\ \mu s$, $M = 4$, and $N = 19200$ pixels, and the total time required is about 768 μs. One additional cascaded 1D filter performs one more scale of filtering in 0.02×2^M µs.

Fig. 6(c) shows the structure of the 1D filtering circuit. This circuit in Fig. 6(c) consists of the following five components: a shift register that stores data of $2^{M+1} + 1$ pixels, two selectors that select the pixel to be multiplied by the weights at the current scale n, two selectors that select the value of the weights for w_0 and w_1, three multipliers, and an adder that sums the three multiplied values. The values described at the nodes of the shift register $(i + 2^m, \ldots, i - 2^m)$ denote the indices of the stored pixel data, and i represents the center of the stored data.

The circuit in Fig. 6(c) operates as follows. First, image data read from the RAM are stored in the shift register. Next, three multipliers multiply values of the pixels with the following indices by weights chosen by the selectors; the indices are i, $i - 2^m$, and $i + 2^m$. Finally, the adder at the bottom sums three multiplied values and rounds the result to a ten-bit value.

4.3. *Evaluation*

We evaluated the multi-scale Gaussian filtering circuit alone and the entire image sensor system with the circuit using the methods described below.

We evaluated the multi-scale Gaussian circuit by comparing the output of the circuit in response to a step input shown in Fig. 8(a) with the result computed with double-precision floating point. In this evaluation, we implemented a circuit that generates a step image shown in Fig. 8(a) in the FPGA. Fig. 8(b)(c) show the differences between the circuit output and the floating-point operation for Gaussian scales 5 and 6, respectively.

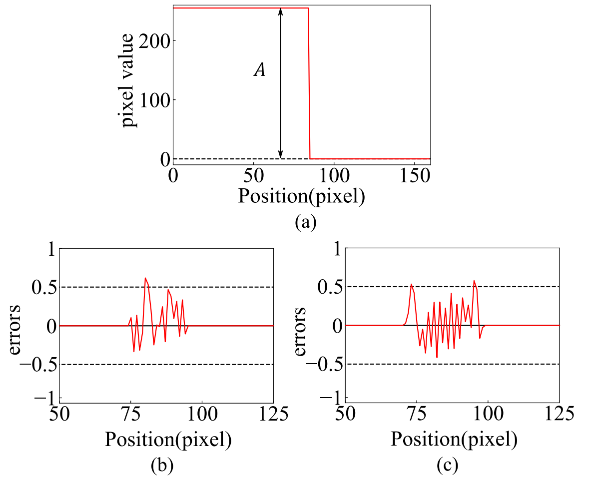


Fig. 8. Difference between the output of the circuit in response to a step input and the result computed with double-precision floating point. (a) Input pattern. The value of ***A*** is 255. (b)(c) The difference for Gaussian scales 5 and 6.

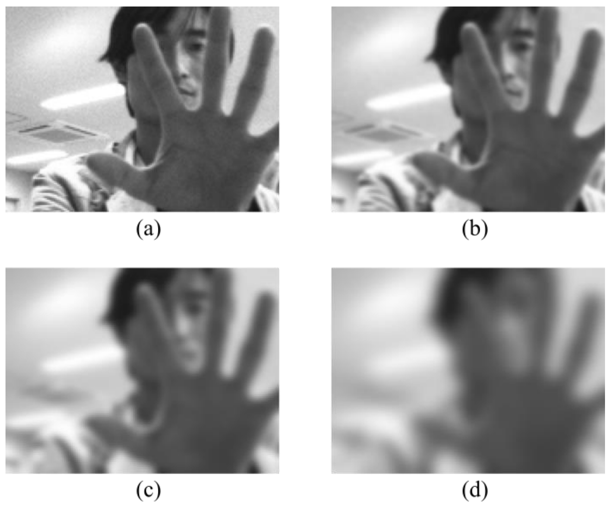


Fig. 9. Output images of the sensor system. (a) Raw image. (b)(c)(d) Gaussian-filtered images whose scales are 2, 4, and 6.

Here, the bit width of the image sent from the circuit is eight, and the input step amplitude is the maximum in eight bits, i.e. A = 255, and therefore, the error was larger than that shown in Fig. 3(a). However, the error was still smaller than the LSB.

We evaluated the entire image sensor system by presenting objects (a human in this experiment) to the image sensor. Fig. 9(a) shows the raw image, and Fig. 9(b), (c), and (d), show the Gaussian-filtered images whose scales are 2, 4, and 6. Gaussian-filtered images with a larger scale are smoothed more widely. These output results were obtained at 156 frames / second.

5. Conclusion

In the present study, we developed a multi-scale Gaussian filtering circuit and developed an image sensor system with the filtering circuit. The filtering circuit alone and the entire system were evaluated by providing a step input and by presenting real-world scenes. The results showed that the filtering circuit provided filtered images whose error is smaller than the LSB in response to a step input and that the image sensor system provided eight-scale Gaussian-filtered images at 156 frames / second.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 19K12916.

References

1. L. Itti, C. Koch, E. Niebur, "A Model of Saliency-Based Visual Attention for Rapid Scene Analysis," IEEE Trans. Pattern Analysis and Machine Intelligence ,Vol.20, pp.1254-1259, 1998.
2. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision, Vol. 60, pp. 91-110, 2004.
3. R. W. Rodieck, "Quantitative analysis of cat retinal ganglion cell response to visual stimuli," Vision Research, Vol. 5, pp. 583-601, 1965.
4. S. Marcelja, "Mathematical description of the responses of simple cortical cells," Journal of the Optical Society of America, Vol. 70, No. 11, pp. 1927-1300, 1980.
5. S. W. Kuffler, "Discharge Patterns and Functional Organization of Mammalian Retina," Journal of Neurophysiology, Vol. 16, pp. 37-68, 1953.
6. D. H. Hubel, T. N. Wiesel, "Receptive fields of single neurones in the cat's striate cortex," The Journal of Physiology, Vol. 148, pp. 574-591, 1959.
7. D. Llamocca, C. Carranza, M. Pattichis, "Separable FIR Filtering in FPGA and GPU Implementations: Energy, Performance, and Accuracy Considerations," Proceedings of the 21st International Conference on Field Programmable Logic and Applications, pp. 363-368, 2011.
8. A. K. Joginipelly, D. Charalampidis, "Efficient separable convolution using field programmable gate arrays," Microprocessors and Microsystems, Vol. 71, 102852, 2019.
9. P. J. Burt, "Fast filter transform for image processing," Computer Graphics and Image Processing, Vol. 16, No. 1, pp. 20-51, 1981.
10. M. Sakauchi, Y. Ohsawa, M. Sone, M. Onoe, "Management of the standard image database for image processing researchs (SIDBA)", ITEJ Technical Report, Vol. 8, No. 38, pp. 7-12, 1984. (in Japanese)

Authors Introduction

Mr. Yuki Yamaji

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.

Mr. Akito Morita

He received his B.S. degree from the Department of Information Science and Technology, Osaka Institute of Technology, Japan in 2020. 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 Robotic vision system emulating fixational eye movements and retinal sampling

Takanori Yotsumoto
Dept. of Human Intelligence Systems, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka, 808-0196, Japan

Yuki Hayashida
Dept. Inform. Eng., Grad. Eng., Mie University,
1577 Kurima-machiya, Tsu, Mie 514-8507*, Japan*

Shinsuke Yasukawa
Dept. of Human Intelligence Systems, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka, 808-0196, Japan

E-mail: yotsumoto.takanori573@mail.kyutech.jp, hayashida@ai.info.mie-u.ac.jp, s-yasukawa@brain.kyutech.ac.jp

Abstract

Recent studies on visual physiology have suggested that fixational eye movements contribute to the information processing function of the retina. In this study, we built a robotic vision system that emulates the characteristics of fixational eye movements and retinal sampling as an assistant tool for clarifying information processing through the collaboration of eye movements and neural circuits. The proposed robotic vision system consists of two galvanometers, a field-programmable gate array (FPGA) board, and a high-speed, high-resolution camera. We measured the frequency response of this robotic vision system, and the results showed that the system satisfied the requirements for emulating the frequency characteristics of biological eye movements. In addition, we generated command signals in the FPGA to emulate biological eye movements and confirmed the satisfactory operation of the system.

Keywords: Fixational eye movement, Robot vision, Galvano mirror, Bio-inspired vision system

1. Introduction

The visual nervous system contains a large number of neurons and synapses, and it is difficult to monitor each of them simultaneously with sufficient resolution in physiological experiments. Therefore, it is useful to develop a neural-circuit model that describes the characteristics of neurons and their interactions based on physiological knowledge, and to use a simulation approach to reproduce the activity of a group of neurons. In simulation experiments, we can systematically investigate the functions and properties of each element of the visual network by freely controlling the parameters that are difficult to control in physiological experiments.

Eye movements, body movements, and adaptations that occur at all levels of the visual system can be viewed as active actions of the biological visual system that constantly seeks to acquire new information. If the biological visual system is considered as a system that performs visual computation while interacting with the external world, it is essential to perform the biological vision simulation in real time.

A methodology that attempts to understand the intricate essence of a system that undergoes complex interactions with the external environment by constructing an artificial system with a similar structure as the hardware and verifying its operation is called a constructivist methodology. Since Mead et al. demonstrated an analog circuit that mimics the structure of the retina [1], research has been extensively conducted to reproduce the structure of retinal neurons and their network circuits as faithfully as possible on silicon. These studies are significant from the standpoint of understanding biological vision and engineering applications (reviewed in [2]). A real-time simulation of the visual system has also been realized [3].

Although there have been attempts at hardware emulation of the visual nervous system, there are few examples of hardware implementation of eye movements, which are the entry points for light input to the visual system, for simulation purposes. As mentioned earlier, if we consider the visual system as an active process, it is important to realize the hardware that can mimic eye movements using a bio-visual simulation method. In addition, findings from visual physiology and visual psychology experiments suggest that fixational eye movement contributes to the realization of retinal neural circuit functions, such as high spatial resolution [4], foreground–background separation [5], and target prediction [6].

In this study, we developed a robotic vision system that can receive visual input considering eye movements, for simulation of biological vision and evaluated the frequency characteristics of the visual input during fidgety microtremors in the robotic vision system.

2. Fixational eye movements and retinal sampling

Human fixation eye movement is composed of three components: tremor, drift, and microsaccade.

Tremor is an eye movement of small amplitude occurring at a frequency of 90 Hz or less, whereas microsaccade is an eye movement with a velocity of approximately 120 °/s occurring at a frequency of 2.0 Hz or less. Drift is a low-speed eye movement that occurs simultaneously with tremor and in between microsaccades [7]. To reproduce these movements using an engineering approach, it is necessary for a robot vision system to have a cutoff frequency of 90 Hz or higher and to move at a speed of 120 °/s or higher.

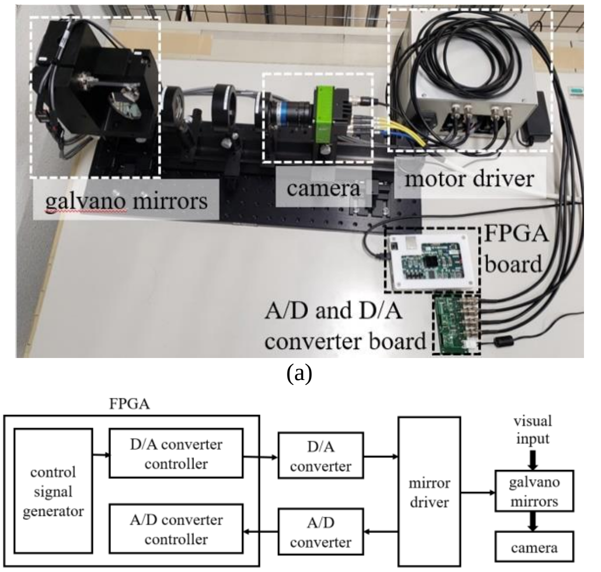
The spatio-temporal characteristics of retinal sampling are used as a reference for selecting a camera module. The density of cone cells in the photoreceptor layer of the central fossa of the retina is approximately 150000 cells/mm^2 [8] and the relationship between the angle and length of the retina is approximately 300 µm/° [9]. Therefore, the spatial-resolution requirement of the robotic vision system is estimated to be 116 pixels/°. In this study, we assumed that one pixel of the image sensor corresponds to one cell of the retinal photoreceptor layer. Next, the temporal characteristics of retinal ganglion cells are considered to determine the frame rate of the camera module. Because the incidence of spike response of ganglion cells in the retina is once every 5 ms (200 Hz) on average, the frame rate requirement of the robotic vision system is 200 Hz or higher.

3. System integration

3.1. *Hardware*

Fig. 1(a) and Fig.1 (b) show the hardware setup and block diagram of the proposed system, respectively. The proposed robotic vision system consists of an optical pan-tilt mechanism with two rotating mirrors, its driver module, digital circuits for control signal generation, A/D and D/A converters, and a high-speed, high-resolution camera.

In general, in a pan/tilt camera, the line of sight is mechanically controlled using a rotary actuator. However, the mechanical control of gaze direction using a rotary actuator does not provide sufficient high-speed performance. In this study, we optically controlled the line-of-sight direction at high speed using a rotating mirror with two axes (pan and tilt), based on the study presented in [10]. In [10], a pupil-transmission system was installed between the camera and the mirror to resolve the trade-off between light intensity and angle of view. Signals for pan-tilt control are generated by digital circuits in the field-programmable gate array (FPGA); these signals are then D/A converted and input to the driver module for the motor driving the mirror. Signals representing the rotation angle of the mirror are output from the encoders installed on each axis and input to the FPGA via A/D converters. The mirror speed



(a)

(b)

Fig. 1 (a) Hardware setup of the proposed system. (b) Block diagram of the proposed system.

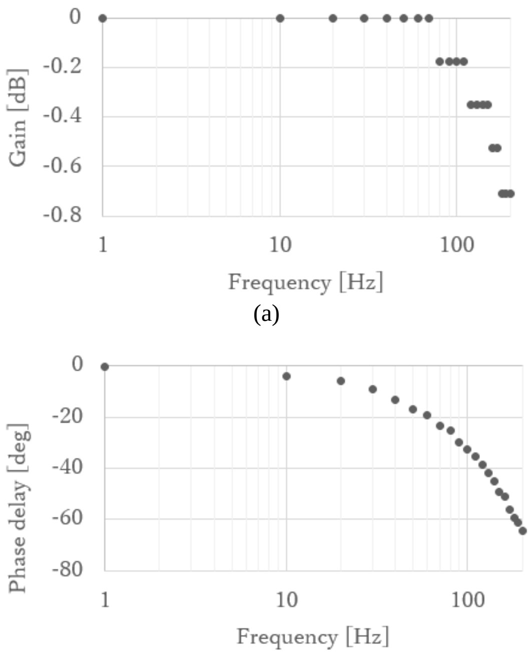
of the system, operating frequency, operating range, and lens diagonal angle of view of the pupil transfer system are set to 280 °/s, 200 Hz, ±20 °, and 38.6°, respectively.

Fig. 2 depicts the measured frequency response of the optical pan-tilt system. Sinusoidal signals of various frequencies with an amplitude of 0.5 V (±1.26°) are input to the optical pan-tilt system, and the angle signal of the mirror is measured by the encoder. The gain and phase characteristics of the optical pan-tilt system are measured on the basis of the relationship between the input and output signals. The cutoff frequency is higher than the target value of 200 Hz, which meets the requirements of the system.

The camera module used is a JAI SP-25000C-CXP4A, which has a resolution of 5120 × 3832 pixels when operated at a resolution of 200 fps. The spatial resolution of this camera is 140 pixel/°, which satisfies the requirement (116 pixel/°) for the robotic vision systems described in Section 2.

3.2. *Control signals for drift motion*

In this study, we focused on the drift motion even during fixational eye movements. Considering the temporal characteristics of the drift motion, we Imple-



(a)

(b)

Fig. 2 Frequency response of optical pan-tilt system. (a) Gain characteristics. (b) Phase characteristics.

mented a digital circuit to generate control signals for the optical pan-tilt system based on the study by Hasegawa et al. [11]. To reproduce the motion, the displacement vector from the gazing center position was used as the control signal for the mirror control motor.

Fig. 3(a) shows a schematic of the circuit configuration used to generate the displacement vector. The circuit consists of a pulse generator (PG), a pair of 63-bit linear feedback shift registers (LFSRx and LFSRy), and a pair of summing modules (Σ). The pulse generator sends a sequence of four pulses as a clock signal to the linear feedback shift registers; this shifts the information in the registers by four bits. LFSRx and LFSRy generate M-sequences that are independent of each other. The summing module outputs the number of bit 0's in each connected linear-feedback shift register. The displacement vectors Δx and Δy are obtained by rounding the number of bits to integer values. Figure 3(b) shows a typical eye position transition pattern, where one pixel corresponds to the size of a cone cell in the central part of the human retina (visual angle of 0.6 arcmin). The width and velocity distribution of the fixational eye movement are reasonable [12].

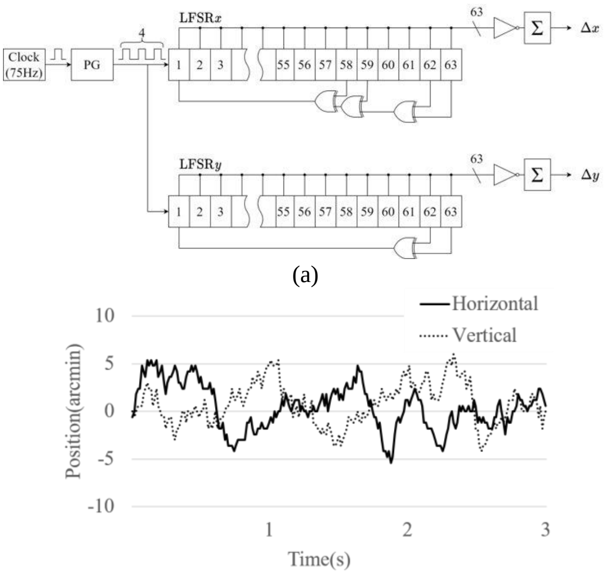


Fig. 3 (a) Digital calculation circuit for displacement vector of fixational eye movements. (b) Variation of the displacement vector generated by the digital circuitry.

4. Experiments and Results

Experiments and analyses were performed using the following procedure: First, the two images shown in Fig.4 were input to the robotic vision system in two cases viz. drifting motion of the mirror and stationary motion of the mirror. We applied a three-dimensional Fourier transform to each of the obtained movies to transform them into a spatio-temporal frequency space. Next, at each time frequency, the data was two-dimensionalized by averaging over each of the same spatial frequencies (center images of Fig. 4). Finally, by integrating in the time direction at each spatial frequency, a graph was created to represent the spatial frequency characteristics of the system.

The results demonstrate that the spatial frequency response of the visual input with the mirror stationary is that of a natural image, with the amplitude spectrum decaying to 1/f. However, the low spatial frequency response of the visual input with the mirror in motion was flatter than that with the mirror stationary. This is similar to the effect of eye movement on the visual input of a biological vision system.

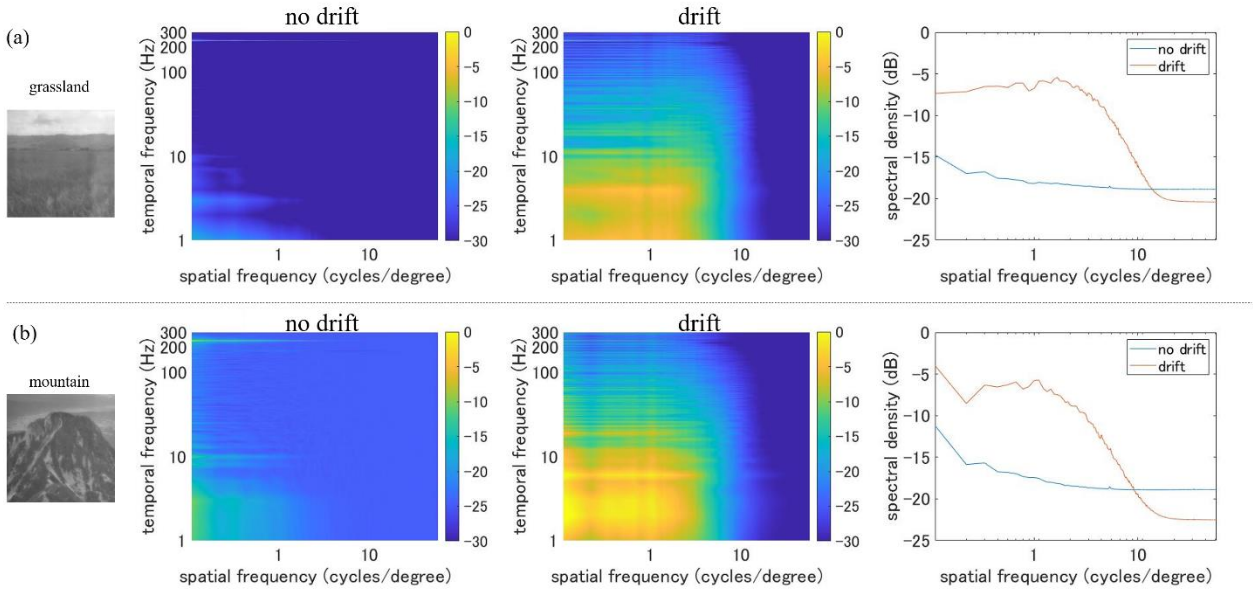


Fig.4 Spatio-temporal frequency characteristics of visual stimuli when the system is presented with images of (a) a meadow and (b) mountain, as input images. The images in the center left and center right show the spatio-temporal frequency characteristics when the mirror is stationary and when the mirror is moving, respectively. The graphs on the right show the spatial frequency response when the amplitude component is integrated in the time frequency direction.

5. Conclusion

We developed a robotic vision system that simulates the characteristics of human fixational eye movements and retinal sampling. To achieve fast response characteristics, such as those of fixational eye movements, we integrated an optical pan-tilt system using galvanometer mirrors and a digital circuit that generates control signals simulating eye movements and emulated dynamic characteristics similar to the drift component of actual human eye movements. In addition, we measured the spatio-temporal characteristics of the retinal input with and without fixational eye movement using the proposed system. From the results, we confirmed that the proposed system can simulate the spatio-temporal frequency characteristics of the visual input during human fixational eye movements. This system is expected to be a useful tool for understanding the relationship between eye movements and visual information processing in the biological visual system.

Acknowledgements

This research was partly supported by JSPS KAKENHI Grant Number 20K19895.

References

1. C. Mead, "Neuromorphic electronic systems." Proceedings of the IEEE, Vol. 78, No. 10, pp. 1629-1636, 1990.
2. S. C. Liu, et al.(Eds.), "Event-based Neuromorphic Systems." John Wiley & Sons, 2014.
3. Y. Hayashida, et al., "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.
4. M. Rucci, & J. D. Victor, "The unsteady eye: an information-processing stage, not a bug." Trends in Neurosciences, Vol. 38, No. 4, pp. 195-206, 2015.
5. S. A. Baccus, et al. "A retinal circuit that computes object motion." Journal of Neuroscience, Vol. 28, No. 27, pp. 6807-6817, 2008.
6. A. Matsumoto, M. Tachibana, "Rapid and coordinated processing of global motion images by local clusters of retinal ganglion cells." Proceedings of the Japan Academy, Series B, Vol. 93, No. 7, pp. 524-524, 2017.
7. S. Martinez-Conde, S. L. Macknik, & D. H. Hubel, "The role of fixational eye movements in visual perception." Nature reviews neuroscience, Vol. 5, No. 3, pp. 229-240, 2004.
8. R. W. Rodieck, "The First Steps in Seeing." Sunderland, MA: Sinauer Associates, Vol. 1, 1998.
9. N. Drasdo, & C. W. Fowler, "Non-linear projection of the retinal image in a wide-angle schematic eye." The British Journal of Ophthalmology, Vol. 58, No. 8, pp. 709-714, 1974.
10. K. Okumura, H. Oku, & M. Ishikawa, "High-speed gaze controller for millisecond-order pan/tilt camera." In 2011 IEEE International Conference on Robotics and Automation, pp. 6186-6191, May. 2011.
11. J. Hasegawa, & T. Yagi, "Emulation of retinal cell responses during fixational eye movements." IEICE Electronics Express, Vol. 7, No. 3, pp. 184-189, 2010.
12. I. Murakami, "Correlations between fixation stability and visual motion sensitivity." Vision Research, Vo. 44, No. 8, pp.751-761, 2004.
13. X. Kuang, et al., "Temporal encoding of spatial information during active visual fixation." Current Biology, Vol. 22, No.6, pp. 510-514, 2012.

Authors Introduction

Mr. Takanori Yotsumoto

He received his B.S. degree from the Department of Biorobotics, Kyushu Sangyo University, Japan in 2020. 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. ShinsukeYasukawa

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 bench-test system of the visual prostheses utilizing retino-morphic spikes as the driver signals of intracortical microstimulation

Ryosuke Okada[1], Shinnosuke Ishikawa[2], Tetsufumi Tasaki[2], Tetsuya Yagi[2,3], Yuki Hayashida[1,2]

[1]*Dept. Inform. Eng., Grad. Eng., Mie University, 1577 Kurima-machiya, Tsu, Mie 5148507, Japan*

[2]*Dept. Electric. Electro. Inform. Eng., Grad. Eng., Osaka University, 2-1 Yamada-oka, Suita, Osaka 5650871, Japan*

[3]*Dept. Electric. Electro. Eng., Sch. Eng., Fukui University of Technology, 3-6-1 Gakuen, Fukui, Fukui 9108505, Japan*

E-mail: 421m506@m.mie-u.ac.jp, ishikawa@neuron.eei.eng.osaka-u.ac.jp, tasaki@neuron.eei.eng.osaka-u.ac.jp, yagi@fukui-ut.ac.jp, hayashida@ai.info.mie-u.ac.jp

Abstract

We developed a prototype bench-test system of the intracortical visual prostheses, in which retino-morphic point-process spike signals from our previously developed retina emulator were utilized for driving the microstimulation applied to the cerebral visual cortex. Substituting the stimulating electrodes with micro-LEDs, the system operations were verified through dry bench tests, in which the spatial pattern of stimulus outputs via 4096 channels was able to be dynamically controlled by stimulus images projected to the retina emulator.

Keywords: Visual prosthesis, Retino-morphic emulator, Point-process spike signal, Intracortical microstimulation

1. Introduction

It has been known for decades that the electrical microstimulation delivered to multiple focal sites in the primary visual cortex can be a reasonable approach for providing artificial vision to people with acquired blindness.[1,2] A recent clinical study in a blind patient reexamined and further confirmed the feasibility of the intra-cortical visual prosthesis.[3,4] Also, a recent psychophysical behavioral study in non-human primates suggested a possibility of evoking the shape perception with phosphenes by means of high-channel-count intra-cortical microstimulation.[5] For realizing a high-channel-count visual prosthesis, it is essential that the incoming information of visual scene should be processed, and delivered to the visual cortex, in a physiologically efficient and psychologically efficacious manner. It is considered that a possible solution for such a prosthesis can be learned from the biological visual system.

In the previous study, we developed a neuromorphic retinal circuit emulator[6] utilizing the Izhikevich model,[7] with which the point-process spike signals are output from 128-by-128 channels, at millisecond time resolution in response to visual events. In other previous studies, we also developed a prototype wireless system supporting the maximum of 4096 channels of intracortical microstimulation for use in animal physiological experiments.[8] In order to further advance the preclinical studies on the visual prosthesis, the present study aimed to integrate the retinal circuit emulator into the multi-channel microstimulation system, and to test the integrated operations of the system.

2. System Integration

As the first step of the system integration, a wired configuration was chosen for the present study. Fig. 1 outlines the wired integration. The first stage is the retinal circuit emulator (Fig. 1, uppermost), which captures the visual scene, performs the information processing, and outputs the point-process spike signals as "spike images". Subsequently, the data packets of the spike images sent from the emulator are decompressed in a single-board computer ("Raspberry Pi" in Fig. 1), and are fed to a Field-Programmable-Gate Array ("FPGA" in Fig. 1) that provides the digital data of the stimulation parameters and the controlling signals for the microstimulator ASIC chip.[8] In the final stage, the ASIC chip generates the microstimulation current pulses through the intracortical stimulating electrodes. In this study, only one chip of the microstimulator was employed for saving the hardware resource and avoiding any complications in experiments on the system operation.

Main roles of the hardware components in the integrated system are explained in the following sections.

2.1 Retinal circuit emulator for image processing and spike encoding

The details of our retinal circuit emulator have been described in the previous paper.[6] In brief, the emulator is an analog-digital hybrid hardware system consisting of a circuit board with the analog VLSI chip called silicon retina[9] (*a* in Fig. 1), an interface board with a SRAM memory chip (*b* in Fig. 1), and a FPGA board (*c* in Fig. 1). In the silicon retina, the incoming images are captured at 200 frame-per-second with the 128-by-128 array of the CMOS active-pixel sensors, and the spatial filtering found in the biological outer retina is emulated by the double-layered resistive network realized by CMOS analog circuits.[9] In the FPGA, the spatial center-surround antagonistic filtering, the linear temporal filtering, and the nonlinear gain function found in the biological inner retina are implemented by the digital signal processing.[10] Moreover, the spike firings in the retinal output neurons,

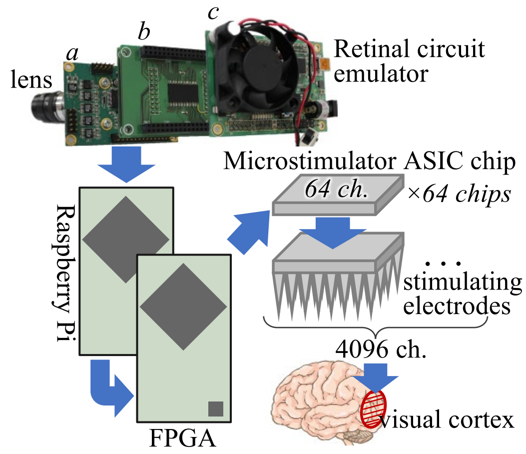


Fig.1. The hardware system configuration integrating the retinal circuit emulator [6] and the multi-channel microstimulation system.[8]

namely retinal ganglion cells, are emulated by the Izhikevich model.[7] By adjusting parameter values of this model and of the above-mentioned filters and functions, the spike responses to light stimuli in the biological retinal ganglion cells (Fig.2, middle trace in Fig. 2, "mouse RGC spike") could be reproduced in reasonable detail by the emulator (lower trace in Fig. 2, "emulator output spike"). Since the array size of the emulated retinal ganglion cells of a certain subtype is 128-by-128 in the original format,[6] the size is down-sampled to 64-by-64 channels to fit the number of the output channels of the microstimulation system[8]. The down sampling is performed in the emulator by summating the values in a 2-by-2 patch and thresholding the summated values.

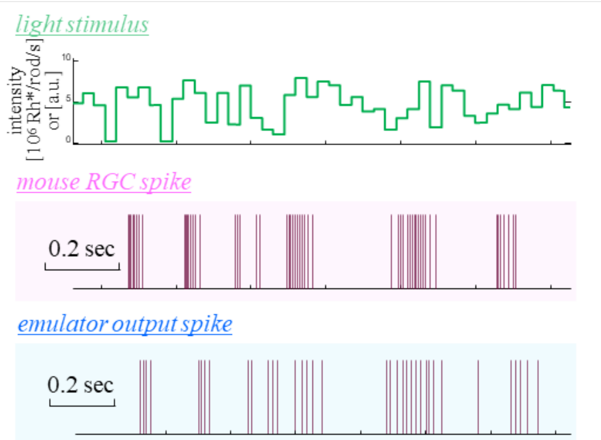


Fig.2. Example of the spike signals measured from the biological retinal ganglion cell (middle) and the output of the retinal circuit emulator (lower) in response to the light stimulus (upper)

2.2 Single-board computer for data managing

The retinal circuit emulator is connected to the single-board computer (Raspberry Pi 4 Model B, Raspberry Pi Foundation, Cambridge, U.K.) by the USB2.0 standard. The data communication is made by using the Python-3.7 USB communication module, PyUSB1.1.1. In the single-board computer, each frame in a 64-by-64 array of the spike images retrieved from the emulator is first sectioned into 64 blocks, each of which is an 8-by-8 array. Subsequently, the spike binary data for each of those blocks are sent to the next stage FPGA to be used as the position data of the microstimulation channels. By sequentially sending all datasets of the 64 blocks, 64-by-64 channels of microstimulation can be achieved. The achievable frame rate in this manner was approximately 30-40 fps.

Although, in the present system, the above-mentioned data managing by the single-board computer is rather simple, it would become complicated if two sets of the retinal circuit emulator are employed for the binocular configuration and if the visual cortices in the both hemispheres are the targets of microstimulation for expanding the field of the artificial vision.

2.3 FPGA for microstimulator ASIC chip control

The single-board computer is connected to the next stage FPGA (Spartan-6, Xilinx, CA, U.S.A.; XEM6010, Opal Kelly, OR, U.S.A.) with using the GPIO pins in the Raspberry Pi and the I/O pins in the FPGA board. In the FPGA, the digital codes in custom format are generated to control the microstimulator ASIC chip. Some of those determine the stimulus parameters and others are to control the circuit operations and memory registrations.[8] The spike binary data of a particular block out of the 64 blocks mentioned above were transcoded to the position data of the microstimulation channels. In addition, since an individual chip of the microstimulator is pre-assigned with an identification (ID) number of the 6-bit code, the ID code is also generated according to the location of a particular block out of the 64 blocks by the FPGA. The position data are once stored in the built-in register of the assigned microstimulator chip, and then stimulus current pulses are injected thought the output channels of the chip depending on the stored position data. The timings of the stimulus current injections are also controlled by the FPGA.

3. Dry Bench-Test of the Integrated Operation

3.1 Experimental setup

Fig. 3 shows a photograph the experimental setup for testing the integrated operation of the system. The upper panel shows the computer disply for the stimulus image presentation and the retinal circuit emulator in front of the display. The lower pannel shows the other hardware componets of the integrated system, nemaly, the Raspberry Pi, the FPGA board, the microstimulator ASIC chip, and a micro-LED array. As shown here, a 8-by-8 micro-LED array is used intead of 64 stimulating electrodes, and thus, spatio-temporal patterns of the microstimulation are examined by imaging the spatio-temporal patterns of the LED light with using a video camera. Since only one pair of the microstimulator chip and the micro-LED array was employed, the LED light patterns were repeatedly imaged for 64 times in sequential sessions, and those images were combined off-line to form the light patterns of 4096 LEDs.

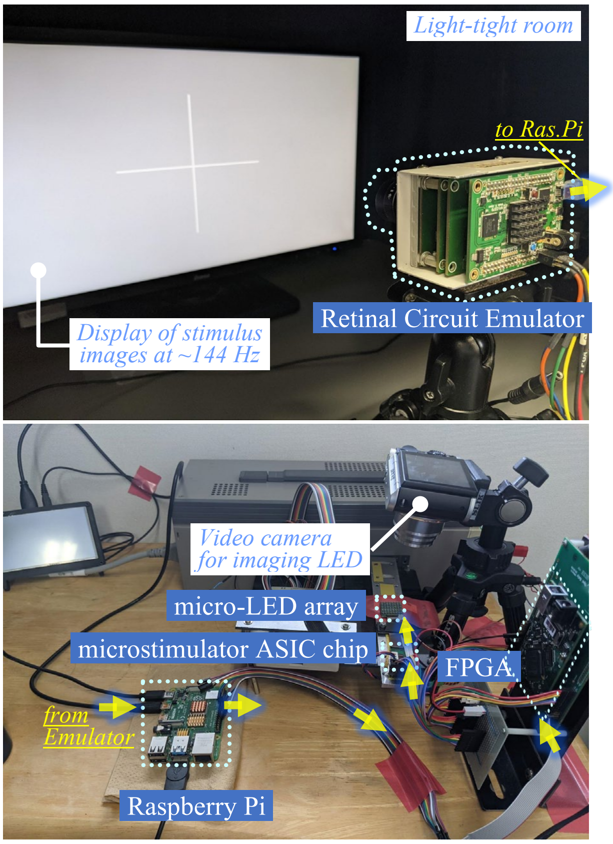


Fig.3. Experimental setup for the dry bench-test of the integrated system.

3.2 Experimental results

In the experiment, the stimulus images were presented as a 10-second movie, as shown in Fig. 4*A*. This movie was repeatedly projected to the retinal circuit emulator for 64 times. As explained in the previous subsection, the LED light were imaged by the video camera, and thus, a movie file created by this video camera was sectioned into 64 movie files, each of which had a duration of 10 seconds, and then was edited off-line. Fig. 4*B-C* shows time lapse images of the spike image data in the 64-by-64 array sent from the retinal circuit emulator to the single-board computer (*B*) and of the output light patterns of the 4096 LEDs (*C*). Since the transient response type of the biological retinal ganglion cells was emulated by the retinal circuit emulator in this experiment, the time lapse images were taken from the time points after the onset of switching the stimulus images from one to the next. As shown in these results, the spatial patterns of the LED light (*C*) showed reasonable correspondences to those of the spike images (*B*). In some portions, the LED light patterns differed from those of the spike images. This was mainly due to mismatch among the refresh timing of the display used for the spike images, the LED blinking timing at ~39 fps (frame interval of ~25.4 msec) and the image capturing timing in the video camera, during the 64 repetitions mentioned above, and also due to the stochastic nature of the spike timings in our retinal circuit emulator.[6] Nevertheless, these results demonstrated the proper operations of the integrated system.

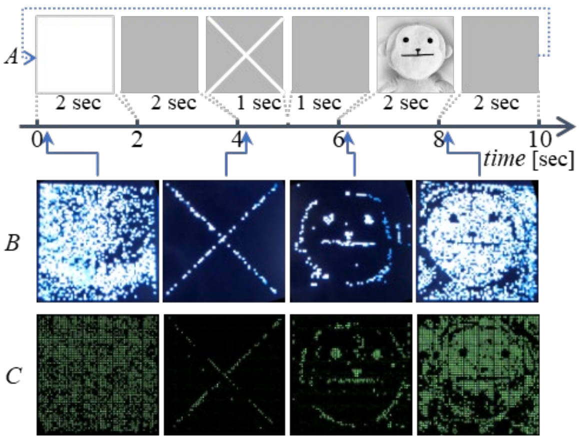


Fig. 4. Example of the experimental results. *A*) The sequence of the stimulus image presentation. *B*) The 64-by-64 spike images in response to the stimulus images shown in *A*. *C*) The final output patterns of the 64-by-64 LED light driven by the spike images shown in *B*.

4. Discussions

In the present study, we integrated the retinal circuit emulator into the multi-channel microstimulation system and verified the system operation through the dry bench-test. If the data communication rate between the retinal circuit emulator and the single-board computer is made faster, and/or if the data managing performed in the single-board computer is implanted in the FPGA, then the frame rate of the stimulus outputs is expected to be shortened. From the physiological point of view, 50 to 100 fps would be sufficient for exciting the neural circuits in the visual cortex with the microstimulation.[11,12]

Acknowledgements

Research partly supported by JSPS KAKENHI Grant Number JP20H00606, JP18K12059.

References

1. E. M. Schmidt, "Finding from chronic optic nerve and cortical stimulation," Visual Prosthetics: Physiology, Bioengineering, Rehabilitation (Ed: G. Dagnelie), New York, Springer, pp. 301-315, 2011.
2. E. M. Schmidt et al., "Feasibility of a visual prosthesis for the blind based on intracortical microstimulation of the visual cortex," Brain 119(2), pp.507-522,1996.
3. E. Fernández et al., "Visual percepts evoked with an intracortical 96-channel microelectrode array inserted in human occipital cortex." J. Clin. Invest. 131(23), e151331(15 pp.), 2021.
4. M.S. Beauchamp et al., "Raising the stakes for cortical visual prostheses." J. Clin. Invest. 131(23), e154983 (4 pp.), 2021.
5. X. Chen X et al., "Shape perception via a high-channel-count neuroprosthesis in monkey visual cortex." Science 370(6521), pp.1191-1196, 2020.
6. Y. Hayashida et al., "Retinal circuit emulator with spatiotemporal spike outputs at millisecond resolution in response to visual events. IEEE Trans. Biomed. Circuits Syst. 11(3), pp.597-611, 2017.
7. E. M. Izhikevich, "Simple model of spiking neurons." IEEE Trans. Neural Netw. 14(6), pp.1569-1572, 2003.
8. Y. Hayashida et al., "Cortical neural excitations in rats in vivo with using a prototype of a wireless multi-channel microstimulation system." Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., pp.1642-1645, 2015.

9. S. Kameda et al., “Real time image processing with an analog vision chip system,” Int. J. Neural Syst., 9(5), pp.423-428, 1999.
10. H. Okuno et al., “Real-time emulator for reproducing graded potentials in vertebrate retina.” IEEE Trans. Biomed. Circuits Syst. 9(2), pp.284-295, 2015.
11. Y. Hayashida et al., “Voltage-sensitive dye imaging of the visual cortices responding to electrical pulses at different intervals in mice in vivo,” Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., pp. 402-405. 2014.
12. T. Nomoto et al., “Imaging of population spikes induced by repetitive stimulus pulses in mouse cerebral slices in vitro,” Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., p.1109, 2017.

Authors Introduction

Mr. Ryosuke Okada

He received his Bachelor's degree from the School of Engineering, Mie University, Japan, in 2021. He is currently a Master course student of the Graduate School of Engineering in Mie University, majoring in information engineering. His research interests include neuromorphic vision system.

Mr. Shinnosuke Ishikawa

He received his Master’s degree from the Graduate School of Engineering, Osaka University, Japan, in 2021. He contributed to the present research when he was a Master course student of the Graduate School, majoring in electrical-electronic-and-information engineering.

Mr. Tetsufumi Tasaki

He received his Bachelor's degree from the School of Engineering, Osaka University, Japan, in 2020. He is currently a Master course student of the Graduate School of Engineering in Osaka University, majoring in electrical-electronic-and-information engineering.

Dr. Tetsuya Yagi

He received his Ph.D. in medical science from Nagoya University, Nagoya, Japan, in 1985. He is currently a Professor with the School of Engineering, Fukui University of Technology, Fukui, Japan. His research interests include neurophysiology on visual system and neuromorphic engineering.

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.

A basic study of how to exchange work shifts using reinforcement learning on a constructive nurse scheduling system

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 this paper, we propose a work revision method using reinforcement learning for a constructive nurse scheduling system. The constructive nurse scheduling system has the characteristic of having easy to understand shift schedule creation procedures and rules because the system does not use the evaluation value for the entire shift schedule. We have confirmed the possibility of improving the quality of the shift schedule by the proposed method.

Keywords: nurse scheduling, reinforcement learning, constructive search, work revision method

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].

In this paper, we propose a work revision method using reinforcement learning[3] for a constructive nurse scheduling system[4]. The constructive nurse scheduling system has the characteristic of having easy to understand shift schedule creation procedures and rules because the system does not use the evaluation value for the entire shift schedule. We have confirmed the possibility of improving the quality of the shift schedule by the proposed method.

2. Constructive nurse scheduling system

2.1. *Features*

The features of the constructive nurse scheduling system[4] 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.

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.[4] describe the following two procedures for the actual modification.

(1) Manual exchange of work shifts.

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) Change the shift schedule manually and create it automatically again.

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 propose a system that leans this exchange procedure using reinforcement learning.

3. Work Revision Method Using Reinforcement Learning

3.1. *Reinforcement learning*

In this section, we introduce Q-learning (QL)[5] which is one of the most popular RL methods. QL works by calculating the quality of a state-action combination, namely the Q-value, that gives the expected utility of performing a given action in a given state. By performing an action $a \in \boldsymbol{A}_{\mathrm{Q}}$, where $\boldsymbol{A}_{\mathrm{Q}} \subset \boldsymbol{A}$ is the set of available actions in QL and $\boldsymbol{A}$ is the action space of the RL agent, the agent can move from state to state. Each state provides the agent with a reward r. The goal of the agent is to maximize its total reward.

The Q-value is updated according to the following formula, when the agent is provided with the reward:

$$\begin{aligned}&Q\big(s(t-1),a(t-1)\big)\\&\leftarrow Q\big(s(t-1),a(t-1)\big)+\alpha_{\mathrm{Q}}\{r(t-1)\\&\quad+\gamma\max_{b\in\boldsymbol{A}_{\mathrm{Q}}}Q(s(t),b)\ -Q(s(t-1),a(t-1))\}\ (2)\end{aligned}$$

where $Q\big(s(t-1),a(t-1)\big)$ is the Q-value for the state and the action at the time step $t-1$, $\alpha_{\mathrm{Q}} \in [0,1]$ is the learning rate of QL, $\gamma \in [0,1]$ is the discount factor.

The agent selects an action according to the stochastic policy $\pi(a|s)$, which is based on the Q-value. $\pi(a|s)$ specifies the probabilities of taking each action a in each state s. Boltzmann selection, which is one of the typical action selection methods, is used in this research. Therefore, the policy $\pi(a|s)$ is calculated as

$$\pi(a|s)=\frac{\exp\left(Q(s,a)/\tau\right)}{\sum_{b\in A_{\mathrm{Q}}}\exp\left(Q(s,b)/\tau\right)}\ (3)$$

where τ is a positive parameter labeled temperature.

3.2. *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 (e.g., 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:

$$min\ \textstyle\sum_{\mathrm{n}}\sum_{\mathrm{v}}\mathrm{V}_{\mathrm{nw}}\qquad (3)$$

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 the semi-night or the 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, the 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 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 the day 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.

In this study, we propose a work revision method to determine an appropriate exchange procedure using reinforcement learning.

3.3. *RL Agent*

QL 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 semi-night, late-night shift, and holiday: $\mathrm{V_{nw}}$ (w=1,2,3,4), to be a Markov decision process. The number of possible actions is 4, which is the exchange of semi-night, late-night, holiday, and night shift.

1 step is defined as 1 exchange including unsuccessful cases, 1episode is defined as the time when the shift schedule reaches the target state. Here, the target state is defined as the sum of violations for all nurses and shifts $\sum_{\mathrm{n}} \sum_{\mathrm{v}} \mathrm{V_{nw}} = 0$, or when the situation does not improve even after an exchange. The positive reinforcement signal $r_t = 10$ (reward) is given only when the target state is reached and the reinforcement signal $r_t = 0$ at any other steps. At the start of each episode, the shift schedule will be in its initial state before the exchange.

4. Computational Experiments

4.1. *Nurse Scheduling Problem*

The proposed method is applied to a nurse scheduling problem similar to that of Kurashige et al.[4]. First, a three-shift system (day shift, semi-night shift, 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 as follows.

- 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 semi-night shift is 5.

Table 1. Evaluation of shift pattern for 2 days.

shift on previous day	shift on the day			
	day	semi-night	late-night	holiday
day	15	1	13	11
semi-night	0	5	0	12
late-night	0	8	5	4
holiday	23	3	0	17

- Constraints on team and skill level for night shifts:
  (5) At least 1 nurse per team should be assigned to each of the semi-night shift and late-night shift.
  (6) At least 3 nurses per team for consecutive semi-night and late-night shifts.
  (7) No more than 2 new nurses may work on the night shifts.
- Restrictions on the position:
  (8) All the work of the head nurse is designated.
  (9) The assistant head nurse works fewer nights.
- Restrictions on shift pattern:
  (10) The interval between holidays is limited to 5 days.
  (11) No more than 4 consecutive days off.
  (12) No more than 2 consecutive days of late-night and semi-night shifts.
  (13) The number of consecutive night shifts is limited to 3 days.

Next, Table 1 shows the evaluation of shift patterns for 2 days with M = 2.

4.2. *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.

The computational experiments have been done with parameters as shown in Table 2. In addition, all initial Q-values are set at 5.0 as the optimistic initial values.

4.3. *Results*

The average of the numbers of steps required to reach the target state and the average of the total number of violations when the target state is reached were observed during learning over 20 simulations, as described in Figs. 1 and 2, respectively.

It can be seen from Figs. 1, 2 that, (1) the total number of violations that can be reached by successive exchanges is 2, (2) the number of violations 3 that can be considered a local solution is sometimes obtained in the early stages of learning, but after more than 100 episodes, the number

Table 2. Parameters for experiments

Parameter	Value
α_Q	0.1
γ	0.9
τ	0.1

of violations 2 that can be considered an optimal solution is obtained in 7 exchanges.

Thus, we confirmed that the proposed method can reduce the number of violations of constraints without using the evaluation value for the entire shift schedule. In addition, we found that 2 nurses had violations of excessive holiday 1 in the modified shift schedule. Since these violations are not concentrated in 1 nurse, it is difficult to think that they lead to a sense of unfairness among general nurses.

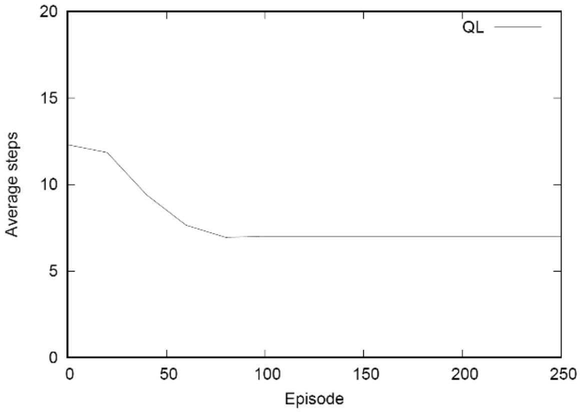


Fig. 1. Required steps to reach the target state.

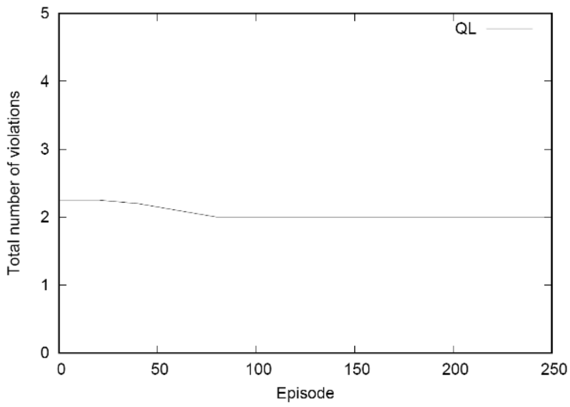


Fig. 2. Total nuber of violations when the target state is reached. Required steps to reach the target state.

5. Conclusion

In this paper, we proposed a work revision method using reinforcement learning for a constructive nurse scheduling system. Through computational experiments, we confirmed the possibility of improving the quality of the shift schedule by the proposed method.

Our future projects include to respond to sudden changes in shift schedule, 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. R. S. Sutton and A. G. Barto, Reinforcement Learning, A Bradford Book, MIT Press, Cambridge, 1998.
4. 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.
5. 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.

Developing Machine Learning and Deep Learning Models for Customer Churn Prediction in Telecommunication Industry*

Teoh Jay Shen†, Abdul Samad Bin Shibghatullah
Institute of Computer Science & Digital Innovation, UCSI University
Cheras, Kuala Lumpur, Malaysia‡

E-mail: 1001539989@ucsiuniversity.edu.my, abdulsamad@ucsiuniversity.edu.my

Abstract

Customer churn is always a significant problem and one of the biggest concerns of telecommunication companies. The companies are attempting to create and design an approach to predict customer churn. This is why determining factors that causes the customer to churn is significant. The proposed models constructed in this work apply both the machine learning and deep learning algorithms. Those models was constructed and run under the Python environment and it used an open sources dataset that are available to everyone on www.kaggle.com. This dataset contained 7043 rows of customer's data with 21 features, and it was applied in the training and testing process of the models development. These models used four different types of machine learning and deep learning algorithms, which are the Artificial Neural Network, Self-Organizing Map, Decision Tree and a hybrid model with the combination of the Self-Organizing Map and Artificial Neural network algorithms.

Keywords: Machine Learning; Deep Learning; Churn Prediction; Telco Industry.

1. Introduction

In Telco industries or other business related field, customer churn is defined as the action of existing customers terminating their subscription of service with the company due to several reasons such as dissatisfaction of the service provided/better price offered by competitors for the same services [1] . Customers churn prediction and management is specifically intense in the mobile telecommunication industry as the market is increasingly saturated by the customers who change their subscription from one network provider to another in a very frequent manner. Furthermore, one of the main reason that customers churn are recognized as the biggest problem in any industry is due to the cost of obtaining new customers is whole lot more than keeping the existing customers [2]. Based on a research conducted by Abbas Keramati and Seyed M.S. Ardabili, it stated that the annual churn rate within telecommunication can be range from 20% to 40%, whereas the cost needed to obtain a single new customer is 5 to 10 times higher than maintaining an existing customer [3]. Therefore, to develop a churn prediction model with high accuracy is very important for a Telco service provider to be successful or bottom-line survival of a Telco service provider in this intensely saturated marketplace. On those grounds, the Telco market is ever changing and getting more and more competitive [4]. In fact, customer churn is not only the essential concern on the marketing side, but also one of the fundamental dimensions of CRM which is also known as customer relationship management [6].Churn prediction model is the model that utilized data mining techniques to transform an enormous amount of data into a meaningful insight and present it in a way that normal people could also understand.

Machine Learning carries a huge potential in the field of data mining and data analytics [6]. Machine Learning can be defined as the ability of a computer to study and discover a set of rules from the given input data or overall build a model that can be used to find correlations between the variables or a structured data set to make predictions [7]. Numerous of machine learning predictive modelling algorithms have been proposed and applied in building churn prediction models, these machine learning algorithms can be ranged from simple linear regression to more complex hybrid methodologies. These algorithms are able to work efficiently in predicting the customers who are most likely to churn. At the past few years, besides machine learning, deep learning has also become one of the most popular trends thanks to its potential in processing big data [8]. Deep Learning, also known as deep structured learning is basically a division of Machine learning that relying on algorithms that trying to model high degree abstractions on data [9]. Deep Learning Algorithms (DLAs) establish a multi layered hierarchical structure of learning and indicating data. DLAs are useful in dealing with a massive amount of unsupervised data as it study and discover the correlations of data in a greedy layer structured method [9]. Thus, this project is going to develop two deep learning churn prediction models, one hybrid churn prediction model (SOM+ANN) and one machine learning churn prediction model using artificial neural network (ANN), self-organizing map (SOM), and decision tree (DT) algorithms. Lastly, CRSIP-DM methodology was applied for the development of this research.

2. Previous Work

Customer churn is one of the significant matters in the Telco industry nowadays. Numerous methods were implemented to predict customer churn in Telco companies. Majority of these methods employed machine learning, deep learning and data mining techniques. Most of the previous work done by researchers focused on implementing only one technique of data mining to retrieve knowledge, where others focused on conducting studies between several ML/DL algorithms to build a churn model. The aims and purposes of the research of such studies differ. Some of the aims are to simply evaluate if machine learning is an applicable approach for customer churn. In others, the aims may be to evaluate and select the best algorithms for prediction.

A comparative study for different type of machine learning algorithms for predicting churn customer for prepaid services was done by Brandusoiu and Toderean in the year 2016 [10]. The machine learning algorithms used in this study includes the Neural Network, Support Vector Machine and Bayes Network. The dataset used in this study consist of 3333 rows of customer call details with 21 independent variables and 1 dependent variable which is the churn status with the values of only YES or NO. Some features also captured the data of incoming and outgoing calls and messages for each customer. In order to reduce the data dimension, the principal component analysis algorithms were used by the author. AUC which is known as the area under the curve is the technique used by the author to evaluate the performance of the ML algorithms. The AUC values were reported 99.10%, 99.70% and 99.55% respectively for Bayes Network, support vector machine and neural network [10]. A similar study was made by Yue He, Zhenglin He and Dan Zhang in 2009 [11], where they designed a new model for churn prediction on the basis of the Neural Network algorithm. This model was proposed with the intention to solve the customer churn issues in a big Chinese Telco company which consists 5 million of customers. The results of the model accuracy rate reached 91.1%, which is very a stunning result [11].

In the year 2015, 9 researchers from China conducted a research of the customer churn problem in the big data platform [12].The goal of this study was to demonstrate that big data can significantly improve the performance of predicting customer churn based on the 3Vs (variety, volume, velocity) of the data. Interpreting the data at China's largest Telco Company, a big data platform is needed to engineer the process and solve the problems.

AUC (area under the curve) was once again used to measure the model performance and Random Forest Algorithm was applied to build the model in this study (12). Other than that, at the same year 2015, a group of researchers from the United States with Maryam M Najafabadi [13] as the leader demonstrated how to use deep learning to interpret big data, extracting and discovering complex correlation hidden under a large dataset. Moreover, they classified and list out certain types of problems that can be solve by applying deep learning. These problems are problems related to prediction and classification, sematic indexing, extracting information from raw dataset and data tagging. They discussed several components of deep learning to address certain issues in big data analytics [13].

Furthermore, Niall McLaughlin and his team used the concept of a CNN, which is also known as convolutional neural network to develop an android malicious software detection system in the year 2017. It was a creative and modern application of deep learning in the domain of malware analysis. This application was proficient in learning to perform malware detection and feature extraction at the same time. Lastly the team came to a conclusion that showed the proposed model is more efficient than an n-gram based malicious software detection system. The last observation of study related to deep learning application in predicting churn was done by Federico Castanedo and his team in 2016. They introduced deep neural network as a perfect tool for customer churn prediction as it constructs numerous hidden layer and it disseminates the weights of each neurons from one layer to the next layer. Deep Learning supports automated process of extracting features with the maximum information (impact), therefore there was a remarkable melioration in the performance of the model in accuracy wise. Billions of call details from a Telco business enterprise were used as the data input to train the predictive model for customer churn prediction. 77.9% of AUC (are under the curve) values were achieved on the validation dataset. Moreover, they extended the idea into the field of fraud detection in various industries such as the banking industry, insurance industry and Telco industry.

3. Methodology and Algorithm

For this research, CRISP-DM Methodology is use to model the customer churn prediction in Telco industry. CRISP-DM contains of the following phases: Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation and Deployment. In general, these phases are implemented subsequently. However in this main stream, various iterative processes can be seen due to the fact that every outcome of each phase impacts the next methodology phase. In other words, after the data understanding phase is completed, data analyst usually needs to go back to the business understanding phase to refine or reconsider the aims and objectives for the project. In the same manner, after the modelling part, mostly there is always a need for the data analyst to conduct a new data pre-processing process in order to enhance the model's performance. Moreover, the results of the evaluation phase could also potentially leads to a new start of the CRISP-DM process in the case that the models do not satisfy the aims and objectives defined previously in the research.

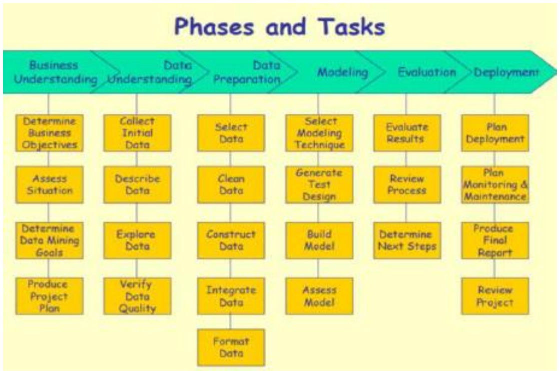


Fig.1: Phases and tasks of the CRISP-DM process

Fig. 1 shows the complete cycle of the CRISP-DM methodology and the phases within the cycle. The details are explained in the following section:

1. Business Understanding: This is the initial phase of CRISP-DM methodology. Three statements are produced in this phase which the statements include the statement of business objective, statement of data mining objective and statement of success criteria. This phase emphasizes on understanding the objectives and requirements of the project from a business point of view, then transforming this information into a problem

statement of the project and the draft of strategy to accomplish the objectives.
2. Data Understanding: The data understanding phase begin with the initial data collection and then follows by the data reading process with the purpose of getting familiarize with the data, identify the data quality, to determine the initial insights to the data and to discover relevance subset to create the hypothesis for hidden knowledge.
3. Data Preparation: This phase includes all the actions to develop the final dataset on the basis of the raw data. The tasks in data preparation phase are commonly been executed several times and does not follow any prescribed order. The tasks include row and feature selection, data filtering, data cleaning and transforming the dataset into the appropriate format for the modelling tools.
4. Modeling: The application of the algorithms and code developing parts are constructed under this phase. Model optimizing process is conducted in order to get the best performance form the model. Usually for one data mining problem, there are multiple algorithms and techniques available to be apply and solve the problems. However, some algorithms require specific requirements from the data. Thus, moving back to the data preparation part is often required.
5. Evaluation: In this phase, a thorough evaluation on the model and the steps of execution to construct the model is conducted in order to make sure that the business objectives are achieved by the model. The main activity of this phase is to assess and identify whether is there any critical business problems that has not been recognized adequately. In the end of this phase, a conclusion on the usage of the model results should be made.
6. Deployment. The deployment phase remains as the last phase in the CRISP-DM process. Where in this phase the data mining results are utilized as the business rule and the findings and knowledge obtained from the model needs to be visualized and presented in a way that business people can understand it.

Artificial Neural Network (ANN) is part of the model of this research. ANN was invented on the basis of a sophisticated biology research regarding neural system and human brain tissue. ANN is applied to stimulate the neural operations of knowledge processing in the human brain [14]. In ANN, the neurons which can also be called as the information processing nodes are formed in a topological structures. Therefore, the neurons disseminate their data and information in a parallel fashion. Multiple nonlinear transfer functions are combined to maps the inputs and measured output responses [15].

Self-Organizing Map (SOM) is classified as a methodology under unsupervised learning, it distribute a group of patterns into clusters or segments. Cluster analysis can be defined as the process of arranging a group of data object into different cluster. Above all, no predefined clusters are allocated. In the year 1987, Teuvo Kalevi Kohonen introduced and illustrated a brand new structure of a neural network architecture named the Self-Organizing Map (SOM). SOM turned out to become incredibly useful in the case that the input dataset are complicated and high dimensions. SOM is applied to study the correlations in a dataset and distribute the data into different cluster based on the similarity of patterns of the data where the modellers are not able to forecast the class of the classification[16]. SOM is classified as a methodology under unsupervised learning, it distribute Decision trees algorithm possess the basis of a graph structure, every decision could potentially generates a new node and eventually developed into a tree-like graph [18].

As a rule, a hybrid model is a combination of two or more machine learning or deep learning algorithms. As an example, the clustering (SOM) and classification (ANN) techniques can be combined in sequence. In other words, clustering techniques are able to be applied as a pre-processing phase to determine different pattern of clusters for later supervised learning [19]. Therefore, the result of the clustering process can either be used to determining the main clusters of a set of data given or be used as a pre-classification of unnamed features. Thus, the result of the clustering process can be included as a feature into the training set to train a prediction model. After the completion of the hybrid model, it has the capability to group or forecast new instances. In other terms, the first stage of the hybrid model is to detect the outlier of the dataset and the second stage of the hybrid model is to make prediction. Since the customer churn prediction is a supervised classification process, thus the hybrid models that are chosen in this paper are the cluster (SOM) to classifier (ANN) model.

3.1. *Data*

Only one dataset was used for this research. In order to obtain a good dataset, an online research are done by the researcher in order to find a free open source dataset that fulfilled the requirements of this research, finally a dataset that fulfils the requirements was found on www.kaggle.com. A total of 7043 rows of customer details with 21 features were provided in the dataset. The variables of the dataset used were stimulated from a Telco company from USA. Variables of the dataset are categorized into two types. Nominal types define a set of values which represents a certain meaning, while quantitative types refer to values that can be calculated and ordered, such as integer and float values. Variable "Churn" is the dependant variable, meaning it is the variable that the model is trying to predict (see Table 1, row 21).

Table 1: Description of the Data

Index	Column	Description	Type
1	Customer ID	Customer ID	Nominal
2	Gender	Gender of the customer (Male/Female)	Nominal
3	SeniorCitizen	Determine the customer is a senior or not (1/0)	Nominal
4	Partner	Determine the customer has a partner or not (Yes/No)	Nominal
5	Dependents	Determine the customer has a dependents or not (Yes/No)	Nominal
6	Tenure	Amount of months that the customer used the company service.	Quantitative
7	PhoneService	Determine the customer has phone service or not (Yes/No)	Nominal
8	MultipleLines	Determine the customer has multiple line or not (Yes/No)	Nominal
9	InternetService	The ISP, Internet Service Provider of the customer (Fiber optic, DSL, No)	Nominal
10	OnlineSecurity	Determine the customer has online (Yes/No/No Internet Service)	Nominal
11	OnlineBackup	Determine the customer has online backup or not (Yes/No/No Internet Service)	Nominal
12	DeviceProtection	Determine the customer protection for the device or not (Yes/No/No Internet Service)	Nominal
13	TechSupport	Determine the customer has technical support or not (Yes/No/No Internet Service)	Nominal
14	StreamingTV	Determine the Customer has streaming TV or not (Yes/No)	Nominal
15	StreamingMovies	Determine the customer has streaming movies or not (Yes/No/No internet service)	Nominal
16	Contract	The contract term of the customer (Month-to-month, One year, Two year)	Nominal
17	PaperlessBilling	Determine the customer has paperless billing or not (Yes/No)	Nominal
18	PaymentMethod	The payment method of the customer (Electronic check, Mailed check, Bank transfer, Credit card)	Nominal
19	MonthlyCharges	The monthly amount charged to the customer	Quantitative
20	TotalCharges	The total amount charged to the customer	Quantitative
21	Churn	Determine the customer churned or not	Nominal

4. Results and Evaluation

The method of evaluation that will be used to evaluate the algorithm's performance in this research is the calculation of the F-measure value. The accuracy, precision, recall and the F-measures value for all of the four proposed models are calculated accurately and populated in Table 2 with its corresponding algorithm.

Table 2 Models performance comparison table.

Algorithm	Accuracy	Precision	Recall	F-measure
Artificial Neural Network (ANN)	78.90	84.03	88.24	86.08
Self-Organizing Map (SOM)	51.89	69.35	70.67	70.00
Decision Trees (DT)	73.41	85.59	80.73	83.08
Hybrid Model (SOM + ANN)	79.53	83.90	89.40	86.56

In Table 2 the performance of each model can be observed clearly. Based on the F-measure score, the hybrid model possesses an 86.56% of F-measure score which is also the highest among the proposed models, and then followed by the Artificial Neural Network model which holds almost the same F-measure score (86.08%) with the hybrid model. In addition, the Decision Tree model holds a 83.08% of F-measure score and the Self-Organizing Map model possesses the lowest F-measure score (70.00%) among the proposed models. As the Self-Organizing Map model applied the unsupervised learning method, thus it is reasonable for the SOM model to have a comparatively weaker performance compared other models that applied supervised learning method in this case. Lastly, by evaluating the results of the models we noticed that the overall performance of the model increased after implementing the Hybrid model approach.

5. Conclusion and Future Work

CRISP-DM methodology was implemented for conducting a prediction of churn customer in the Telco industry. The aim of this research is to compare the algorithms and to determine which is the most suitable and accurate to predict the results of the given problem. A uniquely modified dataset was analysed with different machine learning and deep learning models and its result were compared with an evaluation method. The algorithms used were the Artificial Neural Network

(ANN), Self-Organizing Map, Decision Tree (DT) and a hybrid algorithm which is the combination of SOM and ANN algorithm. After studying and comparing the results of each algorithms, the hybrid model with two deep learning algorithms combined together are concluded as the model that achieved the best outcomes for this research.

Moreover, the likelihood of successfully predicting the customer churn with Machine Learning and Deep Learning depends on the correct choice of data and algorithm. In order to achieve the best outcomes, choosing a suitable machine learning or deep learning algorithm for the problem or topic is crucial. However, by having the algorithm alone is unable to provide the best prediction results. Therefore, having the right variables and types of dataset is also an important factor in getting the best prediction results. Also, by looking at the scores of the feature importance is able to determine how the changes of each independent variable impact on another dependent variable. In the optimization process of the dataset, only highly statistically significance variables/predictors will be selected and grouped in an optimal dataset for independent variable. In fact, the optimal dataset does help the models to come up with better predictions. In addition, based on our literature review, implementing machine learning and deep learning model to predict customer churn is new and prevailing in the Global Telco industry. Feature engineering and data gathering process for predicting churn in the Telco industry in one of the popular and important elements of the current research Thus, it is expected that with an in-depth understanding of characteristic and behaviour of the churn customers, Telco companies can develop new strategies to address churn. These strategies need to be applied on the targeted customers who shared the similar behaviours with the churn customers group. Offering a new plan, provides better quality of service, determine the needs from different segments of customer, and provides tailored offers for different customer groups can be all included in the strategies.

There are some constraints in this research that needs to be recognized and improvement needs to be done against them. Firstly, because of the privacy issues of the Telco companies, real time dataset are not allowed to be used in this research. This research focused on study and compares the performance of different models, thus the future work will focus on the feature selection and the research of the customer needs by applying different data mining techniques. Moreover, future research can be improved by implement others machine learning and deep learning methods such as: K-Nearest Neighbours, Logistic regression, Learning Vector Quantization, Support Vector Machines and Random Forest. They can be used to have a better understanding of each machine learning algorithm results provides a statistical results to aids the algorithm selection process.

Acknowledgements

The authors wish to thank ICSDI, and CERVIE UCSI University for supporting this project.

References

1. Mishra, A., & Reddy, U. S. (2017). A Novel Approach for Churn Prediction Using Deep Learning. 2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC).
2. Wei, C., & Chiu, I. (2002). Turning telecommunications call details to churn prediction: A data mining approach. Expert Systems with Applications, 23(2), 103-112.
3. Keramati, Abbas & M.S. Ardabili, Seyed. (2011). Churn analysis for an Iranian mobile operator. Telecommunications Policy. 35. 344-356. 10.1016/j.telpol.2011.02.009.
4. Mozer, M., Wolniewicz, R., Grimes, D., Johnson, E., & Kaushansky, H. (2000). Predicting subscriber dissatisfaction and improving retention in the wireless telecommunications industry. IEEE Transactions on Neural Networks, 11(3), 690-696.
5. Ngai, E., Xiu, L., & Chau, D. (2009). Application of data mining techniques in customer relationship management: A literature review and classification Expert Systems with Applications, 36(2), 2592-2602.
6. Rathor, A., & Gyanchandani, M. (2017). A review at Machine learning algorithms targeting big data challenges. 2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT).
7. Kotsiantis. (2007). Supervised Machine Learning: A Review of classification Techniques, Informatica. Informatica,31, 249-268.
8. Sejnowski, T. J. (2018). The deep learning revolution. Cambridge, MA: The MIT Press
9. Hordri, Nur & Yuhaniz, Siti & Shamsuddin, Siti Mariyam. (2016). Deep Learning and Its Applications: A Review.

10. Brandusoiu, Ionut & Toderean, G & Beleiu, Horia. (2016). Methods for Churn Prediction in the Pre-paid Mobile Telecommunications Industry. 97-100. 10.1109/ICComm.2016.7528311.
11. He, Y., He, Z., & Zhang, D. (2009). A Study on Prediction of Customer Churn in Fixed Communication Network Based on Data Mining. 2009 Sixth International Conference on Fuzzy Systems and Knowledge Discovery.
12. Huang, Y., Zhu, F., Yuan, M., Deng, K., Li, Y., Ni, B.,Zeng, J. (2015). Telco Churn Prediction with Big Data. Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data - SIGMOD 15.
13. Najafabadi, M. M., Villanustre, F., Khoshgoftaar, T. M., Seliya, N., Wald, R., & Muharemagic, E. (2015). Deep learning applications and challenges in big data analytics. Journal of Big Data, 2(1).
14. Markopoulos, A. P., Manolakos, D. E., & Vaxevanidis, N. M. (2008). Artificial neural network models for the prediction of surface roughness in electrical discharge machining. Journal of Intelligent Manufacturing, 19(3), 283-292.
15. Sharma, A., & Panigrahi, P. K. (2011). A Neural Network based Approach for Predicting Customer Churn in Cellular Network Services. International Journal of Computer Applications, 27(11), 26-31.
16. Kohonen, T. (1987). Adaptive, associative, and self-organizing functions in neural computing. Applied Optics, 26(23), 4910.
17. Tsai, C., & Lu, Y. (2009). Customer churn prediction by hybrid neural networks. Expert Systems with Applications,36(10), 12547-12553.
18. Quinlan, J. (1987). Simplifying decision trees. International Journal of Man-Machine Studies, 27(3), 221-234. official website, https://www.robocup.org/

Authors Introduction

Mr Teoh Jay Shen

He received his Diploma of Information Technology from UCSI University in 2018 and continue his Bachelor of Science in Computing at the same university and completed in 2020. His research interest are in machine learning and artificial intelligent.

Dr. Abdul Samad Bin Shibghatullah

He received the bachelor accounting degree from Universiti Kebangsaan Malaysia, Bangi, Malaysia in 1999, the M.Sc. degree in computer science from the Universiti Teknologi Malaysia, Skudai, Malaysia in 2002, and the Ph.D. degree in computer science from the Brunel University of Uxbridge, United Kingdom. He is currently Associate Professor at Institute of Computer Science & Digital Innovation, UCSI University, Kuala Lumpur, Malaysia. His current research interests include optimization, modelling and scheduling.

Liver Segmentation in CT Images Using Residual U-Net and 3D CRF

Shuntaro Nagano, Tohru Kamiya
Department of Mechanical and Control Engineering, Kyushu Institute of Technology, 1-1 Sensui-cho, Tobata-ku, Kitakyushu-shi, Fukuoka, Japan

Guanguxu Li
Tiangong Univercity, 399 Binshuixi Road, Xiqing District, Tianjin, China
E-mail: nagano.shuntaro244@mail.kyutech.jp

Abstract

In recent years, the development of CAD systems aimed at reducing the burden on doctors and improving diagnostic accuracy has been promoted. In this paper, we propose a segmentation method for the liver site on abdominal CT images as a pretreatment for a CAD system using dynamic CT images. The method consists of two stages. First, we segment the liver with a model based on U-net, a segmentation model using CNN. Next, the 3D CRF (Conditional Random Field) is used to make corrections that take into account the three-dimensional characteristics of the liver to improve the accuracy of segmentation. In the experiment, the accuracy was evaluated for CT images of 20 cases.

Keywords: Liver, Convolutional Neural Network, Segmentation, Conditional Random Field.

1. Introduction

According to GLOBOCAN2020, the International Agency for Research on Cancer (IARC), an external research organization of the World Health Organization (WHO), the number of people suffering from liver cancer worldwide is about 910,000. In humans [1], it ranks 7th among cancer types (4.7% of the total). In addition, the total number of deaths due to liver cancer is about 830,000, which is the second highest among cancer types (8.3% of the total), and the mortality rate for the number of affected people is as high as 92%. Therefore, early detection and appropriate treatment are important.

Imaging diagnosis by CT, MR, etc. plays an important role in early detection of liver disease. In particular, diagnostic radiographic imaging such as dynamic contrast enhanced CT provides useful information for the differential diagnosis of the liver lesions [2]. However, since image diagnosis is performed by subjective judgment based on many years of experience of doctors, there is a problem that the diagnosis results by doctors vary. In addition, since dynamic CT takes multiple images, the number of images taken increases, which raises the burden on doctors. For this reason, research is being conducted on Computer Aided Diagnosis (CAD) systems aimed at reducing the burden on doctors and improving diagnostic accuracy.

In general, the CAD system require many technical issues such as extraction of region of interest (ROI), image alignment, identification of abnormal shadows, and diagnosis. In this paper, we focus on the extraction of ROI from a CT images.

The accuracy of area extraction in medical images is important because it affects the accuracy of alignment and identification performed later. In the extraction of the liver region, there are level set method [3] and active shape model method [4]. However, in these extraction methods, omission of extraction of lesions has become a problem. On the other hand, some Convolutional Neural Network (CNN) approaches [5, 6] are introduced for the extraction of ROI. Therefore, in this paper we use the CNN, which has produced good results in the field of image segmentation, to extract the liver region including lesions on CT images. In the experiment, we performed our proposed method to 20 CT cases and evaluated our performance.

2. Methods

The flow of the proposed method consists of two stages: segmentation using CNN and correction by 3D Conditional Random Field (3DCRF) [7].

2.1. *Segmentation with CNN*

In this paper, we propose a network structure based on Residual U-Net [8], which is a segmentation model by CNN, Residual Block [9] is changed to ResNet-D [10], and added Spatial and Channel Squeeze and Excitation (scSE) [11] to the end of each block of encoder and decoder.

2.1.1. *Residual U-Net*

Residual U-Net is a CNN model used in the field of segmentation that combines U-Net [12] and residual block. Similar to U-Net, the encoder extracts the image features and the decoder restores them. In addition, by connecting a shortcut from the encoder to the decoder, it is possible to restore with the decoder while retaining the local information and position information of the image.

The difference from U-Net is the introduction of residual block in the encoder part and decoder part. With the introduction of residual block, feature extraction becomes possible while suppressing deterioration of image information.

Various structures have been proposed for the residual block, however the proposed method uses residual U-Net 15 with full pre-activation [13].

2.1.2. *ResNet-D*

ResNet-D is a model in which an average pooling layer is added to the shortcut connection of ResNet's residual block. ResNet's residual block causes a 3/4 loss of information in shortcut connection when down sampling. Therefore, by adding an average pooling layer to the shortcut connection, the omission of this information is prevented.

2.1.3. *Spatial and Channel Squeeze and Excitation (scSE)*

Spatial and channel squeeze and excitation is a method proposed to improve the performance of CNN models that perform segmentation. This is an extension of the Squeeze-and-Excitation block (cSE) [14] used in the field of image classification. In cSE block, the entire image is squeeze and the excitation for each channel is calculated, but in segmentation, the information between pixels cannot be enhanced well. Therefore, channel squeeze and spatial excitation was proposed. Contrary to cSE block, sSE block is squeeze in the channel direction and excitation is calculated for each pixel, so that the output takes into account the relationship between pixels. scSE is a combination of both sSE and cSE. The details of scSE are shown in Fig. 1.

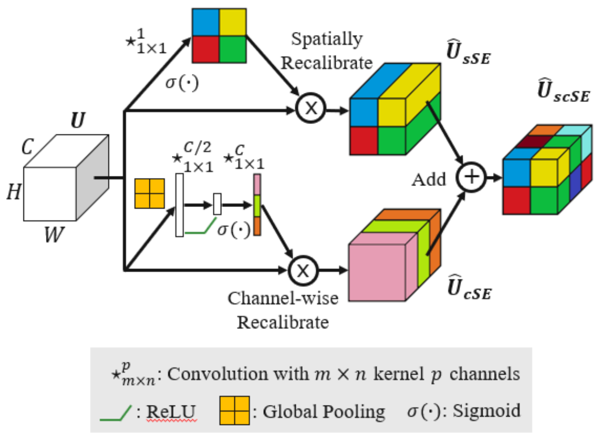


Fig. 1. Outline of scSE Block

2.2. *3D Conditional random Field (3DCRF)*

The CT image is 3D information, but the segmentation model of the proposed method is performed for 2D. In other words, the liver region is extracted for each slice of the CT image. This is because there is a problem due to the number of data sets and GPU specifications for learning in 3D. Therefore, in order to make corrections that take into account the three-dimensional characteristics of the liver, 3DCRF is used in this paper.

The CRF is considered on a complete graph $G = (V, E)$ with the vertices $i \in V$ of each pixel in the image and the edges $e_{ij} \in E = \{(i,j) \forall i,j \in V \text{s.t.} i < j\}$ between the vertices. In addition, the variable vector $x \in L^N$ is used as the label for each vertex, and the labeling is optimized for each vertex by minimizing the following energy function.

$$E(x) = \sum_{i \in V} \phi_i(x_i) + \sum_{(i,j)\in E} \phi_{ij}(x_i, x_j) \quad (1)$$

$$\phi_i(x_i) = -logP(x_i|I) \quad (2)$$

$$\phi_{ij}(x_i, x_j) = \mu\left(w_{pos}\exp\left(-\frac{|p_i - p_j|^2}{2\sigma_{pos}^2}\right)\right.$$
$$\left. + w_{bil}exp\left(-\frac{|p_i - p_j|^2}{2\sigma_{bil}^2} - \frac{|I_i - I_j|^2}{2\sigma_{int}^2}\right)\right) \quad (3)$$

Here, $P(x_i|I)$ is the likelihood of each pixel to the class obtained from the segmentation model of the proposed method. In addition, $\mu(x_i, x_j) = 1(x_i \neq x_j)$, and $\phi_{ij}(x_i, x_j)$ represents the strength of the bond between pixels, and represents the cost considering the color information and distance information of the image. Also, $|p_i - p_j|$ is the spatial distance between pixels i and j, and $|I_i - I_j|$ is the difference in the color information of the image. Each element in this equation can be adjusted by the weights w_{pos}, w_{bil} and the variances σ_{pos}, σ_{bil}, σ_{int}.

In this paper, the following optimization problem is considered with the segmentation output as the optimum labeling for each pixel.

$$x^* = \underset{x \in L^N}{\operatorname{argmin}}\, E(x) \quad (4)$$

By performing the above optimization, the accuracy of segmentation is improved.

3. Experimental Result

As the data set for the experiment, 20 cases of 3DIRCADb data set [15], 2542 sheets were used.

Intersection over Union (IoU) is used as the evaluation method. The formula for IoU is shown below.

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (5)$$

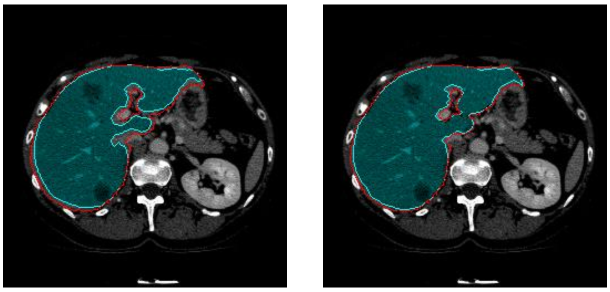
In order to verify the effectiveness of the proposed method, 5-fold cross-validation was performed. The experimental results are shown in Table 1, Fig. 2 and 3.

Table 1. Experimental result

Approach	IoU
U-Net + 3DCRF [6]	0.83
Proposed method Residual U-Net + ResNet-D + seSE +3DCRF	0.87

4. Discussion

From Table 1, the proposed method has higher accuracy than U-Net + 3DCRF.



(a) U-Net + 3DCRF (b) Proposed method

Fig. 2. Comparison of extraction accuracy of liver part

(Blue: Extracted result region, Red: Boundary of ground truth)

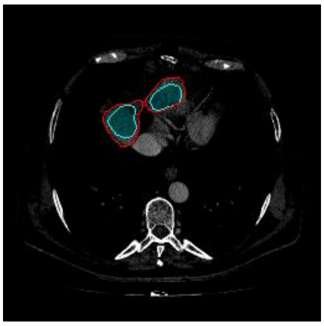


Fig. 3. Extraction result of liver terminal

From Fig. 2, it can be seen that the proposed method can extract the vicinity of the boundary with other organs, which was difficult in the past. We believe that the retention of information by the residual block and the characterization of the spatial position by the scSE block are particularly effective.

Figure 3 shows the extraction results of the terminal part of the liver, but shown in the figure, there are still some parts that have not been successfully extracted. This is because the shape of the liver often varies greatly from person to person, so it is considered that they have not been fully learned. The solution to this is the augment of data.

□

5. Conclusion

In this paper, we proposed segmentation of the liver region using deep learning and 3DCRF. By improving U-Net, which is a segmentation model, the experimental results showed that IoU was 0.87, which exceeded the conventional method of 0.83. Future tasks are further improvement of accuracy and introduction of alignment method and cancer detection method.

References

1. Hyuna Sung et al., "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries", A Cancer Journal for Clinicians, pp.209-249, 2021.
2. Jiarong Zhou et al., "Automatic detection and classification of focal liver lesions based on deep convolutional neural networks: A preliminary study", Frontiers in Oncology, pp.1-11, 2021.
3. Kentaro Saito et al., "Automatic liver segmentation from multiphase CT images by using level set method", International Conference on Control Automation and Systems, pp.1590-1592, 2017.
4. Esfandiarkhani Mina et al., "A generalized active shape model for segmentation of liver in low-contrast CT volumes", Computers in Biology and Medicine, 82, pp.59-70, 2017.
5. Christ Patrick Ferdinand et al., "Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields", International Conference on Medical Image Computing and Computer-Assisted Intervention, pp.415-423, 2016.
6. Kentaro Saito et al., "ROI-based Fully Automated Liver Registration in Multi-phase CT Images", International Conference on Control, Automation and Systems, pp.645-649, 2018.
7. Krähenbühl Philipp et al., "Efficient inference in fully connected crfs with Gaussian edge potentials", NIPS, pp.109-117, 2011.
8. Zhang et al., "Road Extraction by Deep Residual U-Net", IEEE Geoscience and Remote Sensing Letters, Vol.15, Issue.5, pp.749-753, 2018.
9. He et al., "Deep Residual Learning for Image Recognition", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp.770-778. 2016.
10. T. He, "Bag of Tricks for Image Classification with Convolutional Neural Networks", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp.558-567, 2019.
11. Abhijit Guha Roy et al., "Concurrent Spatial and Channel Squeeze & Excitation in Fully Convolutional Networks", arXiv preprint arXiv:1803.02579, 2018.
12. Olaf Ronneberger et al., "U-net: Convolutional networks for biomedical image segmentation." In: MICCAI, vol. 9351, pp. 234–241, 2015.
13. He et al., "Identify Mapping in Deep Residual Networks", European Conference on Computer Vision, pp.630-645, 2016.
14. Jie Hu et al., "Squeeze-and-Excitation Networks", arXiv preprint arXiv: 1709.01507, 2018.
15. Soler L et al., "3d image reconstruction for comparison of algorithm database: a patient-specific anatomical and medical image database", 2012.

Authors Introduction

Mr. Shuntaro Nagano

He graduated from the Department of Mechanical and Control Engineering, Faculty of Engineering, Kyushu Institute of Technology, Japan in 2020. 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.

Dr. Guanguxu Li

Guangxu Li received his B.E. and M.E. of automation engineering from Liaoning Institute of Technology in 2006 and 2009 separately. And Ph.D. degree from Kyushu Institute of Technology in 2013. He joined Tiangong University after graduated. His research interests focus on the geometric methods applied to medical image.

Judgement on Shunt Sounds from Vascular Access using YOLO Deep Learning Model

Kyosuke Fujiwara
Faculty of Engineering, University of Miyazaki,
1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

Takayuki Yamamoto
Graduate School of Engineering, University of Miyazaki,
1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

Lindsey Tate
Faculty of Engineering, University of Miyazaki,
1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

Kazuya Kibune
Department of clinical engineering section, Tokatsu dialysis hospital & clinic
865-2 Hinokuchi, Matsudo-shi, Chiba, 271-0067, Japan

Hiroki Tamura
Faculty of Engineering, University of Miyazaki,
1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

E-mail: hi18041@student.miyazaki-u.ac.jp, hi16046@student.miyazaki-u.ac.jp, teitorinzeirini.c4@cc.miyazaki-u.ac.jp, kmtssk@yahoo.co.jp, htamura@cc.miyazaki-u.ac.jp

Abstract

In order to explore more cost effective and accessible options for screening AV fistulas used as accesses for hemodialysis, we sought to link the results of echocardiograms to audio recorded from their respective shunts. We used machine learning techniques aimed at correctly classifying the audio data into the three classification categories derived from echocardiogram results. In this paper, we compare the discrimination rate of YOLOv2tiny alone to that of the combination of YOLOv2tiny and multivariate analysis. We conclude that bump wavelet analysis and linear multivariate analysis are the most suitable for audio data from hemodialysis shunts.

Keywords: Shunt sounds, Wavelet transform, YOLOv2tiny, Deep learning, Support vector machine

1. Introduction

An arteriovenous (AV) fistula, hereafter referred to as a shunt, is a blood vessel that connects a vein directly to an artery in order to allow a large volume of blood to pass for hemodialysis. The sound of a shunt is the sound made at its junction. Shunts carry a high risk of stenosis or blood clots, which are normally screened for using an echocardiogram (aka, "echo"). However, performing an echo requires oversight by a physician and the use of expensive equipment, which may not always be accessible to patients. In order to explore more cost effective and accessible options for screening, we sought to link the results of echocardiograms to audio recorded from their respective shunts. We used machine learning techniques aimed at correctly classifying the audio data into the classification categories derived from the echo results. Audio was recorded from shunts both before and after stress. We used wavelet transforms to image shunt sounds before and after stress, and we trained YOLOv2tiny to image the sounds. In this study, we found that YOLOv2tiny alone did not improve the discrimination rate. Therefore, we performed multivariate analysis using the multivariate calculated in the judgment of YOLOv2tiny. In this paper, we compare the discrimination rate of the result of YOLOv2tiny alone to the result of the combination of YOLOv2tiny and multivariate analysis, and we explain how to achieve a higher discrimination rate.

2. Proposed method

When examining a shunt, pre-stress measurements (such as peripheral vascular resistance, PI) and post-stress measurements are calculated by echography, and the values are comprehensively taken into account for diagnosis. I hypothesized that by linking the pre- and post-stress characteristics of the shunt sound with the diagnostic results, it would be possible to diagnose the shunt using only audio data. First, the audio data was converted into images using wavelet transform, and YOLOv2tiny was trained with the transform images. Next, we applied the trained YOLOv2tiny to all available data. The discrimination rates before and after stress were calculated for three echocardiogram classifications: normal, gray (uncertain), and abnormal. We used these six discrimination rates to compare analysis methods. In order to further improve discrimination rate, we used multiple regression analysis, decision tree, random forest, and SVM as learning methods.

2.1. *First analysis method*

Using only YOLOv2tiny, the discrimination rate is calculated using a model that is trained on pre-stress and post-stress together (Fig.1).

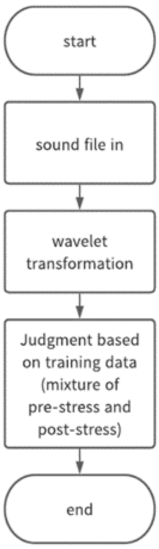


Fig.1 .Flowchart of the first analysis method

2.2. *Second analysis method*

Using multiple regression analysis, the discrimination rate is calculated using a model trained separately for pre-stress and post-stress (Fig.2).

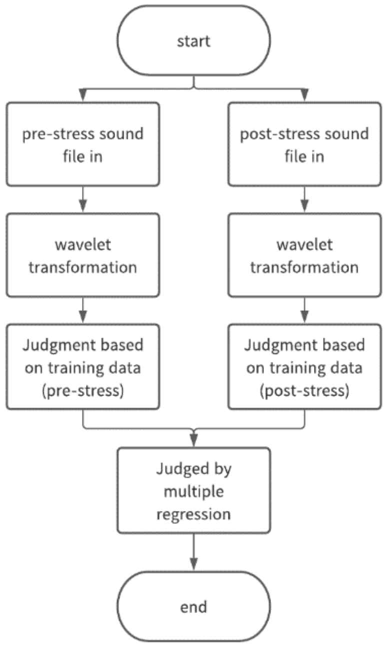


Fig.2 .Flowchart of the second analysis method

2.3. *Third analysis method*

From Fig.2, using separately pre-stress and post- stress, the judgment part has been changed. The discrimination rate was calculated by Decision tree, Random Forest, Linear SVM, and Nonlinear SVM.

3. Experiment

For the YOLOv2tiny model, the ratio of training data to test data was 7:3 for 100 subjects. The discrimination rate was calculated as the percentage of correct answers in the test data. The labels used were: good (i.e., normal; n = 45 subjects), need for follow-up (i.e., gray; n = 15 subjects), and need for treatment (i.e., abnormal; n = 40 subjects). For each pre- and post-stress, there were 100 trials (total 200 trials of audio data).

In the multiple regression analysis, all the data were trained and labeled as: normal = 1, gray = 0, and abnormal = (-1). After calculating the equation, the values were obtained by re-substitution. In the judgment, values lower than (-0.33) were considered abnormal, values between (-0.33) and 0.33 were considered gray, and values above 0.33 were considered normal.

Decision trees, random forests, and SVMs were all trained using MATLAB by MathWorks, and the discrimination rates were verified by single-tailed cross-validation. Morlet, morse, and bump wavelet transforms were tested.

There were 100 observations with 6 variables. Because of the small amount of audio data, the data was padded by clipping 5 seconds from the beginning of the audio and shifting it by 1 second. This resulted in 5,815 observations of image data. There were approximately 30 observations for each pre-stress audio and post-stress audio per person. Since the number of seconds of shunt sound varied, there was some variation in the padded image data, which made it difficult to label the same subject before and after stress. Therefore, we averaged the approximately 30 data mentioned earlier and adjusted the data so that there would be one observation of the three pre-stress variables and one observation of the three post-stress variables per person.

4. Result

The discrimination rate for YOLOv2tiny was calculated by assuming that the correct answer was the one where the judgment and label matched. For decision trees, random forests, and SVMs, the discrimination rate was calculated as the average of 100 cross-validation trials.

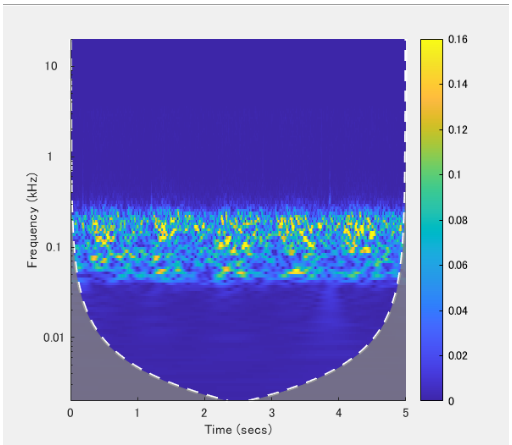


Fig.3 . Representative image of a normal subject after wavelet transform (bump)

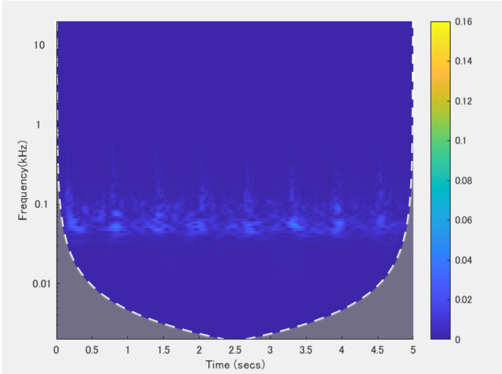


Fig.4 . Representative image of an abnormal subject after wavelet transform (bump)

From both wavelet transformations (Fig. 4 and 5), we observed periodic sounds continuously between 0.01 KHz and 1 KHz. Normal subjects characteristically had continuous waveforms with many high amplitude waveforms between 0.01 KHz and 1 KHz (Fig. 4). On the other hand, the abnormal subjects had only a few small amplitude waveforms and a series of high frequency sounds over a short time period (Fig. 5).

4.1. *Results of the proposed method 1*

The overall results for the proposed method 1 are 64% (morlet),57%(morse)and57%(bump). The discrimination rate of gray is low in all mother wavelets, and the discrimination rate of bump seems to be the most balanced.

Table 1. Discrimination rate with YOLOv2tiny only (morlet)

		Correct		
		normal	gray	abnormal
Method1	normal	84.6	14.4	1
	gray	57	9.3	33.7
	abnormal	22.5	16.2	61.3

4.2. *Results of the proposed method 2*

The overall results for the proposed method 2 are 89% (morlet),93%(morse)and97%(bump). Identification rate of bump was the highest at 97% (Table2). However, do not use cross-validation or other methods, so this result is not sufficiently generic.

Table 2. Discrimination rate when YOLOv2tiny is combined with multiple regression analysis (bump)

		Correct		
		normal	gray	abnormal
Method2	normal	97.8	2.2	0
	gray	0	100	0
	abnormal	0	2.5	97.5

4.3. *Results of the proposed method 3*

The overall results of the proposed method 3 are 88%(morlet), 89%(morse), 89%(bump) for Decision Tree, 87%(morlet), 90%(morse), 91%(bump) for Random Forest, 91%(morlet), 91%(morse), 95%(bump) for Linear SVM, and 90%(morlet), 91%(morse), 90%(bump) for Nonlinear SVM. The discrimination rate of morse and bump was the same 89% in the decision tree (Table3). In Random Forest, bump had the highest discrimination rate at 91% (Table4). In linear SVM, bump had the highest identification rate at 95% (Table5). In nonlinear SVM, morse had the highest identification rate at 91% (Table6). As a result, the results of the linear SVM were high, which indicates that this data is linear data. It can be seen that gray(uncertain) has a low discrimination rate in all results.

Table 3. Discrimination rate when YOLOv2tiny is combined with Decision Trees(bump)

Decision Tree		Correct		
		normal	gray	abnormal
Method3	normal	93.3	6.7	0
	gray	13.3	66.7	20
	abnormal	0	7.5	92.5

Table 4. Discrimination rate when YOLOv2tiny is combined with Random Forest (bump)

Random Forest		Correct		
		normal	gray	abnormal
Method3	normal	100	0.0	0
	gray	6.7	80	13.3
	abnormal	10	5	85

Table 5. Discrimination rate when YOLOv2tiny is combined with Linear Support Vector Machines (bump)

Linear SVM		Correct		
		normal	gray	abnormal
Method3	normal	97.8	2.2	0
	gray	6.7	80.0	13.3
	abnormal	0	2.5	97.5

Table 6. Discrimination rate when YOLOv2tiny is combined with Non-Linear Support Vector Machines (morse)

Nonlinear SVM		Correct		
		normal	gray	abnormal
Method3	normal	91.1	0.0	8.9
	gray	13.3	73.3	13.3
	abnormal	2.5	0.0	97.5

5. Conclusion

Rather than classify wavelet transform images solely with YOLOv2tiny, we improved the discrimination rate of shunt sounds by performing multivariate analysis on the variables derived from the YOLOv2tiny decision.

In the first analysis method, the highest result for the normal classification was 84.6% (Table 1), 63.2% for abnormal (Table 2), and 17% for gray (Table 3). Since YOLOv2tiny alone was not sufficient to improve the discrimination rate, we used multivariate analysis to improve the discrimination rate. In the multiple regression analysis (i.e., the second analysis method), the discrimination rates for bump were quite high (Table 1). However, this result did not allow us to verify the method’s versatility, so we used cross-validation in the third analysis method to verify its generalizability. The results were 89% for decision trees, 91% for random forests, 95% for linear SVM (Table 5), and 91% for nonlinear SVM (Table 6). Overall, the mother wavelet that was best suited for shunt sounds analysis was bump, and linear multivariate analysis was found to be suitable for the data.

Future research will investigate why the discrimination rates of the gray classification were consistently lower than those of abnormal and normal classifications for all methods. We suggest that one of the reasons for this is the small number of gray data; therefore, we will include more gray trials. Another possible explanation is the inherent ambiguity in the gray category, which is an echo undiagnosable as either normal or abnormal; this ambiguity may necessitate subdividing gray into two or more sub-categories. Additionally, the analysis method should be developed further specifically to improve the pre-stress trials with the same accuracy as post-stress trials. Ultimately, the goal of this line of research is to create an automated system for screening arteriovenous (AV) fistulas for stenosis and blood clots without the use of an echocardiogram.

Acknowledgments:

In this experiment, audio data was collected using a device under development provided by Togo Medikit Co., Ltd. We appreciate Togo Medikit Co., Ltd for their cooperation.

References

[1] Ikuo Toufukuzi "The Future of Telemedicine Journal of the Japan Society of Home Economics" 62(2):141-144 ,2011(Japanese)
[2] Mizuho Tagashira and Takahumi Nakagawa "Discrimination Method for Vascular Stenosis by Acoustic Analysis of Dialysis Shunt Sounds" Biomedical Engineering 59(1):31-39,2021(Japanese)
[3] Joseph Redmon, Ali Farhadi; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 7263-7271
[4] G. Li, Z. Song and Q. Fu, "A New Method of Image Detection for Small Datasets under the Framework of YOLO Network," 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), 2018, pp. 1031-1035,
[5] Sifuzzaman, M,slam, M R, Ali, M Z, 2009, Application of Wavelet Transform and its Advantages Compared to Fourier Transform. Journal of Physical Science;Vol 13 [2009] 121-134
[6] Ngui, W.K., Leong, M.S., Hee, L.M., Abdelrhman, A.M., 2013. Wavelet Analysis: Mother Wavelet Selection Methods. AMM 393, 953–958.
[7] S. R. Safavian and D. Landgrebe, "A survey of decision tree classifier methodology," in IEEE Transactions on Systems, Man, and Cybernetics, vol. 21, no. 3, pp. 660-674, May-June 1991
[8] Breiman, L. Random Forests. Machine Learning 45, 5–32 (2001).
[9] Haifeng Wang and Dejin Hu, "Comparison of SVM and LS-SVM for Regression," 2005 International Conference on Neural Networks and Brain, 2005, pp. 279-283,
[10] A. Mathur and G. M. Foody, "Multiclass and Binary SVM Classification: Implications for Training and Classification Users," in IEEE Geoscience and Remote Sensing Letters, vol. 5, no. 2, pp. 241-245, April 2008,

Authors Introduction

Mr. Kyousuke Fujiwara

Kyosuke Fujiwara born in 2000. He is currently studying in department of environmental robotics, and will receive the B.Eng from University of Miyazaki in 2022.
His current research is Analysis of Shunt Sound from Vascular Access in Machine Learning.

Mr. Takayuki Yamamoto

Takayuki Yamamoto born in 1998. He is currenttly studying in Graduate School of Engineering, and will receive the M.Eng from University of Miyazaki in 2022.
His current research is A Study on the Algorithm of Auscultation SoundAnalysis for Heart Sounds.

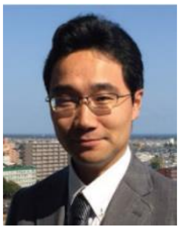
Dr. Lindsey Tate

She received a B.A. in Psychology (Neuroscience) from St. Edward's University in Austin, TX, USA in 2013. In 2016 and 2018, respectively, she received her M.S. and PhD in Cognitive Psychology from the University of Oklahoma, where she conducted perceptual and behavioral neuroscience research using EEG and MEG. Now she works as a researcher in the Faculty of Engineering at the University of Miyazaki. Her main research interests are applying machine learning and other algorithms to real-time non-invasive physiological data for use in clinical and non-clinical biofeedback systems.

Mr.Kazuya Kibune

He is currently a department of clinical engineering section manager in tokatsu dialysis hospital & clinics. He works as the clinical engineer and the clinical vascular technologist.

Prof. Hiroki Tamura



He received the B.E. and M.E. degree from Miyazaki University in 1998 and 2000, respectively. From 2000 to 2001, he was an Engineer in Asahi Kasei Corporation, Japan. In 2001, he joined Toyama University, Toyama, Japan, where he was a Technical Official in the Department of Intellectual Information Systems. In 2006, he joined Miyazaki University, Miyazaki, Japan, where he was an Assistant Professor in the Department of Electrical and Electronic Engineering. Since 2015, he is currently a Professor in the Department of Environmental Robotics. His main research interests are Neural Networks and Optimization Problems. In recent years, he has had interest in Biomedical Signal Processing using Soft Computing.

Research of Classification of Palmprint Based on Deep Learning

Kunyu Yu
System Information, Ashikaga University , Ashikaga, City, 3260822, Japan

Hiroshi Matsuki
System Information, Ashikaga University , Ashikaga, City, 3260822, Japan
E-mail: g20692@g.ashikaga.ac.jp, matsuki.hiroshi@g.ashikaga.ac.jp
www.ashitech.ac.jp.

Abstract

After many years of clinical research in traditional Chinese medicine, it was found that the large thenar part of the palm was related to allergic reactions such as the asthma. This paper classifies the thenar part of the palm according to the characteristics of the wrinkles through the transfer learning in the depth learning, so as to play an assistant role in the diagnosis.

Keywords: Deep Learning, Inception V3, Great thenar palmprint, CNN.

1. Introduction

Traditional Chinese medicine has a long history in Asia. With the development of modern science and technology, the intelligentization of Chinese medicine has gradually become a boom. As early as the 1980s, *Sun Tongjiao* developed a diagnosis and treatment system for hepatitis B experts based on research on hepatitis B. *Zou Yunxiang* After conducting in-depth research on diseases of the kidney system in Chinese medicine, a computerized diagnosis and treatment and nursing system was developed. Over 140 Chinese medicine diagnosis expert systems have been developed nationwide during this period. Nowadays, with the rapid development of AI technology, the development of artificial intelligence in medicine can be seen everywhere. For example, some sub-healthy physical condition detectors can quickly detect whether the human body is in the sub-health stage and the coefficients of various physical indicators in just a few minutes, and give a test result and intimate reminders. Not only these aspects, AI also has certain achievements in medical treatment: such as medical robots, intelligent drug development, intelligent diagnosis and treatment systems, and auxiliary diagnosis systems. The upsurge of intelligent informatization of Chinese medicine diagnosis is rapidly developing, which has further promoted the modernization of Chinese medicine and provided many conveniences for people's daily life.

In Chinese medicine, after years of diagnostic experience, Chinese medicine experts have discovered that the thenar part of human palm prints can be found to be related to asthma and other allergic diseases in terms of its lines and roughness. Traditional Chinese medicine can provide diagnostic evidence for some allergic diseases through the characteristics of palm prints in the thenar area. However, TCM judges the condition through observation, pulse diagnosis, inquiry and research, which is subjective. If it is possible to identify and classify the collected large thenar palm print pictures through a computer, according to certain quantification rules, extract the characteristics of the yin and yang large thenar palm prints, and then classify them. Because computer recognition has a certain degree of objectivity, it can play a certain auxiliary diagnostic role in TCM diagnosis and provide convenience for TCM diagnosis.

This article mainly uses deep learning to classify palm prints in the thenar area. According to the relationship between the palm prints of the thenar and asthma and other allergic diseases and their characteristics, collect data sets, preprocess the data images, intercept and segment the thenar regions, and use the TensorFlow framework and GoogleNet's inception V3 model to compare the large thenar part of the palm is classified, so that the model can correctly classify the negative and positive large thenar palm prints under a certain high probability, so as to assist in the diagnosis. However, due to the small difference in characteristics between the two types of thenar palm prints, especially for some pictures of thenar palm prints on both sides of the yin and yang big thenar palm prints, it is also difficult to distinguish between the two types of thenar palm prints by human eyes. , So the computer can be used as an objective reference when classifying in this case.

2. Design of Thenar Palmprint Classifier

In this paper, the TensorFlow framework is used for image numerical calculation, and the inception V3 model is used to classify and train large thenar palm prints.

2.1. *Pretreatment for classification of thenar palm prints*

In the long-term clinical practice of Chinese physicians, it is found that the palm prints of the big thenar can be divided into negative (level I, II) and positive (level III and IV) according to their texture characteristics. Among them, the feminine thenar palm prints are smaller in spacing, presenting a small lattice pattern, with fine mesh, shallow grooves, and more delicate touch; the positive thenar palm prints are larger in spacing, presenting a large lattice pattern, and are distributed Uniform, deep grooves, clear lines, rough touch.

Fig.1 Negative palmprint and Positive palmprint

2.2. *Design of classifier based on convolutional neural network*

The network structure of convolutional neural network is divided into input layer, convolutional layer, ReLU layer, pooling layer and fully connected layer. The network structure of convolutional neural network is divided into input layer, convolution layer, ReLU layer, pooling layer, fully connected layer and output layer. However, in practical applications, the convolutional layer and the ReLU layer are often collectively referred to as the convolutional layer.

2.2.1. *Convolutional layer*

In the convolutional layer, the feature map of the previous layer is convolved by a learnable convolution kernel, and then through an activation function, the output feature map can be obtained. Each output feature map can be combined to convolve multiple feature maps value.

$$\begin{aligned} x_j^l &= f(u_j^l) \\ u_j^l &= \sum_{i \in M_j} x_i^{l-1} * k_{ij}^l + b_j^l \end{aligned} \tag{1}$$

u_j^i The net activation of the jth channel of the convolutional layer l.

x_i^{l-1} Output feature map of the previous layer.

x_j^i Is the output of the jth channel of the convolutional layer l.

*Is the convolution symbol.

k_{ij}^l Is the convolution kernel matrix

b_j^l Is the bias of the feature map after convolution.

M_j Is a subset of the calculated input feature map.

2.2.2. *Pooling layer*

After the original data image of the thenar palmprints is calculated by convolution, it is necessary to perform the pooling operation on the feature map of the thenar palmprints obtained by the convolution to perform spatial size compression. Commonly used pooling operation methods are average pooling and maximum pooling. The pooling layer of the thenar palmprint classification in this paper selects the maximum pooling.

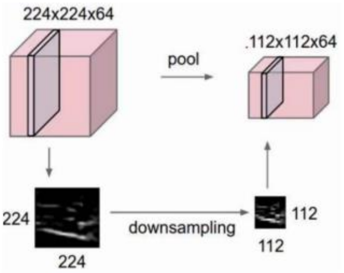


Fig.2 Pooling layer processing effect

2.2.3. *Fully connected layer*

Through convolution and pooling, the local features of the big thenar palm print image are extracted. Full connection uses the weight matrix to combine the local features of the previous big thenar palm print image into a complete image.

$$x^l = f(u^l) \quad u^l = w^l x^{l-1} + b^l \tag{2}$$

u^l The net activation of the fully connected layer l.

w^l Is the weight coefficient of the fully connected network.

b^l Is the bias term of the fully connected layer l.

2.2.4. *Activation function*

In the training process, a function is needed to convert the input large thenar palm print image into an output value to facilitate more intuitive classification. This function is called an activation function. The function of the activation function is to add some non-linear factors to solve the difficulties that cannot be solved by the linear model. The activation function used in this article is the ReLU function.

$$\mathrm{h}(x) = \begin{cases} x & (x > 0) \\ 0 & (x \le 0) \end{cases} \tag{3}$$

2.2.5. *Softmax classifier*

In the output layer, the activation function corresponding to the output layer needs to be selected according to the corresponding goal to be achieved. Generally speaking, if it is a regression problem, use the identity function; if it is applied to image classification research, use the softmax function.

For the study of palmprint classification in this article, the SoftMax function, also known as the normalized exponential function, is used, and the result is a real number between 0.0 and 1.0. The distance between the two probability distributions obtained from the cross entropy function is processed by the Softmax classifier, and the probability distribution of each mutually exclusive output class is returned, so that the range of each element is between (0,1) and the sum of all elements The time is 1.

$$Softmax(z_i) = \frac{e^{z_i}}{\sum_{c=1}^{C} e^{z_c}} \tag{4}$$

z_i Is the output value of the i-th node.
C Is the number of output nodes, that is, the number of categories.

2.3 . *Chapter summary*

This chapter mainly describes the data collection, classification and processing of palmprints in the thenar region, the levels of convolutional neural networks in deep learning, and the functions and functions of each level. In actual operation, feature extraction is performed on palmprint images of the thenar area during the training process, and then the size of the extracted feature maps is compressed, and the category of a certain thenar palmprint image is output through the objective function in the fully connected layer. The probability value of the trend.

3. Results and analysis

Create a new test folder and use the data pictures in the test file to test the accuracy of the model. The test results are as follows.

Table1 .Model Test Result.

Classification of palmprint	Recognition accuracy
Negative	88. 2%
Positive	85. 7%

From the results, it can be seen that using inception V3 has a high accuracy in the palmprint classification of the large thenar area, and it is an effective model.

3.1 Analysis

When the test set pictures are used for verification, the recognition of the more obvious thenar negative palmprints and positive palmprints can reach more than 90%.

Fig. 3 Test result

4. Summary and Reflection

This article mainly applies deep learning to the classification of palm prints of the big thenar, preprocessing the collected data, and intercepting the big thenar area of the palm. According to the characteristics of palm prints in the large thenar area of asthma and other similar diseases, the data set is divided into two types, negative and positive, for training the model.

However, due to the complexity of the thenar palmprint image and the limitations of existing conditions, there are still some shortcomings and areas for improvement. First of all, it is necessary to continuously enrich the data set, collect more data of large thenar palm prints, and establish a database to make the training results more convincing. Second, choose a better model or algorithm, which can improve the classification accuracy.

References

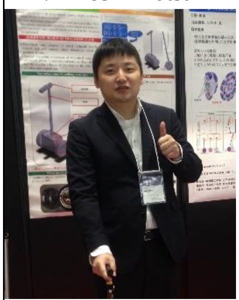
1. "Detailed explanation of convolutional neural network (CNN)website, ", https://zhuanlan.zhihu.com/p/47184529
2. Zhou Zhaoshan. The correlation between palm print morphology and asthma [A]. Respiratory Disease Committee of China Integrative Medicine Association. The Seventh National Conference on Respiratory Diseases of Integrated Traditional Chinese and Western Medicine (1) [C]. Respiratory Disease Professional Committee of Chinese Integrative Medicine Association: China Integrative Medicine Association, 2004: 5.(In Chinese)
3. Xi Yang,Tan Wu,Lei Zhang,Dong Yang,Nannan Wang,Bin Song,Xinbo Gao. CNN with spatio-temporal information for fast suspicious object detection and recognition in THz security images[J]. Signal Processing,2019, 160..(In Chinese)
4. Kemal Adem,Serhat Kiliçarslan,Onur Cömert. Classification and Diagnosis of Cervical Cancer with Softmax classification with stacked autoencoder[J]. Expert Systems With Applications,2018.
5. Fu Yanli.PM2.5 prediction based on error back propagation neural network[J].Computer Knowledge and Technology,2018,14(33):180-181.(In Chinese)
6. .Xiaoying Han,Ruodu Wang. Computation of credit portfolio loss distribution by a cross entropy method[J]. Applied Computing,2016,52(1-2).

Authors Introduction

Ms. Kunyu Yu

She is graduated from Ludong University in China in 2014, and went to Japan to study in the same year. Currently studying for a master's degree at Ashikaga University in Japan

Dr. Hiroshi Matsuki



He received the B.E., M.E., and Ph.D. degrees in electrical and computer engineering from Yokohama National University, 773 Yokohama, Japan, in 2009, 2011, and 2014, respectively. 775 Since 2019, he has been with the Division of Systems and Information Engineering, Ashikaga University, Ashikaga, Japan. His research interests include machine learning, optimization problems, and artificial intelligence.

Online Deep Reinforcement Learning on Assigned Weight Spaghetti Grasping in One Time using Soft Actor-Critic

Prem Gamolped[1], Sakmongkon Chumkamon[1], Chanapol Piyavichyanon[2], and Eiji Hayashi[1]
[1]Department of Mechanical Information Science and Technology, Kyushu Institute of Technology, 680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan
[2]Department of Creative Informatics, Kyushu Institute of Technology, 680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan

Abbe Mowshowitz[3]
[3]Department of Computer Science, The City College of New York, 160 Convent Avenue, New York, NY 10031, USA

E-mail: gamolped.prem129@mail.kyutech.jp, m-san@mmcs.mse.kyutech.ac.jp, piyavichayanon.chanapol402@mail.kyutech.jp, haya@mse.kyutech.ac.jp, amowshowitz@ccny.cuny.edu
http://www.kyutech.ac.jp/

Abstract

Artificial Intelligence and Robotics have become essential and widely used to package food. Packaging an assigned weight spaghetti into a lunch box at one time can be difficult. This paper proposes a solution for one-time grasping using Deep Reinforcement Learning (DRL) based on the Soft Actor-Critic algorithm on the manipulator. Spaghetti detection and segmentation are implemented from the RGB-D camera for the observation. We conclude that the experiment shows the effectively grasped result can almost succeed within 10% of the target weight in the experimental environment.

Keywords: One-Time Grasping, Deep Reinforcement Learning (DRL), Soft Actor-Critic, RGB-D Camera

1. Introduction

In the past years, the technology of robotics in the food industry has tremendous advanced for the task mainly focused on packing, picking, palletizing the food, etc. Many researchers implement artificial intelligence technology with robotics to create self-decision based on observation. The task is repetitive since in packaging the specific amount of spaghetti into the lunch box, the human can pick up the spaghetti and measure the weight before put into the lunch box, as shown in Fig 1. This work proposes a solution for one-time spaghetti grasping using a self-learning system called Reinforcement Learning (RL) based on the off-policy learning RL called

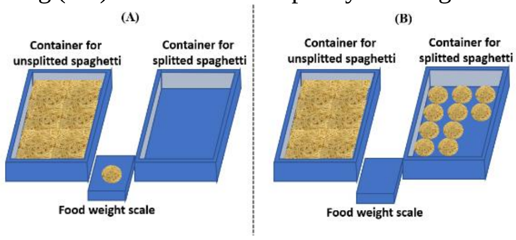


Fig. 1. On the A-side, the worker picks spaghetti from the left container, measures weight, and puts it in the right container, as shown on the B-side.

Soft Actor-Critic (SAC). The spaghetti detection and segmentation system from the RGB-D camera are implemented to recognize the state for observation of RL. We introduce the learning and evaluation system framework for a one-time grasping of the following task. The result of our proposed solution shows that the grasped result can almost succeed within 10% of the target weight.

2. Related Work

The spaghetti grasping task in a manipulated robot can be enhanced by deciding the action to take based on the robot's environment, also called self-decision capability. The RL learning system is proposed in many research papers because it can perform the learning task from the ground up without giving any data to learn at the initial state.

2.1. *System Configuration*

Performing the spaghetti grasping task on the manipulated robot. The controlling, perceiving, and feedback system should be considered, especially the grasping system in the part of the gripper that should match with an object to grasp. Spaghetti is regarded as a soft-body object that can be damaged quickly and causes the worse result in training. The robot arm we use is the Motoman SIA5F articulated arm, as shown in the Fig. 2.

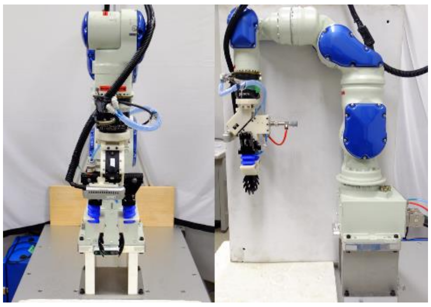


Fig. 2. The Motoman SIA5F, 7-Axis Articulated arm robot with 5KG Payload. Attached with gripper system for the spaghetti grasping task.

Fig. 3 shows the gripper system that consists of the following

- The Twintool gripper can switch between two attached tools motorized using Servo motor
- Soft Robotics gripper is the pneumatic actuated soft gripper for food grasping and an additional noodle tong for spaghetti grasping.
- The Vacuum tool to vacuum the flat face food such as ham, cut bread, bologna, etc.
- The RGB-D camera will provide the color and depth image for the robot's perception.

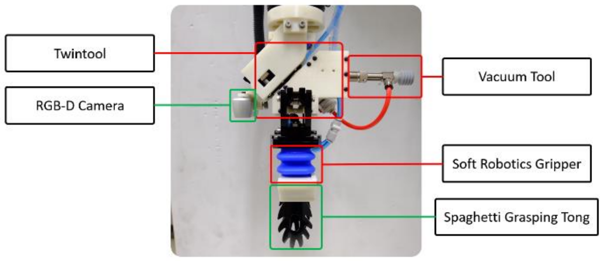


Fig. 3. Gripper system

We created the spaghetti detection system to identify the probability of being the spaghetti and the spaghetti instance segmentation to get a centroid of the shape to help perform the limit of the environment in RL for the agent. We collected the lunch box and spaghetti dataset then manually labeled and annotated it using the COCO annotator. The dataset was trained using the Hybrid Task Cascade (HTC) method on the MMDetection framework. Fig. 4 shows the vision system for spaghetti. The left is the target box, and the right side is the spaghetti container.

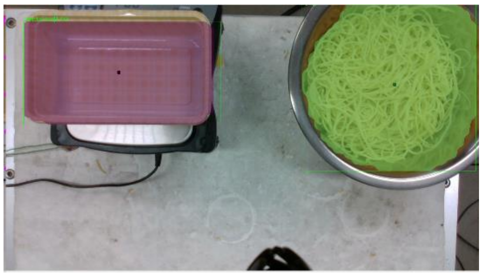


Fig. 4. Spaghetti detection and instance segmentation system

2.2. *Deep Reinforcement Learning*

The self-learning system that does not require any instruction to take action; instead, the agent will perceive the environment from the sensor and take action based on the observation to get the maximum reward. The agent will do the trial-and-error, which yields the positive or negative reward based on the action that agent takes through the reward function. DRL mainly consists of two parts shown in Fig. 5

- Agent (Human or Robot)
- Environment

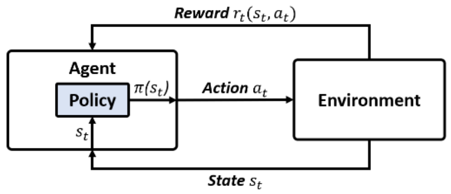


Fig. 5. Spaghetti detection and instance segmentation system

In DRL, we will consider in timestep t according to the finite-horizon discounted Markov Decision Process (MDP) wherein each t; the agent is going to get state s_t from the environment, the policy function is going to map from what the agent observes to action a_t, in this timestep t, the reward $r_t(s_t, a_t)$ will determine the good or bad that the agent took action. In our research, timestep t is set to be 1. Thus episode length will be one. The method of DRL we use is Soft Actor-Critic (SAC) [1]. The SAC is the off-policy algorithm that can handle the complex environment. The algorithm is well-known in the features of the following:

- SAC has the entropy regularization that agent will get the best of the trade-off between explore-exploit, which will encourage the agent to explore more.
- A consequence of the explore-exploit trade-off yields the training faster later because the agent gets to know a lot in the environment.
- SAC can work on the continuous action space.

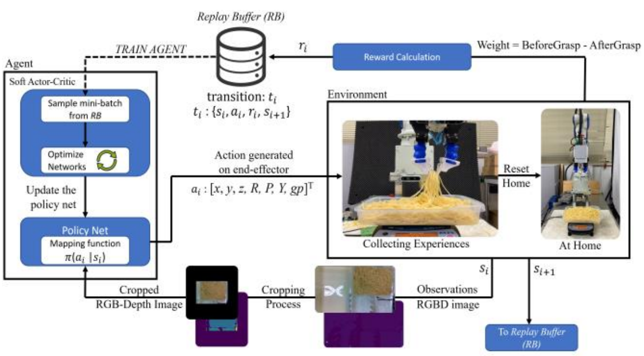


Fig. 6. The DRL framework of the training and evaluation system on one-time spaghetti grasping task using DRL

The proposed training framework of one-time spaghetti grasping will imitate the MDP process that wherein each timestep t; we will take an RGB-D image of the spaghetti and do the image processing by cropping unwanted area to get the image that has a high percent of spaghetti included as a state t, and the policy network we use is the convolutional neural network (CNN) to map from state t to action a_t, the action space of our framework consists of the following values: position, orientation, and gripper width of the end-effector; thus, our action a_t is x, y, z, R, P, Y, gp. The command will be sent to the sequence of the motion we create for picking the spaghetti and measuring the weight, then randomly put spaghetti back to the container. The reason to randomly put is to avoid the same observation every timestep t. The reward $r_t(s_t, a_t)$ is calculated from the following formula.

$$r_t(s_t, a_t) = \frac{(target - |measured - target|)}{target} \qquad (1)$$

The agent will calculate the reward function based on Eq. (1) if the measured or the weight from grasped spaghetti is in 10% of the target weight. In our research, the target weight is 50-gram; thus, the reward function will be activated when the measurement is between 45-55 grams. Otherwise, the $r_t = -0.5$ and after the reward, the next state s_{t+1} will be taken with the cropping process to complete the transition for replay buffer to complete one timestep.

3. Experiment

To perform the experiment of the one-time spaghetti grasping, we set up the experiment inside the laboratory environment.

3.1. *Setup the Experiment*

We will start to prepare spaghetti for 300-gram and boil for 7 minutes; after that, we will shake off the water until almost dry and put the oil in a bit to prevent sticking spaghetti because it can damage and make training results worse. Fig. 7 shows the prepared environment and spaghetti after boiling.

Fig. 7. Prepared environment before starting training

3.2. *Training*

We have set the training for two sessions; each session run through 2,000 episodes, the first one of A, we use the reward function that has no negative and second run of B, the negative reward is included as shown in Eq. (1). However, the spaghetti will get dried very fast. Thus, the spaghetti entangling problem can happen, and the training can worsen because the physical weight will change a lot from its undamaged spaghetti. To solve this problem, we spray a little bit of water onto the spaghetti face every 20 minutes and change new spaghetti every 200 episodes to make sure of training the undamaged spaghetti. Fig. 8 shows the training result of the two 8sessions described above. After that, we get the Episode Reward Mean to conduct the training trend. We can see that the model that negative reward is included is better because we can filter out the unimportance of action that led to always getting the positive reward; also, we can say that it is a Sparse Reward. Table 1 shows the comparison result of two training sessions. Table 2 shows the Mean Absolute Percent Error (MAPE) which is used to measure the percent error of model prediction.

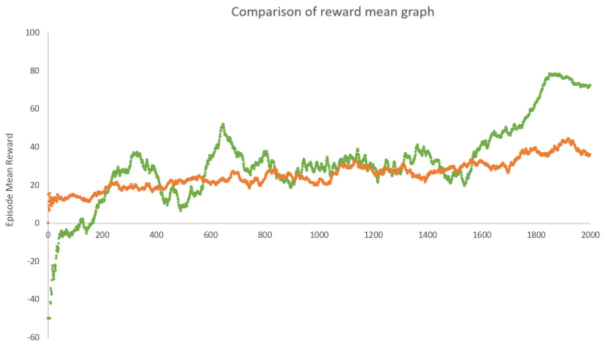


Fig. 8. The reward mean graph from two training sessions

Table 1. Training result comparison of two sessions

Parameters	A [g]	B [g]
Mean	46.72	48.90
Absolute Mean	19.29	14.08
Standard Deviation	23.41	17.91

Table 2 shows that the model with the negative reward included could do lower the MAPE, which mean that with the lower value, it is possible to get near the 10% of target weight

Table 2. MAPE comparison of two sessions

Parameters	A [%]	B [%]
MAPE	38.97	28.44

4. Evaluation

The same framework is used to evaluate the model with lower MAPE or model B to show how accurate the model is; thus, we set up the same environment and proceed with the same training step. We acquired only 100 episodes of the session, calculated how many samples are in the 10% of the target weight, and plotted the evaluated model's distribution. Fig. 9 shows the distribution of the evaluated model. Table 3 shows the following described above.

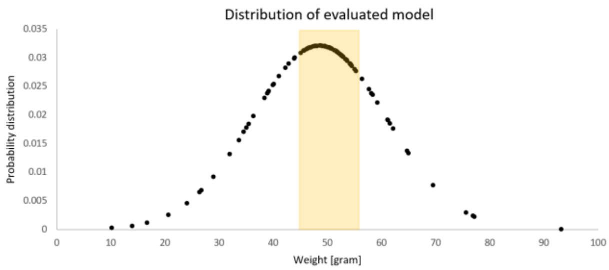


Fig. 9. The distribution of evaluated model

Table 2. Result of evaluated model

Parameters	B	unit
Mean	48.52	grams
Absolute Mean	8.31	grams
Standard Deviation	12.41	grams
Grasp in the range of acceptable error	58	times

5. Conclusion

In this research, we proposed the solution to grasp a one-time spaghetti at the assigned, and the result shows that we can almost succeed within 10% of the target weight. However, there is still a problem that spaghetti gets to dried quickly and it can cause the entangling problem and can worsen the training result. Also the reward function should have the negative reward to prevent the sparse reward. In the future work, we will try to control the condition of spaghetti such as time and the humidity of spaghetti.

References

1. [1801.01290] Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor (arxiv.org)

Authors Introduction

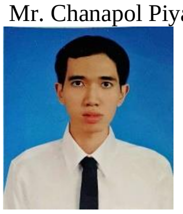
Prem Gamolped

He received bachelor degree in Engineering in 2021 from Robotics and Automation Engineering, King Mongkut's University of Technolgoy North Bangkok in Thailand. 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.

Mr. Chanapol Piyavichayanon



He received bachelor degree in Engineering in 2021 from Mechanical Engineering, Chulalongkorn University in Japan. He is currently a Master student at Kyushu Institute of Technology and conducts research at Koga 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

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. -Inference in the same phrase including ostinato-

Haruna Yamasaki
Department of Mechanical Information Science and Technology, 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

Eiji Hayashi
Department of Mechanical Information Science and Technology, Kyushu Institute of Technology
680-4, Kawazu, Iizuka-City, Fukuoka, 820-8502, Japan

E-mail:
http://www.kyutech.ac.jp/

Abstract

We have developed a system to reproduce human-like piano performance with an automatic performance device. In order to perform on a par with a human, it is necessary to express the intensity of the notes and other aspects of the performance. Therefore, using F. Chopin's music as a subject, we studied the changes in performance expression when similar phrases such as ostinato are repeated. Then, we developed an interactive music editing system to realize efficient music editing.

Keywords: Player Piano, DP Matching, Knowledge Database, Computer Music, Music Interface

1. Introduction

The purpose of this research is to reproduce human-like piano performance with an automatic performance system. The automatic musical performance system developed by Hayashi et al. in 1996 has the same performance capability as a human, but it does not have the ability to infer musical expressions such as the strength of notes[1]. Therefore, we inferred the musical expression of the same performer in an unperformed piece of music using the performer's existing performance information and score information. In this paper, we study the changes in performance expression when similar phrases such as ostinato are repeated in a piece by F. Chopin, and describe the inference method.

2. Editing Support System

2.1 Performance Information

The songs listed in Table 1 of CrestMusePEDB, which were created from the actual performance of Vladimir Davidovich Ashkenazy (Russia), were used as performance information. From now on, the names of the songs are A to F in Table 1.

Table.1 Song information

Character	Song Title	Tonality
A	Prelude Op.28 No.7	A-Dur
B	Prelude Op.28 No.4	E-Moll
C	Prelude Op.28 No.15	Des-Dur
D	Prelude Op.28 No.20	c-Moll
E	Etude Op.10 No.4	Cis-Moll
F	Nocturnes Op.9 No.2	b-Moll

In this study, song A was inferred. Therefore, we also used the data of the eight performers who played the song A in Table 2. Hereafter, the names of the performers will be referred to as ① through ⑧.

Table.2 Pianists who played song A

Number	Pianist Name	Nationality
①	Maria Martha Argerich	Argentina
②	Claudio Arrau León	Chile
③	Jorge Bolet	Cuba
④	Rafał Blechacz	Poland
⑤	Alfred Denis Cortot	Switzerland
⑥	Dang Thai Son	Viet Nam
⑦	Maria João Pires	Brazil
⑧	Ivo Pogorelich	Croatia

2.2 Score information

Table 3 shows the variables used for musical notation information in the MIDI (Musical Instrument Digital Interface) standard, where Key is the position of the keyboard, Velo is the strength of the note, Gate is the length of the note, Step is the interval to the next note, and Time is the playing time.

Table.3 Note Information

Para meter	Key	Velo	Gate	Step	Time	Bar
Unit	—	—	ms	ms	ms	—
Refer ence	21～108	1～127	—	—	—	—

2.3 Performance Reproduction Flow

Figure 1 shows the flow of reproducing a piece of music that has not been played by the performer. The features of the performer extracted from the performance and score information are applied to the score information of the unplayed piece to infer and reproduce the performance expression.

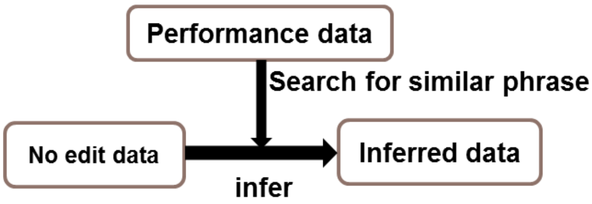


Fig.1 Flow of performance reproduction

2.4 Search systems

From previous research, it was found that phrases with the same note value sequence are similar in performance expression. Based on this theorem, we have developed a search system that uses DP(Dynamic Programming) matching to find phrases (similar phrases) that have a similar sequence of note values to the phrase of a song that we want to reproduce (search phrase)[2]. This method calculates the shortest distance and path between two patterns, and expresses the similarity between them in terms of distance. The range of the search phrase was set arbitrarily. However, phrases that are considered to be musically coherent, such as phrases connected by "slurs," are set as articulations, but since these phrases were not found in this study, the search was divided into three phrases as shown in Figure 2.

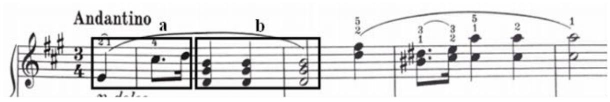


Fig2. Examples of search phrases

2.5 Selection system

When multiple similar phrases with the same DP matching score are found, the selection system selects the phrase that is most suitable for inference (the optimal phrase) using the following five selection indices.

(1) Selection based on dynamics symbols and Velo
(2) Selection by StepRate
(3) Selection by staccato removal
(4) Selection by similarity of pitch change
(5) Selection by Gate

2.5.1 Selection based on dynamics symbols and Velo

Dynamics symbols are thought to affect performance expression, We selected the ones with matching dynamics symbols. Then, those with too large or too small Velo were subsequently excluded.

2.5.2 Selection by StepRate

The StepRate is the ratio of the theoretical time in the score to the Step. The performance data of the performers in Table 2 is used to select phrases with high SR correlation from similar phrases. StepRate will be abbreviated as SR from now on.

2.5.3 Selection by staccato removal

Compare the length of the notes before and after the staccato part of the B~F song. Exclude any search phrase that exceeds the average of its results.

2.5.4 Selection by similarity of pitch change

Phrases with similar pitch changes are considered to have similar performance expressions, and phrases with higher similarity in pitch changes are selected.

2.5.5 Selection by Gate

From the average of 8 pianists as well as SR The phrase with the highest correlation of Gate was selected.

2.6 Inference system

The inference system performs Step and Velo inference.

2.6.1 Inference of Step

The Step inference uses n_SR, where the nth n_SR is as in “Eq. (1)”.

$$\mathrm{n\_SR} = \mathrm{SR}(\mathrm{n}+1) \;/\; \mathrm{SR}(\mathrm{n}) \tag{1}$$

This way, once the SR for the first note is determined, the SR for the following notes can be determined. n_SR is used from the optimal phrase selected in 2.5. Since the previous inference method gave the same inference results for all eight phrases, we used the data of the performers in Table 2 to infer each phrase. First, to determine the SR of the first note of each phrase, we calculated the BPM (Beats Per Minute) of songs B~F and the performers in Table 2. Since the BPMs of songs A~F are different, it is necessary to match the BPM of song A. The average BPM of the performers of song A in Table 2 is 37. Using this, the ratio to Ashkenazy's BPM of songs B~F is calculated and listed in Table 4. Using the BPM ratios in Table 5, the inferred SR from the average SR of songs B~F is obtained from “Eq. (2)” and is shown in Table 4.

$$\mathrm{Inference\ SR} = \mathrm{SR\ average} \;/\; \mathrm{BPMratio} \tag{2}$$

From the data of the performers in Table 2, calculate the average SR for each phrase using only the values whose average SR is between 3.4 and 4.2 as calculated by inferred SR. This value is used as the SR of the first note of each phrase, and the SR of the second and subsequent notes is calculated using n_SR to obtain the inferred value of Step.

Table.4 BPM ratio and SR average

	B	C	D	E	F
BPM ratio	0.66	0.52	1.06	0.21	0.55
SR average	2.49	1.83	3.63	0.72	2.31
Inference SR	3.77	3.51	3.44	3.39	4.19

2.6.1 Inference of Velo

For song A, the dynamics symbol is p only. The range of p is from pp to f. As in the Step inference, we determine the inferred value of Velo by calculating the average per phrase and average per note from the performers in Table 2. n_Velo for the nth note is Eq (3).

$$\mathrm{n\_Velo} = \mathrm{Velo}(\mathrm{n}) \;/\; \mathrm{Velo}(\mathrm{n}-1) \tag{3}$$

The first note of each phrase is assigned the average value of the phrases that fall within the range of p of the performers in Table 2; from the second note onward, the inferred value of Velo is calculated by calculating n_Velo from the average value of each note of the performers in Table 2.

3. Inference experiment

3.1 Experimental Method

We conducted an experiment to compare the actual performance data of song A with the inference results of the system.

3.2 Results of experiment

The right-hand graph comparing Step's inferred values, the actual performance, and the data before editing is shown in Figure 3. Similarly, the right-hand graph of Velo is shown in Figure 4.

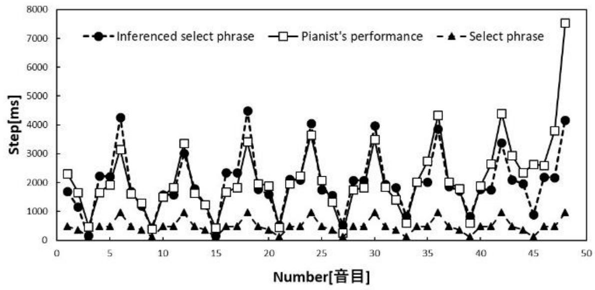


Fig.3 Comparison of the right-hand Step

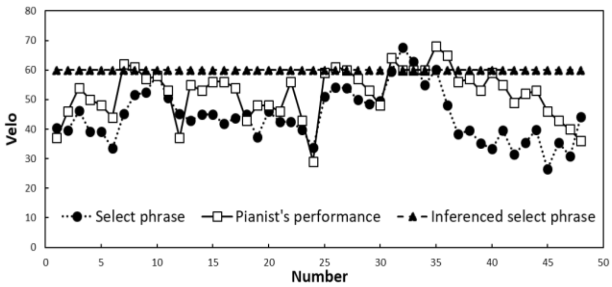


Fig.4 Comparison of the right-hand Velo

4. Consideration

The selection system and inference system used data from other pianists as well, resulting in more accurate results than before. From this, we believe that we were able to demonstrate the effectiveness of this system. However, we have not yet been able to reliably mimic the loudness of the notes or the way the notes are extended.

We believe that this is due to the search system that judges similar phrases based only on the length of the sound, although the connection before and after is also important in musical expression.

5. Conclusion

This year, we introduced an inference system that combines the performance data of other performers with the performance information editing support system. The problem was that when the same phrase was repeated twice, the inference results were exactly the same, but in this year's experiment, we were able to infer each phrase by combining the performance data of the performer we wanted to infer with that of other performers. However, since this method relies on the performance data of other performers, it cannot be used for music with only a few performers or for music for which there is no data to begin with, which is an issue for the future.

We would also like to study the development of an alternative search system to DP matching that takes into account the connections before and after.

Therefore, we will consider developing a more versatile system with the addition of another method.

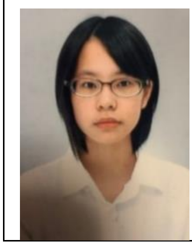
6. References

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

Ms. Haruna Yamasaki



She received bachelor degree in Engineering in 2021 from mechanical system engineering, Kyushu Institute of Technology in Japan. 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).

Weight estimation for noodle products in food layout of a home replacement meal

Tomofumi Tsuji
Department of Mechanical Information Science and Technology, Kyushu Institute of Technology
680-4, Kawazu, Iizuka-City, Fukuoka, 820-8502, Japan

Sakmongkon Chumkamon[1], Chanapol Piyavichayanon[2], Ayumu Tominaga[3], Ryusuke Fujisawa[1], Abbe Mowshowitz[4], and Eiji Hayashi[1]
[1] Department of Mechanical Information Science and Technology, Kyushu Institute of Technology,
680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan
[2]Department of Creative Informatics, Kyushu Institute of Technology,
680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan
[3]Department of Creative Engineering Robotics and Mechatronics Course, National Institute of Technology Kitakyushu Colllege, 5-20-1 Shii, Kokuraminamiku, Kitakyushu, Fukuoka, 802-0985, Japan
[4]Department of Computer Science, The City College of New York,
160 Convent Avenue, New York, NY 10031, USA

E-mail: tsuji.tomofumi552@mail.kyutech.jp, m-san@mmcs.mse.kyutech.ac.jp,
piyavichayanon.chanapol402@mail.kyutech.jp, tominaga@kct.ac.jp, amowshowitz@ccny.cuny.edu,
fujisawa@ces.kyutech.ac.jp, haya@mse.kyutech.ac.jp

http://www.kyutech.ac.jp/

Abstract

In recent years, there has been an increasing demand for robot automation to improve productivity in the food layout of lunchboxes and prepared foods in the Japanese home replacement meal industry. In this research, we are developing autonomous work robots that can perform home replacement meal tasks and developing the technology for industrial food automation using Artificial Intelligence (AI) to improve productivity, security, and safety. In this paper, we perform weight estimation of the served object to identify the amount of spaghetti grasped. We created our dataset of spaghetti used for weight estimation. The spaghetti is in different types of containers, the various weight of spaghetti, and the random position of spaghetti in the robot workspace. We also propose a deep learning method using RGB-D cameras for weight estimation and describe its validation and evaluation.

Keywords: Factory Automation Robots, image processing, noodle products, Weight Estimation, Deep Learning, General Dataset

1. Introduction

In recent years, there has been an increasing demand for robot automation to improve productivity in the food layout of lunchboxes and prepared foods in the Japanese home replacement meal industry. In this research, we are developing autonomous work robots that can perform home replacement meal tasks

In a previous study, we developed a system that can adequately grasp and serve solidified foods such as fried chicken and rice balls. However, in addition to solid foods, noodle products such as spaghetti are used as ingredients for lunch boxes and prepared foods in the foodservice industry. In addition, it is necessary to consider the quantitative properties of noodle products when placing them in containers. To realize quantitative serving of noodle products, functions that apply previous studies are required.

In this research, we perform weight estimation of the served object to identify the amount of spaghetti grasped. We also propose a deep learning method using RGB-D cameras for weight estimation and describe its validation and evaluation.

2. System Overview

2.1. *System configuration*

The system configuration of this research is shown in Fig. 1. In this system, a robot grasps spaghetti, arranges it in a specific container, and estimates its weight. The system consists of a visualization processing unit and a robot control unit. First, PC1 receives the information from the RGB-D camera mounted on the robot and performs the visualization process. The visualization process includes real-time object recognition and 3D processing of food. Next, PC2 receives the information from PC1 and sends the motion planning signal to the controller, which enables the robot to grasp and serve the food. PC1 also performs weight estimation after the spaghetti is served.

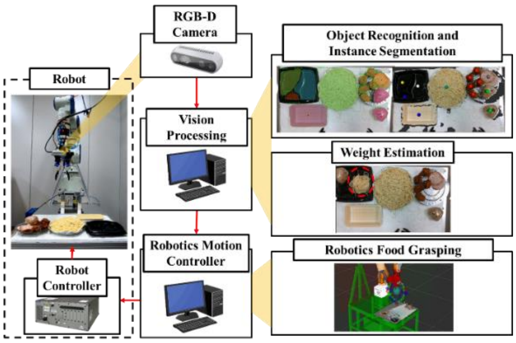


Fig. 1. System configuration diagram

2.2. *Weight Estimation System in Spaghetti*

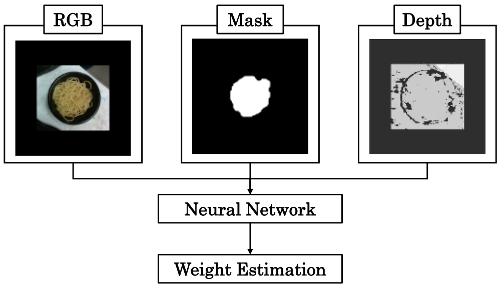


Fig. 2. Weight estimation system configuration diagram

The weight of the spaghetti is estimated by using the information from the RGB-D camera as well as the object recognition system. By using the three data as feature values, when spaghetti is placed in a container, the system can identify the object for weight estimation by RGB and Mask, and it can also specify the surface shape of the object in three dimensions by Depth. The goal of the neural network is to capture the surface shape of the object and estimate the density of the entire object, including the container.

3. Robot configuration

The appearance of the robot in this study is shown in Fig. 3. The robot is a 7-axis vertically articulated robot that has a shape like a human arm and can move with a high degree of freedom because it is required to perform the same work as a human in a food factory. The robot is equipped with an RGB-D camera, an end-effector, and a force torque sensor. The robot is equipped with two types of grippers to grasp the object. The tongue gripper is used for grasping foods with rough surfaces such as fried foods, rice balls, and spaghetti. Vacuum gripper is for gripping objects with a smooth surface, such as ham and containers. The gripper can be selected instantly by the servo motor according to the target food.

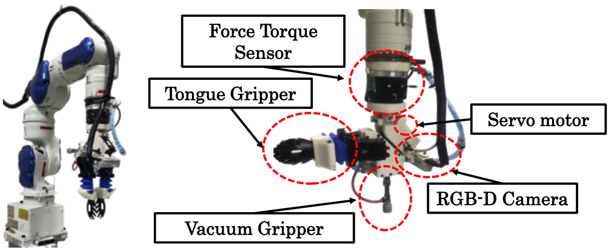


Fig. 3. Appearance of the robot

4. Building Deep Learning Models

Since the training data for our system are RGB, Mask, and Depth, we need to consider the complexity of the neural network and the reduction of computational complexity for a large amount of data. Therefore, we used the idea of ResNeXt[1], a deep learning model, to build a model that enables weight estimation. ResNeXt is an extension of ResNet[2] (Residual Network), which solves the problem of gradient loss due to too many layers in a neural network by using a residual block consisting of a convolutional layer input and a Shortcut Connection. The Shortcut Connection outputs residuals across layers to prevent gradient decay. In addition, ResNeXt extends the residual blocks of ResNet to branch out a new dimension of cardinality, which improves the expressive power. Fig. 4 shows the structure of the deep learning model we constructed. For each input RGB, Mask, and Depth, we designed a CNN convolutional layer, a maximum pooling layer, and a Basic Block (Cardinality=32) which is the basis of ResNeXt. The structure of the Basic Block is shown in Fig. 5. The Basic Block is designed again by fusing these three inputs, and then the weight of spaghetti is predicted by the average pooling layer and the total combination layer.

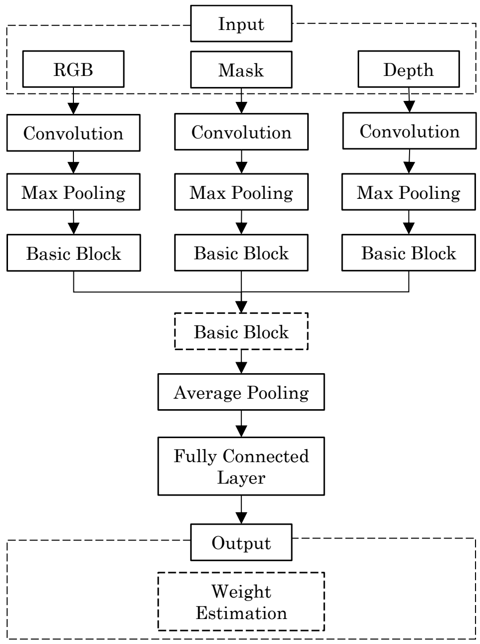


Fig. 4. Weight estimation system configuration diagram

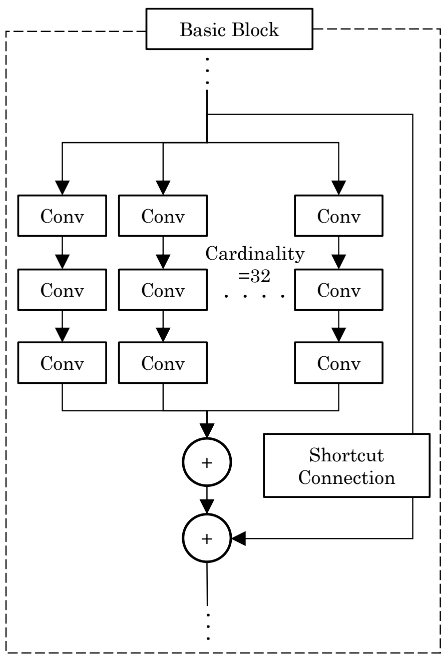


Fig. 5. Configuration of Basic Block

The activation function immediately after the convolutional layer is an ELU (Exponential Linear Unit), and the relationship between the output $f(x)$ and the input x of the ELU is shown in Eq. (1). in the ELU, if the input value is negative, the output value also returns a negative value. In ELU, if the input value is negative, the output value will also return a negative value, thus speeding up the convergence of the loss and improving the stability of the signal.

$$f(x) = \begin{cases} \alpha(e^x - 1), & x \le 0 \\ x, & x > 0 \end{cases} \quad (\alpha > 0) \qquad (1)$$

5. Experiment

5.1. *Dataset*

We created a dataset with weight as the correct answer label to develop a weight estimation system. The dataset consisted of one pair as RGB, Mask, and Depth for objects whose weight was measured in 0.1g increments. Since spaghetti is irregularly shaped, we used six different containers of different shapes and depths to obtain the data. The weight of the containers ranged from 49.8g to 200.3g, considering the maximum grasping capacity of the robot and the spaghetti in a medium meal. The dataset consisted of 3,806 pairs with respect to weight and container. An example of the dataset is shown in Fig. 6. The dataset was randomly divided into 80% (3044 pairs) training data, and 10% (381 pairs) validation data, and 10% (381 pairs) test data. This dataset is a general-purpose dataset that considers each data set independent of the surrounding background, camera position, and height.

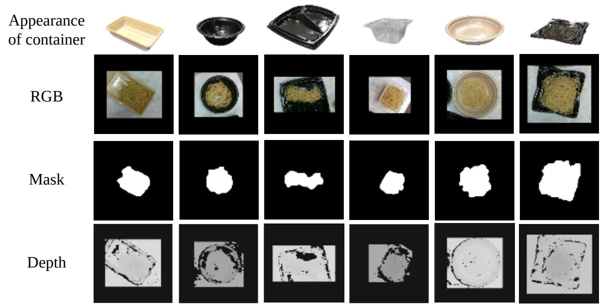
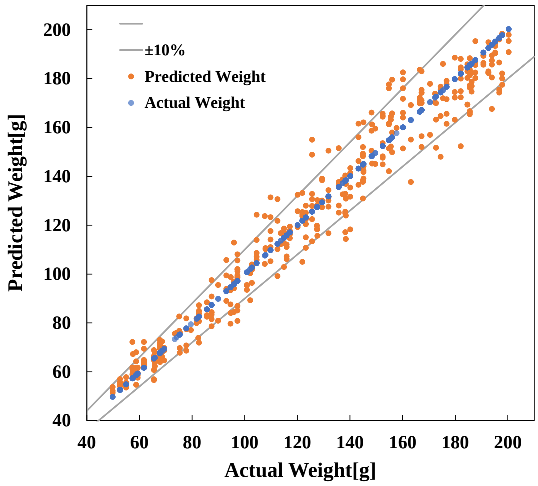


Fig. 6. Example of the dataset for the spaghetti weight estimation, including the various types of containers.

5.2. *Result*

Weight estimation with test data was performed using data trained by a deep learning model. The number of training sessions was set to 3,000, the learning rate to 0.001, and the batch size to 16. The results of the predictions against the actual measurements are shown in Fig. 7. Histogram of percent error and Frequency of percent error is shown in Fig. 8 and Table 1, respectively. The mean absolute error MAE[g], the mean absolute percent error MAPE[%], and the standard deviation, maximum error, and minimum error for each unit are shown in Table 2. Eq. (2) and (3) are used to calculate MAE and MAPE, respectively.

Fig. 7. Distribution of predicted values to actual measured values



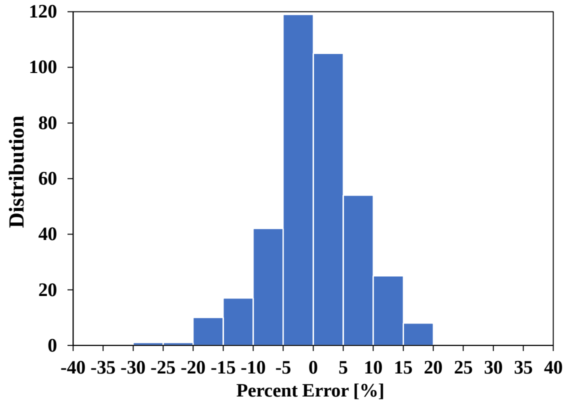


Fig. 8. Histogram of percent error

Table 1. Frequency of percent error

Percent Error [%]	Distribution
-35 ~ -30	0
-30 ~ -25	1
-25 ~ -20	1
-20 ~ -15	10
-15 ~ -10	17
-10 ~ -5	42
-5 ~ 0	119
0 ~ 5	105
5 ~ 10	54
10 ~ 15	25
15 ~ 20	8
20 ~ 25	0
25 ~ 30	0
30 ~ 35	0

$$MAE = \frac{1}{N}\sum |A_t - F_t| \quad (2)$$

$$MAPE = \frac{1}{N}\sum \frac{|A_t - F_t|}{A_t} \quad (3)$$

Table 2. Evaluation index of predicted values

Parameters	Error [g]	Percent Error [%]
Mean Absolute	6.65	5.38
Standard Deviation	6.02	4.52
Maximum Error	29.7	26.1
Minimum Error	0.00961	0.0112

5.3. *Discussion*

In this experiment, from Fig. 7 and Table 2, the MEPE and standard deviations are 5.38% and 4.52%, respectively, which means that the weight estimation can be done with an accuracy of ±10% percent error. The histogram in Fig.8 also shows that -5%~5% is the most significant percentage, and the percentage with large errors is small, which indicates that the learning and data are valid. As a comparison to the ResNeXt model, Table 3 show the results of the ResNet (Cardinality=1 in Fig. 4) model, where the MAPE and standard deviations are 6.40% and 4.91%, respectively, indicating that the accuracy of the ResNeXt model is higher than that of the ResNeXt model. This shows the importance of the complexity of the neural network.

Since there was no consistency in the data with relatively large errors in both models, it is expected that the accuracy can be improved by increasing the number of such data as training data and by reviewing the training model and parameter design.

Table 3. Evaluation index of predicted values (ResNet)

Parameters	Error [g]	Percent Error [%]
Mean Absolute	7.77	6.40
Standard Deviation	6.15	4.91
Maximum Error	33.4	31.3
Minimum Error	0.0271	0.0194

6. Conclusion

In this study, we developed a system that can estimate the weight of noodle products. This system enables the identification of the grasping amount of noodle products. The versatility of the dataset and the balance between the complexity of the neural network and the amount of computation are examined theoretically, and as a result, it is possible to estimate the weight of spaghetti with an accuracy of generally less than 10% error.

References

1. S.Xie, R.Girshick, P. Doller, Z.Tu, K.He, UC San Diego. "Aggregated Residual Transformations for Deep Neural Networks." In CVPR, 2017.
2. K. He, X. Zhang, S. Ren, and J. Sun. "Deep residual learning for image recognition." In CVPR, 2016.

Authors Introduction

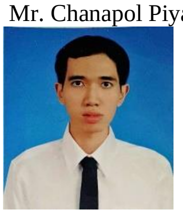
Mr. Tomofumi Tsuji

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.

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.

Mr. Chanapol Piyavichayanon



He received bachelor degree in Engineering in 2021 from Mechanical Engineering, Chulalongkorn University in Japan. He is currently a Master student at Kyushu Institute of Technology and conducts research at Koga 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.

Assoc. Prof. Ryusuke Fujisawa

Assoc. Prof. Prof. Ryusuke Fujisawa is an associate professor in the Department of Intelligent and Control Systems at Kyushu Institute of Technology. He received the Ph.D. (Dr. Eng.) degree from The University of Electro-Communications in 2009. His research interests include Swarm intelligence, Swarm robotics and Distributed autonomous system. He is a member of Japan Society of Civil Engineers (JSCE), Information Processing Society of Japan (IPSJ), The Society of Instrument and Control Engineers (SICE) and The Robotics Society of Japan (RSJ).

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

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).

Cognition of surrounding conditions for a field robot
- Slope detection using a multilayer perceptron classifier with point cloud as input-

Takumi Tomokawa
Department of Mechanical Information Science and Technology, Kyushu Institute of Technology
680-4, Kawazu, Iizuka-City, Fukuoka, 820-8502, Japan

Sakmongkon Chumkamon[1], Ayumu Tominaga[2], Sylvain Geiser[1], Ryusuke Fujisawa[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: tomokawa.takumi163@mail.kyutech.jp, m-san@mmcs.mse.kyutech.ac.jp, tominaga@kct.ac.jp, geiser.nathan-sylvain@mail.kyutech.jp, fujisawa@ces.kyutech.ac.jp, haya@mse.kyutech.ac.jp,
http://www.kyutech.ac.jp/

Abstract

In the Japanese forestry industry, automation of work to supplement labor is desired to achieve sustainable forest management. In this study, the field robot for the automation of forestry is developed. In the field robot, recognition of the surrounding situation is an important function for safe movement. In this paper, we focus on the recognition of terrain. The terrain in a mountainous area has various conditions such as slope, presence of weeds and trees, and unevenness. In this study, the classifier for ground and sloped surfaces using Multi Layered Perceptron (MLP) is developed. This classifier classifies each point of the 3D point cloud acquired from the RGB-D camera into the ground plane of the robot and the slope plane where the robot cannot climb. The accuracy of the classification was verified by training the classifier on a dataset acquired in a real environment.

Keywords: PointCloud, Multi Layered Perceptron, Field Robot, Forestry, ROS

1. Introduction

In the Japanese forest industry, automation of the work to decline the labor is desired to realize sustainable forestry(1). In this study, the forestry field robot was developed. Recognition of the surrounding environment such as a terrain and objects are an essential function for autonomous moving. We focus on the recognition process of the terrain because the terrain in a mountainous has various conditions such as slope, presence of weeds and trees, and unevenness. To recognize and classify these conditions, Multi Layered Perceptron (MLP) was used 3D point cloud as input which acquired from the RGB-D camera. The accuracy of the classification was verified by training the classifier on a dataset acquired in a real environment.

2. Overview of the robot

The developed robot is shown in Fig. 1, which built on an ATV (All-Terrain Vehicle) as its platform. RGB-D camera (D435, Intel Ltd. Co.) is equipped on the front of the robot. The total height of the robot is about 1.7 m, and the sensor is mounted at a height of 1.1 m. The 3D point cloud was obtained from the RGB image and Depth image obtained from this camera.

Fig. 1. Overview of the robot

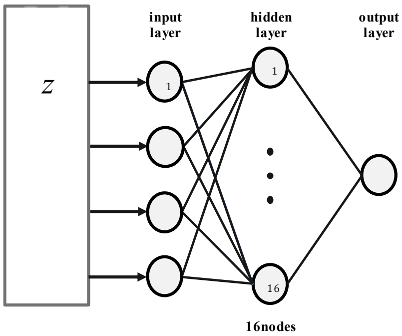


Fig. 2. Overview of the MLP built

$$X = \left[z, n_x, n_y, n_z\right] \tag{1}$$

$$Y = \begin{cases} 1 \ (\rho \geq 0.5) \\ 0 \ (\rho < 0.5) \end{cases} \tag{2}$$

3. Classification of ground plane and inclined plane by MLP

MLP is a machine learning framework that connects multiple simple perceptron to determine nonlinear discriminative boundaries in feature space. In this study, the slope surface is detected by treating 3D point cloud as a class classification with input data.

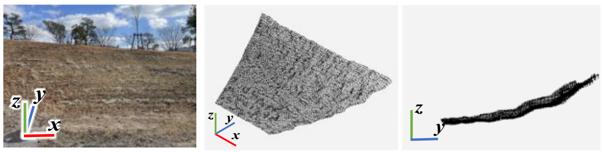
3.1. *MLP Design*

Fig. 2 shows the overview of designed MLP in this study. This MLP would be classify the ground plane and the inclined plane. In creating the classifier, the feature value of each point is a four-dimensional feature value shown in equation(1), which consists of the z-coordinate value (vertical direction) and each element n_x, n_y, and n_z of the local normal vectors n_i. The local normal vector is the normal vector of the approximate plane obtained from the point group contained within the radius r[m] from p_i for any point $p_i = (x_i, y_i, z_i)$ in the point group. By using such features, we can handle the relative height and angle difference between each observed point and the robot. The middle layer has only one layer with 16 nodes, and the activation function is the ReLU[(2)] function. The output layer has 1 node and the activation function is the Sigmoid function. The output ρ of the output layer takes values from 0.0 to 1.0, where less than 0.5 indicates that it is classified as a ground plane, and more than 0.5 indicates that it is classified as an inclined plane. Therefore, the teacher label for the ground plane was set to 0 and the teacher label for the slope plane was set to 1 using equation(2).

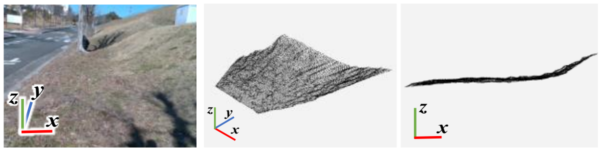
3.2. *Creating a Data Set*

To create the dataset, we collected point clouds of the ground surface covered with weeds, including the ground surface where the robot can run and the inclined surface where the robot cannot climb (inclination angle of more than 20 degrees) within the angle of view of the camera. As shown in Fig. 3, we prepared five sets of point cloud data each for MLP training, where the inclined surface was in front, right, and left of the robot. The z-coordinate values and local normal vectors of each point were divided into the ground plane and the inclined plane in advance, and 0 and 1 supervisory labels were given.
Fig. 4 shows an example of the training data set. The evaluation data was also prepared in the same way as the training data, with five sets of each state.

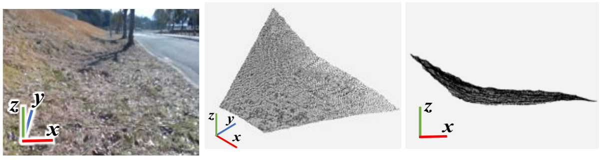
The numerical value of the local normal vector, the input to the MLP, depends on the radius of the neighborhood r[m]. Therefore, it is thought that small irregularities on the ground surface are less likely to be misidentified as slopes. Therefore, we prepared three different datasets with r values of 0.05m, 0.1m, and 0.5m, and trained the MLP with each input vector X. The number of times the MLP was trained was determined by the number of input vectors. The number of times the MLP was trained was set to 10 in each case.



(a) An inclined plane exists in front of the robot



(b) An inclined plane exists on the right side of the robot



(c) An inclined plane exists on the left side of the robot

Fig. 3. Examples of point clouds used to create the data set

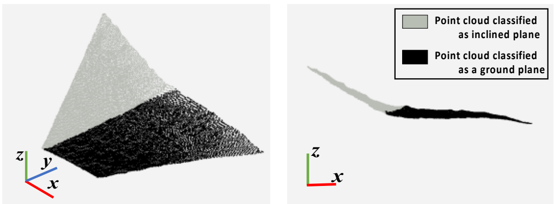


Fig. 4. Examples of created training (An inclined plane exists on the left)

4. Classification results by MLP

The classification accuracy was evaluated by IoU[(3)] (Intersection over Union), which takes a value from 0 to 100%. The larger this value is, the more correctly the input data is classified by MLP. Fig. 5 shows the values of IoU for each neighborhood radius r. The inclination angle of the plane classified as an inclined plane by MLP was calculated by least-squares plane fitting. If the inclination angle is more than 20 degrees, the inclination is impossible for the robot to climb. The distribution of the estimated inclination angle with respect to the neighboring radius r is shown in Figure 6. Figure 6 shows the distribution of the estimated inclination angles with respect to the nearest neighbor radius r. Figure 7 shows an example of the results of classifying the evaluation data by MLP.

5. Consideration

The IoU values in Fig. 5 do not appear to change significantly for any of the neighborhood radius r.
In this verification, however, a 1% increase in the IoU value indicates that about 1200 points were successfully classified. Fig. 7 shows that the number of points that were incorrectly classified as belonging to the ground plane decreased in the regions A and B in the classification results. The number of points incorrectly classified as belonging to the ground plane decreased. This suggests that the classification accuracy of the MLP created in this study may depend on the neighborhood radius r when the local normal vector is used as a feature of the 3D point cloud.

As shown in Fig. 6, the inclination angles estimated by the point clouds classified as inclined surfaces were distributed in the range of 18 to 31 degrees. Of these, 32 out of the 45 data for evaluation were estimated to be more than 20 degrees, suggesting that it is possible to detect unclimbable slopes with this classifier.

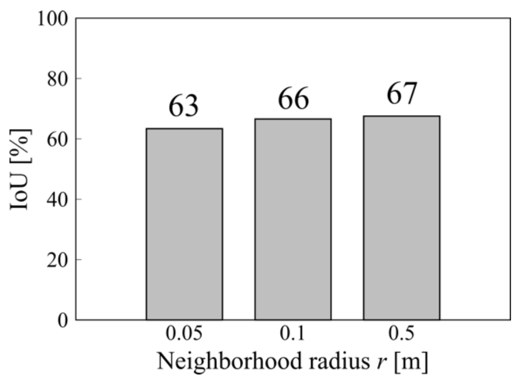


Fig. 5. IoU calculated from classification results

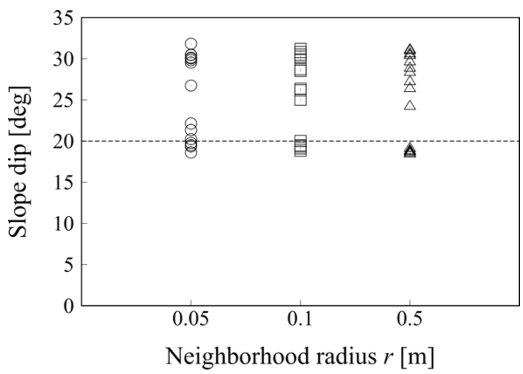
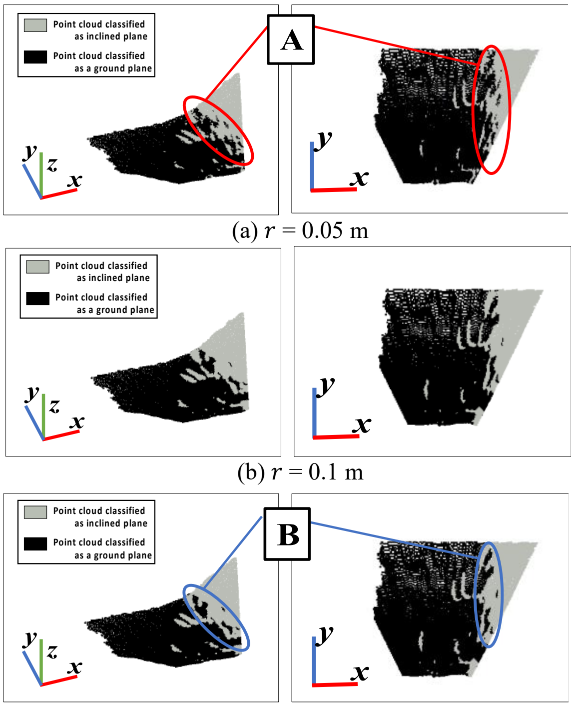


Fig. 6. Distribution of slope angle estimated from point clouds classified as inclined planes



(a) r = 0.05 m

(b) r = 0.1 m

(c) r = 0.5 m

Fig. 7. Classification Results

6. summary

In this study, we created a classifier using MLP with each point of the point cloud as input and attempted to detect inclined surfaces that cannot be traveled on. As a result, it was suggested that the radius of the neighborhood of the local normal vector, which is the input vector of the classifier, can reduce the false detection of unevenness of the ground surface and that MLP can detect slopes of more than 20 degrees.

In the future, we will study the optimal feature values to further improve the classification accuracy. In addition, we will verify the real-time performance of the system, which is important when the robot moves autonomously in mountainous areas, and we will also study terrain recognition in other environments composed of sand, such as the coast.

References

1. R. Parker, K. Bayne , P. W. Clinton, “Robotics in forestry,” New Zealand Journal of Forestry, Volume 60, 4 号, pp. 8-14, 2016.
2. V. Nair and G. E. Hinton, "Rectified Linear Units Improverestricted Boltzmann Machines", In ICML, (2010), pp,807-81
3. Rezatofighi, H.; Tsoi, N.; Gwak, J.; Sadeghian, A.; Reid, I.; and Savarese, S. 2019. Generalized intersection over union: A metric and a loss for bounding box regression. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)

Authors Introduction

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.

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. 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.

Assoc. Prof. Ryusuke Fujisawa

Assoc. Prof. Prof. Ryusuke Fujisawa is an associate professor in the Department of Intelligent and Control Systems at Kyushu Institute of Technology. He received the Ph.D. (Dr. Eng.) degree from The University of Electro-Communications in 2009. His research interests include Swarm intelligence, Swarm robotics and Distributed autonomous system. He is a member of Japan Society of Civil Engineers (JSCE), Information Processing Society of Japan (IPSJ), The Society of Instrument and Control Engineers (SICE) and The Robotics Society of Japan (RSJ).

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).

Particle Filter Based SLAM 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

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-sylvain@mail.kyutech.jp, m-san@mmcs.mse.kyutech.ac.jp, tominaga@kct.ac.jp, tomokawa.takumi163@mail.kyutech.jp, haya@.mse.kyutech.ac.jp,
http://www.kyutech.ac.jp/

Abstract

In Japan, the forestry workforce is dramatically declining. Therefore, field robots are investigated to replace humans for dangerous actions. Task execution with such mobile robots requires localization and mapping. This research focuses on online SLAM implemented on SOMA forestry robot developed at Hayashi Laboratory. In this approach, the core algorithm is a Rao-Blackwellized particle filter. A realistic simulation has been build using Gazebo and the results of first experiments speak for real-time capability.

Keywords: Field Robot, online SLAM, Forestry, Particle Filter

1. Introduction

Japanese forestry workers have a more and more advanced age. Consequently, because little youths want to apply for a job in this sector, the number of people in the field is shrinking. At the same time, the need of forest management is increasing. Owing to all this factors, in addition to hazardous nature of some forestry tasks, introducing robots on the field could be beneficial. One prototype for this purpose is currently designed at

[*]7 rue Peyron, Vienne (38200), France

Hayashi Laboratory. This robot, named SOMA,[1] has been built from an All-Terrain Vehicle (ATV) and many sensors like odometers, GPS (Global Positioning System), IMU (Inertial Measurement Unit), RGB-D cameras and lidar have been fixed to it.

Every mobile robot – including forestry ones – requires some navigation system to be able to execute its tasks and Simultaneous Localization and Mapping (SLAM) is an essential part of it. Unless most applications where full SLAM is first solved only for mapping and then localization is performed within the previously built map, this research is investigating online SLAM for forestry robots. This approach is motivated by the fact that building a map in advance demands time. Furthermore, it needs to be done every time the environment where the robot evolves changes. As a result, it enables computation savings only when the world stays identical for a long time. However, forests are constantly changing because of trees being planted, growing and being felled. In addition, the robot can be led to work in different areas. This point explains the choice of solving online SLAM for this specific application.

Among various algorithms available, FastSLAM was chosen because of its multimodal beliefs management ability, inherent to its use of a particle filter.[2]

This article begins with the description of the approach and how simulation experiments were conducted. Then, the obtained results are detailed and analyzed, before conclusion and opening on future research being drawn.

2. FastSLAM for Forestry Robot

The FastSLAM algorithm is based on a Rao-Blackwellized particle filter in which a particle represents a joint hypothesis of the pose of the robot and the map through the positions of landmarks.[2] The uncertainty of the map is represented by associating a Kalman filter to each feature in it. Thus, each particle contains a weight, a pose and a collection of Kalman filters depicting the environment.

Different parts of the algorithm need to be specified according to the forestry context. First, since trees are very common and easily distinguishable in a forest, there were chosen as landmarks to which the environment is mapped.

A particle filter being a variety of Bayes filter, it also requires motion and observation models to be implemented.

2.1. *Motion*

The motion model is based on the odometry captured by the rotary encoders mounted on the wheels of the robot. Each movement is decomposed in three parts: a first rotation, a translation and a second rotation, and linear gaussian noise is assumed for each of them.

2.2. *Observation*

With regard to observation, the sole sensor used is the lidar on top of the robot and thereby a range-bearing model was chosen. The coordinates of the trees are determined by extracting the centers of 3D clusters made from lidar pointcloud, after having clipped the latter in height to remove ground and foliage and keep only tree trunks (Fig. 1). Assuming linear gaussian noise for individual readings of the lidar, these coordinates can be proved to be also corrupted with gaussian noise.

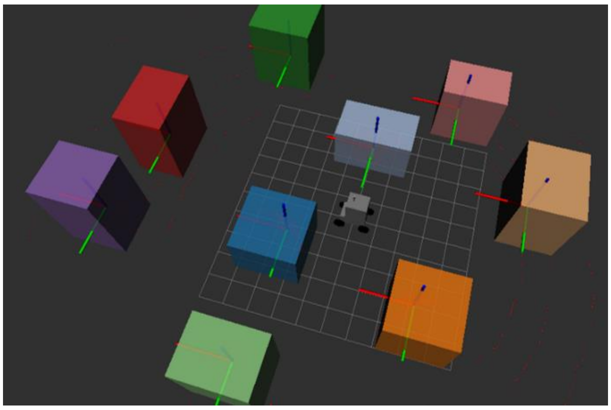


Fig. 1. Clusters made from lidar pointcloud

2.3. *Correspondences and map management*

Another important part of every SLAM solver is how to get correspondences between observations and features in the map being built. The largely used maximum likelihood approach was chosen. In FastSLAM, correspondences are determined per particle and not for the whole particle set, enabling some diversity of data association.

2.1.1. *Multiple observations*

Multiple observations are often split into individual ones and processed sequentially. However, it can lead to wrong fusions of landmarks, because two observations generated by two different landmarks can be assigned to the same feature in the map. On account of this issue, we here try to handle multiple observations at once thanks to the Gale-Shapley algorithm.[3]

The likelihood of each feature in the lidar visibility scope is computed for each observation and a list of features ordered by decreasing probability is made for the latter. Then, each observation is appaired with the first feature in its list. If more than one observation is linked with the same feature, only the observation with the highest likelihood keeps it and the second-ranked ones are attributed to the others. This step is repeated until each observation is appaired with a different feature. When the likelihood of an attributed feature is less than a threshold or when there are no more features in the list to continue the process, a new feature is created for the corresponding observation.

2.1.1. *Features deletion*

Features can also be subject to deletion when it appears that they do not correspond to a tree anymore. This action is realized along with correspondences establishment and the same threshold is used. Of course, it is not applied to features which are not considered to be in the lidar visibility scope.

2.4. *Particles handling*

Particle filters are often subject to what is named particle deprivation, that is the decrease of the diversity of the particle set over time. Therefore, resampling is used to counteract this drawback. In addition, resampled particles are slightly randomized by applying gaussian noise with the same order of magnitude than motion noise.

3. Experiments

For simulation purposes, a model of SOMA robot has been developed in Rviz and Gazebo using ROS, along with a realistic forest environment. The latter is made of a ground and 16 pine trees and measures 30 by 30 meters (Fig. 2).

Fig. 2. Robot model in Gazebo environment

In order to simulate motor noise, a ROS node was introduced between the steering controller and Gazebo. This node adds linear gaussian noise to linear and angular velocity commands before publishing them to a new topic to which Gazebo subscribes (Fig. 3).

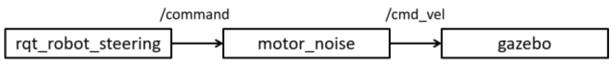


Fig. 3. Graph of motion nodes and topics

Different experiments have been conducted in order to determine the effect of the number of particles and the number of trees on the overall precision and update rate of the implementation. For each of them, the robot is going straight from the left to the right in the forest environment.

4. Discussion

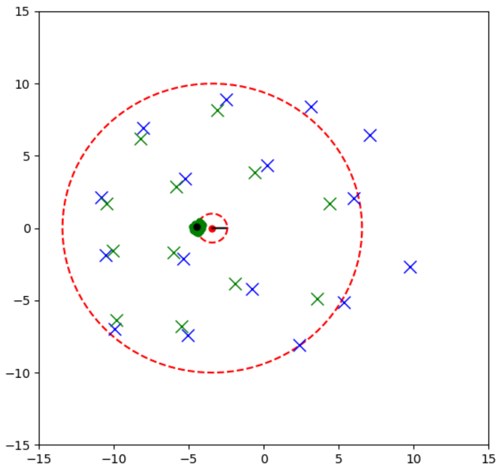


Fig. 4. Display of simulation (red point is real position, black line in front of red point is real heading, space between the two dotted red circles is lidar visibility scope, blue crosses are trees, green points are positions of particles, black point is position of the most probable particle, green crosses are features in the map of the most probable particle)

With realistic motion and observation noises and 100 particles, the distance between real and estimated poses at the end of the simulation is about 2.39 m, and the average distance between trees and associated features is 0.81 m (Fig. 4). Five trees over 16 are represented by multiple features and one feature does not correspond to any tree. In addition, the update rate in this case is about 1 Hz.

Overall, the robot is well tracked along its path. However, we can notice that the pose estimation is increasingly late compared to the real pose. This systematic shift can be explained by two factors. First, because this effect is observed even without motion noise, we can suppose odometry to be captured late. However, the main cause seems to be the imprecision of the computation of the coordinates of the trees. Indeed, the latter is based on the clusters made from lidar pointcloud. The centers of these do not coincide with the centers of trunks, because only a part of each tree is seen by the robot at one time. As a consequence, the center of each cluster is changing a lot as the robot is moving around the corresponding tree (Fig. 5).

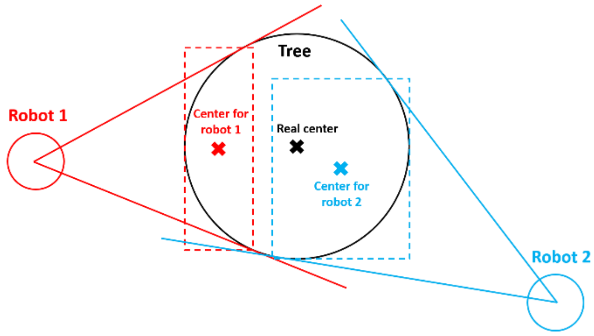


Fig. 5. Centers of clusters depending on the pose of the robot

Thus, depending on the ratio between motion and observation noises, the pose estimation can be overly updated according to this modification of the coordinates of the features, and induce a shift between real and estimated poses. This is why the use of circle pattern recognition should be used in each cluster to get the right coordinates of trees in future research.

4.1. *Effect of the number of particles*

The number of particles does not seem to have a very significant effect on the overall precision of the implementation. Even 5 particles are sufficient to get quite accurate final pose and map. As expected, the update rate decreases with the number of particles. Finally, adding more particles do not solve the systematic shift issue.

4.2. *Effect of the number of trees*

The closer the robot goes next to the trees, the more important the effect of systematic shift, because less of the tree is seen at one time. Then, the amplitude of change of the coordinates of the center of the cluster is larger in this case. Thereby, a higher density of trees leads to a lower precision. Moreover, as already known in the context of SLAM, more landmarks induce more wrong data association. Finally, the total number of trees does not affect the update rate of the implementation, because only features in the lidar visibility scope are updated. Nevertheless, the denser the forest, the lower the update rate.

4.3. *Future research*

As previously mentioned, a circle pattern recognition should be used to compute the coordinates of the trees from the lidar pointcloud, in order to remove the described systematic shift between the real and estimated poses.

In addition, resampled particles are here randomized to prevent particles deprivation, but more advanced methods exist such as mixture MCL, which is much more efficient in particular when motion noise dominates observation noise.

Finally, more experiments including real ones should be conducted to completely evaluate the approach described in this paper.

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. D. Gale, L. S. Shapley, “College Admissions and the Stability of Marriage”, The American Mathematical Monthly, Vol. 69, No. 1 (Jan., 1962), pp. 9-15

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.

Prof. Ayumu Tominaga

Prof. Ayumu Tominaga 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 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).

Anomaly Detection using Autoencoder with Gramian Angular Summation Field in Multivariate Time Series Data

Umaporn Yokkampon[1], Abbe Mowshowitz[2], Sakmongkon Chumkamon[1], Eiji Hayashi[1]
[1]Graduate School of Computer Science and Systems Engineering, Kyushu Institute of Technology, 680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan
[2]Department of Computer Science, The City College of New York, 160 Convent Avenue, New York, NY 10031, USA
E-mail: may@mmcs.mse.kyutech.ac.jp, amowshowitz@ccny.cuny.edu, m-san@mmcs.mse.kyutech.ac.jp, haya@mse.kyutech.ac.jp
www.kyutech.ac.jp

Abstract

Uncertainty is ubiquitous in data and constitutes a challenge in real-life data analysis applications. To deal with this challenge, we propose a novel method for detecting anomalies in time series data based on the Autoencoder method, which encodes a multivariate time series as images by means of the Gramian Angular Summation Field (GASF). Multivariate time series data is represented as 2D image data to enhance the performance of anomaly detection. The proposed method is validated with four time-series data sets. Experimental results show that our proposed method can improve validity and accuracy on all criteria. Therefore, effective anomaly detection in multivariate time series data can be achieved by combining the methods of Autoencoder and Gramian Angular Summation Field.

Keywords: Anomaly detection, Factory automation, Autoencoder, Multivariate time series.

1. Introduction

Anomaly detection involves discovering unusual patterns in time series data. Such detection is an important and challenging task, widely applied in various fields, such as credit card processing, medical diagnosis, sensor network operations, intrusion detection, and other areas.[1] Many algorithms have been developed to detect anomalies, but most of them are limited to analysis of univariate time series.

Recently, deep learning techniques have been applied to anomaly detection, the most common being Principal Component Analysis (PCA), which is a linear dimensionality reduction technique. PCA projects high-dimensional time series into a low-dimensional sequence. However, PCA is not flexible and cannot perform non-linear operations. A newer dimensionality reduction method is autoencoder[2,3], which has become a popular method for anomaly detection. Autoencoder is used to perform dimension reduction by stacking up layers to form a deep autoencoder. The hidden units are expected to extract features that represent the data faithfully by reducing the number of units in the hidden layer. Autoencoder judges whether something is abnormal by considering the difference between encoded data and the original.[4,5] Moreover, the autoencoder can perform both linear and non-linear operations.[6]

In this paper, we propose an anomaly detection method based on the autoencoder which encodes the multivariate time series data as 2D images designed to enhance performance. Our proposed method is validated by comparing the encoded results with the original data using four standard datasets. Four validation criteria have been used, namely, precision, recall, F1-score, and g-mean.

2. Proposed Method

This section introduces our proposed method for improving anomaly detection performance, and discusses the datasets and evaluation metrics used in this paper. Fig. 1 displays the overall structure of the proposed method. First, four time series data sets are converted into 2D images using Gramian Angular Summation Field (GASF), and then input into the autoencoder to identify anomalies by constructing the encoder and decoder. The results are then obtained from the autoencoder reconstruction values.

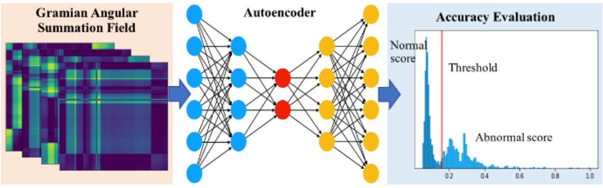


Fig. 1. Concept of the proposed method.

2.1. *Autoencoder*

Autoencoder is an unsupervised artificial neural network, which can leverage neural networks for the task of representation learning[7]. Autoencoder learns how to compress and encode data efficiently and then to reconstruct the data from the reduced encoded representation to one that is as close to the original input as possible. It consists of an encoder, a latent variable, and a decoder.

The encoder maps input data $x \in \mathbb{R}^d$ to a latent variable (code) $z \in \mathbb{R}^d$ and the decoder maps back from latent variable to input space. The autoencoder training procedure uses backpropagation to minimize the network's reconstruction loss. The loss function measures the differences between the original input and the consequent reconstruction. The loss function is defined as follows:

$$L(x,\hat{x}) = \|x-\hat{x}\|^2 \tag{1}$$

After the training procedure, the autoencoder uses the reconstruction error as the anomaly score. The data with high anomaly scores are considered anomalies because only the normal data are used to train the autoencoder. The autoencoder will reconstruct normal data very well while failing to do so when confronted with anomalous data.

2.2. *Gramian Angular Summation Field (GASF)*

A Gramian Angular Field(GAF)[8] is an image obtained from a 1-dimensional time series, representing some temporal correlation between each time point. There are two methods available: Gramian Angular Summation Field(GASF) and Gramian Angular Difference Field. In this work, Gramian Angular Summation Field (GASF) is used to transform our time series data into 2D images, using a polar coordinate system.

In the Gramian matrix, each element is the cosine of the summation of pairwise temporal values. Given a time series $X = (x_1, x_2, ..., x_n)$ of n observations, we rescale X so that all values fall in the interval [0, 1]:

The signal is warped in the transform domain. After this, each time point in polar coordinates is compared with every other point for temporal correlation.

$$\tilde{X}_0^i = \frac{x_i - \min(X)}{\max(X)-\min(X)} \tag{2}$$

At this point, we convert the obtained values into the polar coordinate system as follows:

$$\begin{cases} \phi = \arccos(\tilde{x}_i), \quad -1 \le \tilde{x}_i \le 1,\ \tilde{x}_i \in \tilde{X} \\ r = \dfrac{t_i}{N},\ t_i \in N \end{cases} \tag{3}$$

where t_i represents the time stamp and N is a constant factor used to regularize the polar coordinate system. After transforming the rescaled time series into the polar coordinate system, GASF can be written as follows:

$$\text{GASF} = \begin{bmatrix} \cos(\phi_1+\phi_1) & \cdots & \cos(\phi_1+\phi_n) \\ \cos(\phi_2+\phi_1) & \cdots & \cos(\phi_2+\phi_n) \\ \vdots & \ddots & \vdots \\ \cos(\phi_n+\phi_1) & \cdots & \cos(\phi_n+\phi_n) \end{bmatrix} \tag{4}$$

$$\text{GASF} = \tilde{X}' \cdot \tilde{X} - \sqrt{I-\tilde{X}^2}' \cdot \sqrt{I-\tilde{X}^2} \tag{5}$$

where I is the unit row vector.

3. Experiments

In this section, we first describe the details of the datasets[9] used in our experiments, and present the performance metrics for evaluating the performance of our system.

3.1. *Datasets*

To demonstrate the effectiveness of the proposed method, we conducted experiments on four datasets: Satellite, SonyAIBORobotSurface2, ItalyPowerDemand, and Wafer. Table 1. shows the details of the datasets. We use 80% of the normal data for training, and the remaining 20% of the normal data is used for testing purposes.

Table 1. Details of the datasets.

Datasets	Length	Number of instances	Anomaly Ratio
Satellite	36	6435	0.32
SonyAIBORobotSurface2	65	980	0.38
ItalyPowerDemand	24	1096	0.49
Wafer	152	7164	0.11

3.2. *Performance Evaluation*

To evaluate the detection performance of our proposed method, we employed four metrics, namely, Precision, Recall, F1-Score, and G-mean[10] which are defined as follows:

$$\text{Precision} = \frac{TP}{TP+FP} \tag{6}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{7}$$

$$\text{F1-Score} = 2 \times \frac{\text{Precision識 ecall}}{\text{Precision}+\text{Recall}} \tag{8}$$

$$G-mean = \sqrt{\frac{TP}{TP+FN} \times \frac{TN}{TN+FP}} \tag{9}$$

where TP is the correctly predicted anomaly, FP is the incorrectly predicted anomaly, TN is the correctly assigned normal, and FN is the incorrectly assigned normal.

4. Results and Discussion

In this section, we evaluate the proposed method for detecting anomalies in a time series data set using the Autoencoder method. First, we transform all time series data sets into 2D images using GASF. Fig. 2 shows the GASF images example, in which the image represents some temporal correlation between each time point. The result is from the wafer dataset.

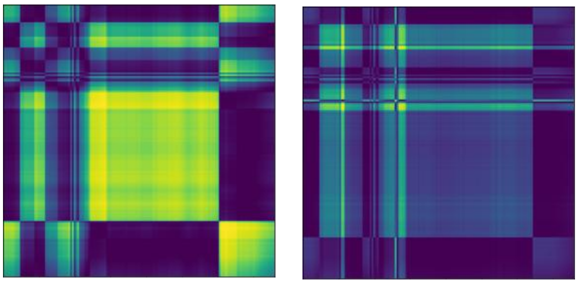


Fig. 2. The example of GASF images.

Our proposed method is trained with the normal samples, and then the model is verified using testing samples, including both normal and abnormal data. We conducted our experiment to improve the anomaly detection performance in four time series data sets using precision, recall, F1-score, and G-mean criteria. Table 2. shows the summary and comparison results of anomaly detection. The bold fonts in Table 2 indicate our proposed method could improve performance relative to relying on original data. For the Wafer dataset, the F1-Score value associated with the original data was 84.64%, whereas our method achieved 99.06%. In addition, the G-mean of the original data was only 51.61% compared with 99.48% for our proposed method. However, the SonyAIBORobotSurface2 and ItalyPowerDemand datasets of original data gave precision results better than our proposed method.

Table 2. Autoencoder compare of original data and our proposed method

Datasets	Original Data				Our proposed method			
	Precision	Recall	F1-Score	G-mean	Precision	Recall	F1-Score	G-mean
Satellite	0.7075	0.6779	0.7761	0.5074	0.7369	**0.9989**	0.8481	0.8325
SonyAIBORobotSurface2	0.9043	0.7647	0.8287	0.7903	0.8891	**0.9982**	0.9405	0.9171
ItalyPowerDemand	0.7091	0.5166	0.5977	0.5263	0.6818	0.7813	0.7282	**0.8559**
Wafer	0.7349	0.9979	0.8464	0.5161	0.9820	**0.9992**	0.9906	0.9948

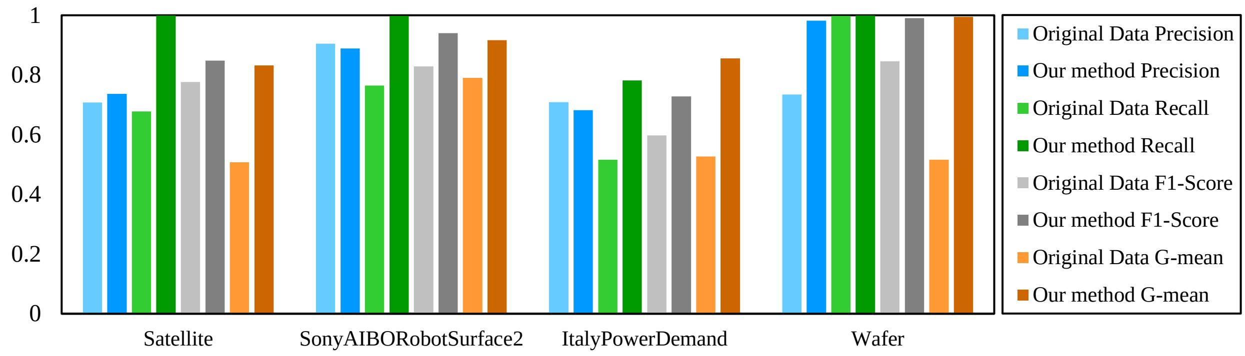


Fig. 3. All criteria comparisons of four datasets on Autoencoder

Fig. 3. shows the chart of all criteria comparisons of four datasets on Autoencoder. The light color line represents the original data, and the dark color line represents our proposed method. It is quite clear that our proposed method outperforms original data.

5. Conclusion

In this paper, we have presented an anomaly detection method using an autoencoder to compare the original data represented as GASF images (2D images). Performance has been evaluated by the metrics Precision, Recall, F1-Score, and G-mean. The experiments on anomaly detection show that our proposed method could improve the accuracy of detecting anomalies for time series data. Therefore, anomaly detection in multivariate time series, judged by the criteria of precision, recall, F1-score, and G-mean, can be improved by using Gramian Angular Summation Field (GASF) image encoding.

Acknowledgements

This work was supported by the Religion revitalization project by the Japanese Government.

References

1. V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey", *ACM computing surveys (CSUR)*, Vol. 41, No. 3, pp. 1-58, 2009.
2. A. Ng, "Sparse autoencoder", *CS294A Lecture notes*, Vol. 72, No. 2011, pp.1-19, 2011.
3. C. Zhou, and R.C. Paffenroth, "Anomaly detection with robust deep autoencoders", In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 665-674, August, 2017.
4. A. Borghesi, A. Bartolini, M. Lombardi, M. Milano, and L. Benini, "Anomaly detection using autoencoders in high performance computing systems", In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33, No. 01, pp. 9428-9433, July, 2019.
5. B. Zong, Q. Song, M.R. Min, W. Cheng, C. Lumezanu, D. Cho, and H. Chen, "Deep autoencoding gaussian mixture model for unsupervised anomaly detection", In *International conference on learning representations*, February, 2018.
6. M. Sakurada, and T. Yairi, "Anomaly detection using autoencoders with nonlinear dimensionality reduction", In *Proceedings of the MLSDA 2014 2nd workshop on machine learning for s ensory data analysis*, pp. 4-11, December, 2014.
7. Z. Chen, C.K. Yeo, B.S. Lee, and C.T. Lau, "Autoencoder-based network anomaly detection", In *Wireless Telecommunications Symposium (WTS)*, pp. 1-5, April, 2018.
8. Z. Wang, and T. Oates, "Imaging time-series to improve classification and imputation", In *Twenty-Fourth International Joint Conference on Artificial Intelligence*, June, 2015.
9. H.A. Dau, A. Bagnall, K. Kamgar, C.C.M. Yeh, Y. Zhu, S. Gharghabi, C.A. Ratanamahatana, and E. Keogh, 2019. "The UCR time series archive", *IEEE/CAA Journal of Automatica Sinica*, Vol. 6, No. 6, pp.1293-1305, 2019.
10. X. Hua, Y. Cheng, H. Wang, Y. Qin, and Y. Li, "Geometric means and medians with applications to target detection", *IET Signal Processing*, Vol. 11, No.6, pp.711-720, 2017.

Authors Introduction

Ms. Umaporn Yokkampon

Ms. Umaporn Yokkampon received Master of Science degree in Applied Statistics from King Mongkut's University of Technology North Bangkok in 2019. Presently she is a doctoral student in the Department of Computer Science and Systems Engineering at Kyushu Institute of Technology. Her research interests include data mining, anomaly detection and time series analysis.

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.

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).

Autonomous Robotics Packaging Ready Meal in Conveyor Production Line

Sakmongkon Chumkamon[1], Tomofumi Tsuji[1]
Prem Gamolped[1], Chanapol Piyavichyanon[2], Umaporn Yokkampon[1], Abbe Mowshowitz[3], Eiji Hayashi[1]
[1]Department of Mechanical Information Science and Technology, Kyushu Institute of Technology,
680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan
[2]Department of Creative Informatics, Kyushu Institute of Technology,
680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan
[3]Department of Computer Science, The City College of New York,
160 Convent Avenue, New York, NY 10031, USA

E-mail: gamolped.prem129@mail.kyutech.jp, m-san@mmcs.mse.kyutech.ac.jp,
piyavichayanon.chanapol402@mail.kyutech.jp, haya@mse.kyutech.ac.jp, amowshowitz@ccny.cuny.edu
http://www.kyutech.ac.jp/

Abstract

Food automation technology has become increasingly important in industrial and scientific research, especially in a social distance of pandemics. In this paper, we investigate the robotic motion planning and grasping for the food, which is non-rigid, nonpattern, and soft since the food is challenging to pick. Moreover, in automation, the process has to be organized on time since the automation uses the conveyor production line. This paper contributes three points: online motion planning while the production line conveyor operates. The second is real-time non-pattern food segmentation. The third is grasping non-rigid, nonpattern, and soft objects of food. Additionally, the robotics grasping also proposes the finishing decorating spaghetti packaging by rolling the spaghetti while the conveyor moving which is our novel proposed.

Keywords: Robotics, Pick-and-Place, Food Automation.

1. Introduction

Nowadays, technology has played an increasingly important role in facilitating human life and work. Technology development has been utilized rapidly and widespread in agriculture, medicine, and automation in the recent decade. That mainly focuses on developing renewable energy considering the environment and helping to reduce pollution problems. The industry has continually improved and set in terms of tools, equipment, and electrical appliances rapidly since the technology of Artificial intelligence, Machine Learning, and Robotics are researched and developed pervasively.

Recently, the robot has been widely and extensively applied to improve the quality, performance, and precision in a production factory. Additionally, one popular function of robots is assembling various parts. The robot would pick the component to assembly in this process, typically called pick-and-place. The critical point in this process is that the robot needs to be assigned to the position and orientation of the robot and the picked object since the robot would be programmed and instructed to move from pick position to place position.

The manipulator robot is controlled by joint position; however, we usually manipulate the robot in Cartesian space. Therefore, we calculate the inverse kinematic (IK) from the cartesian space to the joint space. The robot is

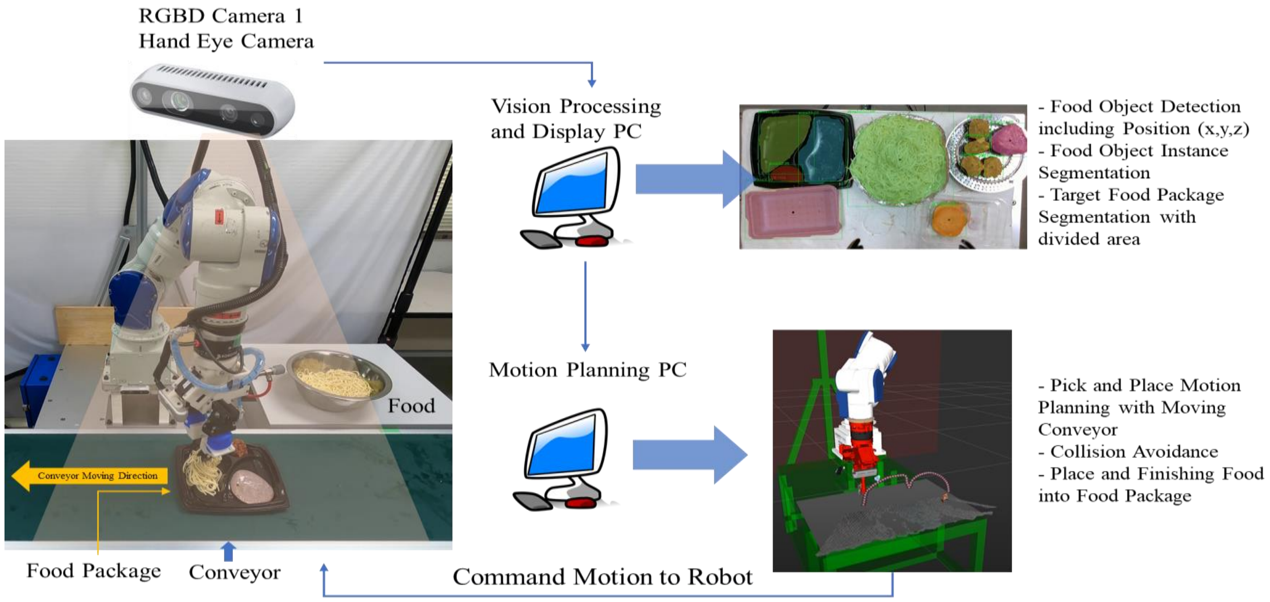


Fig. 1. The overview system of the proposed scheme includes a robot, RGBD camera, PCs, Foods and Conveyor

required to plan the trajectory according to the environment, such as avoiding the collision or moving to the target. The recent motion planning method is the Open Motion Planning Library, which is popular, reliable, and fast for solving the trajectory motion in various movements such as robot arm motion, vehicle, drone.

In this paper, we propose and contribute the robot automation for food packaging that can generate the robotics motion according to the food and moving container on the conveyor. Then we apply the planned move to the actual robot motion. Besides, this system uses Trac-IK for kinematic[1], and OMPL[2] for path planning based on ROS. Fig. 1 shows the overview of our proposed, which consists of the robot, food container on moving conveyor, and the food tray which contains the preparing food for packaging.

2. Approach

In this section, we explain the overview system, including robot hardware configuration, Conveyor, food instance segmentation, food package segmentation with the divided area, And the robot placing the food into the food package and finishing.

2.1. *Food Automation Setup with Conveyor*

Our proposed food automation setup consists of the robot, conveyor, food tray, food container package, and camera. The robot is used to operate the pick and place task and support the vision system since the camera is installed on the end-effector of the robot. The conveyor is utilized to move the food package to the robot then the robot would assembly the food into the container. For the vision system, we installed the RGBD camera into the robot to perceive the position of food in the tray and the food package on the conveyor. For the image processing system, we use one PC to operate the deep learning architecture to extract the segmentation and position of foods. Additionally, we use the second PC to manage the robotics motion planning for picking and placing the food into the package. The hardware system

Fig. 2. The example image of food and package instance segmentation includes three section areas of the package, spaghetti, Ham, chicken fries, rice ball.

configuration is shown in Fig 1.

2.2. *Foods and Package Instance Segmentation*

Firstly, we implement the Foods Instance Segmentation for the perception to detect and segment the food area and extract the position. We introduce the deep learning architecture of Cascade Mask R-CNN, implemented with PyTorch[3]. The backbone is ResNet-50. Our detection can detect the foods such as chicken fries, Ham, rice balls, spaghetti. We also implemented the food package divided into three section areas. We used this method to detect the food area in the color image after extracting the position using point cloud since we installed the robot's RGBD camera. The example of food and package section areas segmentations are shown in Fig.2.

2.3. *Robot Trajectory and Motion*

Since the robot knew the hand motion trajectory, we can then plan the motion according to the workspace constraint and the trajectory. Firstly, we set the home pose to initialize the robot joints to prepare a starting position to move following the trajectory. The robot will plan motion using the Open Motion Planning Library (OMPL) and Trac-IK inverse kinematic when the robot gets the new trajectory from the demonstration. In this implementation, we utilize the MoveIt tool with Robotics Operation System to organize motion planning and collision protection among the workspace environments. After the robot could detect the food and the package on the conveyor, the robot would move following the planned motion. Moreover, placing the food into the packing on the moving conveyor, we introduce the finishing process for spaghetti to allow the spaghetti not messy by making the robot roll the spaghetti while the conveyor moves.

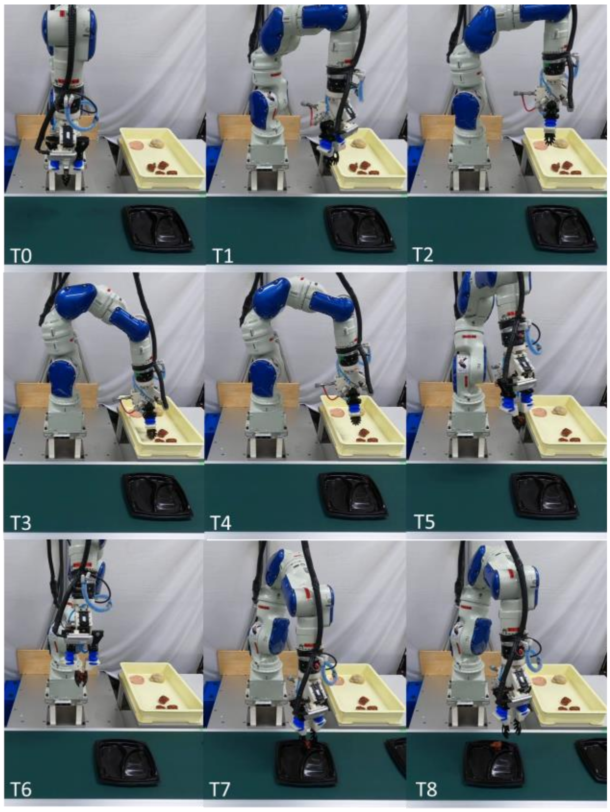


Fig. 3. The robotics motion of chicken fries pick and place with moving conveyor.

2.4. *Experiment Setup*

We use the industrial robot with 7 degrees of freedom and a soft gripper tool of end-effector for the hardware configuration system since we aim our robot to grasp various food. For robot control and motion planning system is implemented in the desktop PC. Additionally, we implement the image procession in another PC based on Deep learning Instance Segmentation. The robot is mounted on the robot base station with the camera mounted at the robot's end-effector in the experiment setup. We also calibrate the camera position translation to the robot position and orientation. Finally, the conveyor is set up in front of the robot.

3. Experiment and Results

The robot performs the food packing with a moving conveyor in this experiment. The robot would perceive the food and container position and then make a pick and place motion. After that, the robot will grasp the food into the package according to the speed of the conveyor.

First, the robot assembly the ready meal package by picking the chicken fries in the food package on the conveyor. Fig. 3 shows the robot assembling the chicken fries into the food package. Second, the robot assemblies ham to the food package, which is the same method of assembly chicken fries, except the robot would switch the tool by rotating the actuator at the end-effector to be a vacuum pad. The robot assembling Ham is represented in Fig. 4. Finally, the robot assemblies

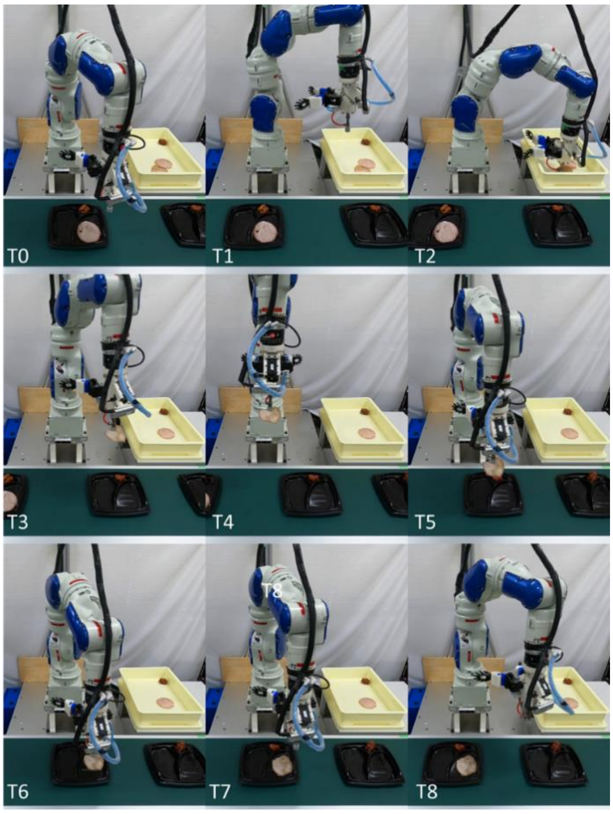


Fig. 4. The robotics motion of Ham pick-and-place with moving conveyor.

spaghetti, as shown in Fig. 5. However, in this process, the robot adds a finishing process by rotating the end-effector tool to roll the spaghetti, making spaghetti look in shape, as shown in Fig 4 from T8 to T9.

4. Conclusion

We introduce the robot food packaging in automation with the finishing process. We present and implement the robot framework of autonomous robot motion generation by perceiving the foods and the package moving on the conveyor using deep learning-based food and package instance segmentation based on Cascade Mask-RCNN. The robot could succeed in the food packaging especially assembling spaghetti with finishing process to keep it in shape while the conveyor moves.

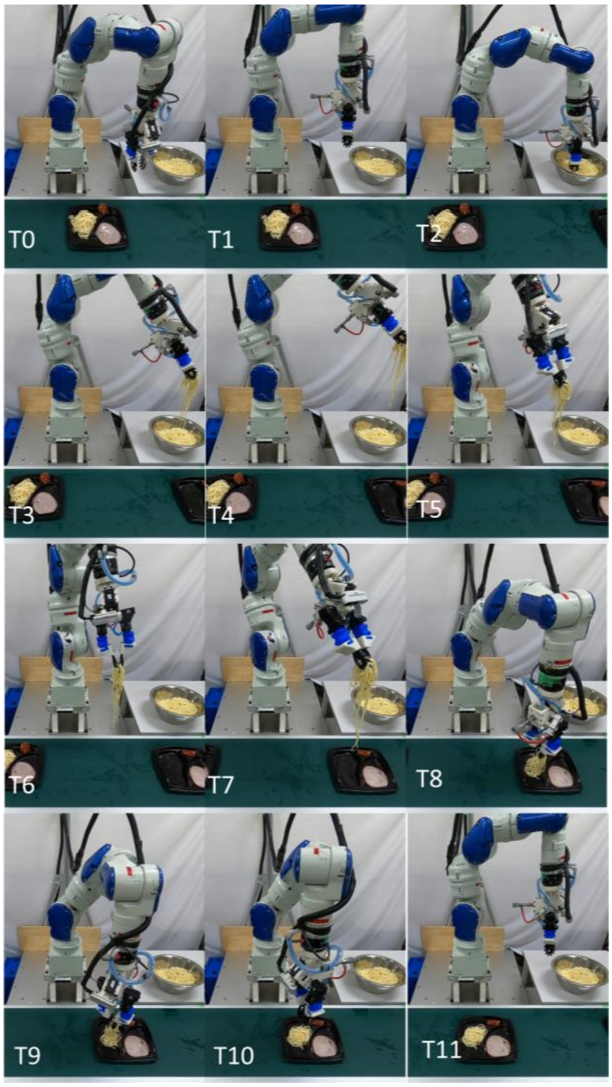


Fig. 5. The robotics motion of spaghetti pick and place with moving conveyor.

Acknowledgments

This research was partially supported by the Religion revitalization project in Kitakyushu by the Cabinet Office of Japan.

References

1. Beeson, Patrick, and Barrett Ames. TRAC-IK: An open-source library for improved solving of generic inverse kinematics. In 2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids), pp. 928-935. IEEE, 2015.

2. Sucan, I.A., Moll, M. and Kavraki, L.E., 2012. The open motion planning library. IEEE Robotics & Automation Magazine, 19(4), pp.72-82.
3. Cai, Zhaowei, and Nuno Vasconcelos. "Cascade r-cnn: Delving into high quality object detection." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 6154-6162. 2018.

Authors Introduction

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.

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.

Prem Gamolped

He received a bachelor degree in Engineering in 2021 from Robotics and Automation Engineering, King Mongkut's University of Technology North Bangkok in Thailand. He is currently a Master student at Kyushu Institute of Technology and researches at Hayashi Laboratory.

Mr. Chanapol Piyavichayanon

He received a bachelor degree in Engineering in 2021 from Mechanical Engineering, Chulalongkorn University in Japan. He is currently a Master student at Kyushu Institute of Technology and researches at Koga Laboratory.

Ms. Umaporn Yokkatmpon

Ms. Umaporn Yokkampon received Master of Science degree in Applied Statistics from King Mongkut's University of Technology North Bangkok in 2019. Presently she is a doctoral student in the Department of Computer Science and Systems Engineering at Kyushu Institute of Technology. Her research interests include data mining, anomaly detection and time series analysis.

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

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).

An Acoustic Artificial Life System Using the Game of Life and its Application for Performing Arts

Yasuhiro Suzuki
Graduate School of Informatics, Nagoya University, Furocho Chikusa
*Nagoya City, Aichi Prefecture 464/0814, Japan**

E-mail: ysuzuki@i.nagoya-u.ac.jp
www.ysuzuki.info

Abstract

Sound and vibration are forms of energy propagation. We have constructed an artificial life form that takes sound as energy and produces energy. The system converts sound into a two-dimensional pattern and uses it as input to the Game of Life. After n steps, the Game of Life is re-transformed into sound, and the sound is output. The sound allows the artificial life in the PC to interact with the outside world. We used this system to create an artwork that interacts with a dancer interactively through sound.

Keywords: *Butoh, The Game of Life, Sound, Deep Micro Vibrotactile, DMV*

1. Introduction

Tactile sense is the common language of Nature, so the tactile language should allow us to integrate the PC with Nature. We are trying to integrate the PC with Nature through haptic interaction.

If we want to integrate the PC with Nature, what do need to "transmit"? Unlike the Internet, the natural system is not the PC, and the Internet cannot directly connect the PC and the natural system. PCs, in general, do not have a sense of touch, so tactile interaction requires tactile sensors and output devices.

One way to connect the PC to Nature is through sound: the PC has a microphone and a speaker, and through the microphone, the PC can convert sound into electrical signals, process the information, and output the sound through the speaker.

What is sound? A sound is a form of energy. Energy is converted into waves, which travel through a medium. The propagation of the waves depends on the Nature of the catalyst. If we can sense the waves as they travel through the medium, we can convert them back into energy. In other words, by connecting a natural system to a PC via waves, energy can be exchanged between the natural system and the PC.

In the natural system, the energy is converted into waves, which travel through space and reach the PC's microphone. The energy is converted into an electrical signal in the microphone, processed in the PC and converted back into an audio signal. The audio signal is then sent out through the speakers, allowing the PC to transmit energy in waves to the natural world.

2. The Sound Creature

To connect Nature and PC via waves, we created the ***Sound creature*, SC**. This creature obtains energy from outside via waves and metabolizes them, and generate another wave. In other words, this artificial creature is an energy transformer. SC consists of a transformation from sound to two-dimensional patterns and vice versa and the game of life.

John Conway, who proposed the Life Game, strongly rejected the deterministic position in an invited talk at ALife conference, the international conference on artificial life in 2014. He emphasized contingency is much more critical than determinism; he said that "t*here is full of contingencies in the world*."

The universe of the Game of Life [1] is an infinite, two-dimensional orthogonal grid of square cells, each of which is in one of two possible states, alive or dead (or populated and unpolluted, respectively). Every cell interacts with its eight neighbours, which are the cells that are horizontally, vertically, or diagonally adjacent. At each step in time, the following transitions occur:

i)Any live cell with two or three live neighbours survives,
ii)Any dead cell with three live neighbours becomes a live cell,
iii)All other live cells die in the next generation. Similarly, all other dead cells stay dead.

The initial pattern constitutes the seed of the system. The first generation is created by applying the above rules simultaneously to every cell in the seed, live or dead; births and deaths coincide, and the discrete moment at which this happens is sometimes called a tick. Each generation is a pure function of the preceding one. The rules continue to be applied repeatedly to create further generations [1].

The game of life is entirely deterministic; there is no room for "contingency," So we integrate the game of life and the sound surrounding a PC, where the game of life is working.

The sound of volume (amplitude) converts the ambient sound of the PC into a two-dimensional pattern. Since the amplitude of the input sound is between -1 and 1, we divide this into N equal parts, where N is the sensitivity of the SC. The average amplitude of the input sound during Δt (about 0.01 seconds) is measured and converted into a Tactile score. Then, the input speech is converted into a two-dimensional pattern based on the Tactile Score at time t and $t + \Delta t$

This two-dimensional pattern is the volume change pattern of the input speech, and the game of life converts the two-dimensional pattern into other patterns. Recall the episode of the introduction of the Tactile Score. What was the difference between a cheerful good morning and a cheerless good morning? The difference was the way they said it, that is, the pattern of the volume change.

To change the pattern of the volume change is to change the way we say things. In other words, when the game of life converts a volume change into a two-dimensional pattern into a different volume change, it is equivalent to changing the wording.

The SC outputs the input environmental sound differently. In order to output sound, we need vocal cords. So Life game converted the 2 dimensional, 2D pattern into a Musical Instrument Digital Interface, MIDI code. This MIDI code was then used as input to a Digital Audio Workstation, DAW. We used Ableton Live as the DAW. We then generated "musical noise" from the continuous MIDI input.

2.1. *Garandoh no Niwa (empty garden)*

We worked with *Butoh* dancer *Norihito Ishii* (*Sankai jyuku* company) to create a dance piece using SC. Butoh is a form of Japanese dance theatre that encompasses a diverse range of activities, techniques and motivations for dance, performance, or movement [2].

A founder of Butoh, *Tatsumi Hijikata* explored the transmutation of the human body into other forms, such as those of animals. He also developed a poetic and surreal choreographic language, *butoh-fu* (*fu* means "score" in *Japanese),* to help the dancer transform into other states of being.

Butoh Score is a choreographic pair of counts and poetic statements. For example, it can be described as follows;

"A body lies in a heap of flesh, 40 counts.
The corpse comes to life. 20 counts.
......"

2.2. *A fusion of Butoh Score and Tactile Score*

We read and recorded the Butoh Score according to the count. The recorded sound was then subjected to a low-pass filter to extract very low-frequency sounds below 20Hz. We created a sound art by layering nature sounds recorded in the forest-based on the very low-frequency sound.

The length of this work follows the count of Butoh Score, so it is almost the same as the dance length. This sound art piece was played at the beginning of the dance. The Butoh Score was transformed into the Tactile Score via voice and fed into the Sound creature, thus creating a fusion of Butoh Score and Tactile Score. sound generated by vibrations of 20 times or less per second.

3. *Method*

We used two microphones, two PCs, loudspeakers and a projector. One microphone was placed on the floor to collect sound from the floor and output to the PC. Another one was on the floor to collect sound from the space and production to the PC (Fig.1).

Each PC runs the Sound creature; they chat with each other through sound. One PC outputs the audio of their

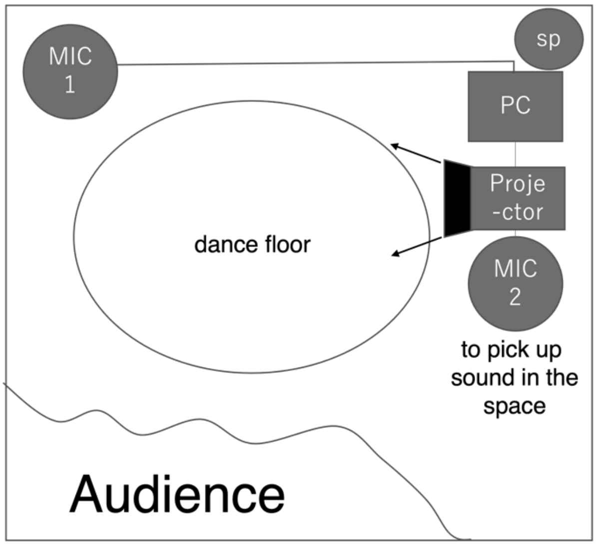


Fig. 1.Floor map of the performance

conversation to the audience, while the other PC outputs the process of creating sounds in the SCs. The process is a pattern of the game of life (Fig.2).

4. *Result*

We had a live performance with Norihito Ishii, Sound

Fig. 2. from live performance, left) patterns of process of sound creating / transforming of SC1 and SC2 right) Norihito Ishii (Butoh dancer).

creatures and the audience. Sound creatures run as autonomous computer programs with no operator in the performance (Fig 2).

References

1. https://www.conwaylife.com/wiki/Conway's_Game_of_Life
2. https://en.wikipedia.org/wiki/Butoh

Authors Introduction

Dr. Yasuhiro Suzuki

He is an Associate Professor of Graduate School of Informatics, Nagoya University Japan. graduated from Japan Advanced Institute of Science and Technology in 1995, received D. Informatics from Kyoto University in 2001.

The Effect of Non-audible Low Frequency, Deep Micro Vibrotactile, DMV Sounds on Music

Yasuhiro Suzuki
Graduate School of Informatics, Nagoya University, Furocho Chikusa
*Nagoya City, Aichi Prefecture 464/0814, Japan**

E-mail: ysuzuki@i.nagoya-u.ac.jp
www.ysuzuki.info

Abstract

Much of the research on low frequencies is on biological effects or noise pollution. We have confirmed that adding low frequencies to music increases autocorrelation and enhances the higher frequency components. This paper ensures that the addition of low frequencies to music produces 1/f fluctuations.

Keywords: deep micro vibrotactile, DMV, low-frequency sound, 1/f fluctuation

1. *Introduction*

Infrasound is a sound generated by vibrations of 20 times or less per second. In general, the sound loses energy owing to vibration as it passes through a medium. Therefore, higher frequencies rapidly lose power and are attenuated, whereas lower frequencies are less attenuated. Consequently, it can travel thousands of kilometres. Low frequency sounds are less directional than high frequency sounds; they proceed in concentric circles from the source, and go around any obstacles [1)]. Low frequency sounds have been examined as noise pollution, and their effects on the human body have been studied [3)]. However, a unified view has not yet culminated as the sensitivity to very low frequency sounds can vary from person to person [2)].

Very low frequencies have long been regarded as noise pollution. However, in nature, low frequency sounds are produced by a variety of natural phenomena, including tectonic movements, such as, volcanoes [5)], celestial movements, such as, asteroids, meteorites, meteors, and fireballs entering the atmosphere [6)], meteorological phenomena, such as, typhoons and lightning [7)], and water movement such as waterfalls and rivers [8)].

In the music industry, inaudible low frequencies below 20 Hz generate vibrations and distort the sound in the audible range. Hence, infrasound is removed from music during production. However, low frequency sounds are not completely eliminated, and may be important; hence, they are coordinated not to overload.

We investigated the case of a 40 Hz sine wave mixed with music [9]. It was confirmed that autocorrelation structures arise when the amplitude of the 40 Hz sine wave is large. When the high-frequency component of the music is small and the amplitude of the mixed 40 Hz sine wave is relatively large, the superposition of the waves strengthens the high-frequency component. This effect did not occur with white noise, which suggesting the effect of wave superposition.

In some natural phenomena or music, we can observe a power spectrum whose distribution is inversely proportional to the frequency. In case the slope of the regression line of the spectrum, λ, is around -1.0, it is called the 1/f fluctuation.

2. *Material*

The music was played by the attached speakers in a MacBook Pro laptop. Low frequency sounds were played using a Denon DSW37K subwoofer. A Zoom H6 handy recorder was used for recording. The sound source was Schuman's *Symphony No. 3 in E-Flat Major, Op. 97 "Rhenish": II. Scherzo,* performed by Sent Luis Symphony orchestra, (from *"Schumann: The 4 Symphonies"* Label: Vox). The low frequency sound of the sine wave was generated using Audacity version 3.0.2.

As the experiment was carried out in a room without soundproofing, the background noise of the room was examined. The frequency spectrum was measured without music and with a 40 Hz sinewave. The number of trials was three and the average intensity of was derived. An interval estimation, with a confidence interval of 95%, showed that the intensities at 40 Hz, 20 Hz, and 16.5 Hz were significantly different from the background noise of the room.

3. *Method*

A comparison was drawn between music from speakers and low frequencies of sine waves from a subwoofer both playing at the same time. The speakers and subwoofer were placed in the same position at a distance of 5.45m from each other during the recording. Owing to the influence of external noise, three measurements were taken: one with music solely and one with music and 40 Hz of sine wave played simultaneously.

We compared the power spectrum of a sound source and a sound source mixed with a 40Hz pure tone. We transform the obtained power spectrum to the logarithmic axis. The envelopes of the power spectra were linearly regressed. The slope of the linear regression was set to λ.

4. *Result*

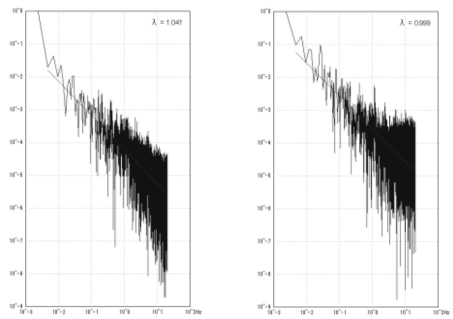


Fig. 1. right) Power spectrum of the sound source. The slope of the regression line, λ=1.041. left) 40 Hz sinusoidal pure tone added to the original. λ=0.999.

We measured the power spectrum of the original sound source and performed a regression analysis of the distribution. The slope of the regression line, λ = -1.041. we also added a pure tone sine wave of 40 Hz to the mixdown. The slope of the regression line of the power spectrum, λ = -0.999. The amplitude of the added sine wave is 0.1; as the amplitude is increased, λ deviates from -1.0.

5. *Discussion*

The role of low frequency sounds in music was studied. We confirmed that low frequency sound enhances the autocorrelative structure of sound. In this study it induces 1/f fluctuation in music.

Acknowledgement

We thank the reviewers for their valuable suggestions. We owe a debt of gratitude to Dr. Annette Grathoff for her thought-provoking discussions.

This study was supported by the Grant in Aid for Scientific Research No. 21K12108.

References

1. Araújo Alves, J., Neto Paiva, F., Torres Silva, L., & Remoaldo, P. (2020). "Low-frequency noise and its main

effects on human health—A review of the literature between 2016 and 2019." *Applied Sciences*, 10(15), 5205.

2. J. Yuei. (2004). "Outline of the 'Guide to dealing with infrasound problems' (in Japanese)" , *Journal of Enviromental Conservation Engineering*, vol. 33, 10, p. 753-757.
3. Koch, C. (2017). "Hearing beyond the limit: Measurement, perception, and impact of infrasound and ultrasonic noise." In *the 12th ICBEN Congress on Noise as a Public Health Problem, Zurich*.
4. Green, D. N., Neuberg, J. (2006). "Waveform classification of volcanic low-frequency earthquake swarms and their implications for Soufrière Hills Volcano, Montserrat." *Journal of Volcanology and Geothermal Research*, 153(1-2), 51-63.
5. Le Pichon, L. Ceranna, C. Pilger, P. Mialle, D. Brown, P. Herry, and N. Brachet. (2013). "The 2013 Russian fireball was the largest ever detected by the CTBTO infrasound sensors." *Geophysical Research Letters*, 40(14), 3732-3737.
6. Bedard, A. J. (2005). "Low-frequency atmospheric acoustic energy associated with vortices produced by thunderstorms", *Monthly weather review*, 133(1), 241-263. Retrieved on the 17th of July 2021
7. Feng, H. & Yichun, Yang & Chunchuzov, Igor & Teng, Pengxiao. (2014). "Study on infrasound from a water dam Acta Acustica united with Acustica." 100. 10.3813/AAA.918702.
8. Tolonen, T., & Karjalainen, M. (2000). "A computationally efficient multipitch analysis model." *IEEE Transactions on Speech and Audio Processing*, 8(6), 708-716.
9. Morikawa, Masahiro. "Low-Frequency Characterization of Music Sounds--Ultra-Bass Richness from the Sound Wave Beats." arXiv preprint arXiv:2104.08872 (2021).

Authors Introduction

Dr. Yasuhiro Suzuki

He is an Associate Professor of Graduate School of Informatics, Nagoya University Japan. graduated from Japan Advanced Institute of Science and Technology in 1995, received D. Informatics from Kyoto University in 2001.

Response of Yeast to Low Frequency Sound Exposure

Yasuhiro Suzuki
Graduate School of Informatics, Nagoya University, Furocho Chikusa
Nagoya City, Aichi Prefecture 464/0814, Japan[*]

E-mail: ysuzuki@nagoya-u.jp.com
www.ysuzuki.info

Abstract

The response of cells and organisms to exposure to ultrasound and low frequencies has been investigated. The present study investigates the increased number of colonies when physically vibrated by a 40 Hz sine wave and exposed as sound. The number of colonies decreased when exposed to vibration compared to silence. When exposed to the sound, the number of colonies increased.

Keywords: Yeast, Infrasonic, Vibration, very-low frequency

1. Introduction

Infrasound is a sound generated by vibrations of 20 times or less per second. In general, the sound loses energy due to vibration as it passes through the medium. Therefore, high frequencies quickly lose power and are attenuated, while low frequencies are less attenuated. Consequently, it can travel up to thousands of kilometres. Low frequencies are less directional than high frequencies; it proceeds in concentric circles from the source and even if there are obstacles, go around 1).

There is low-frequency sound in nature. Low-frequency sound in nature is produced by a variety of natural phenomena, including tectonic movements such as volcanoes [5)], celestial movements such as asteroids, meteorites, meteors and fireballs entering the atmosphere [6)], meteorological phenomena such as typhoons and lightning [7)], and water flows such as waterfalls and rivers [8)]. Low-frequency sound in nature can be so powerful that it can break windows and damage buildings.

It has reported that sound can influence organisms, that do not possess auditory pathways to detect 2). And vibrations below 20Hz (infrasound) and above 20,000Hz (ultrasound), significantly affect the growth and development of microorganisms [2)-5)].

In the brewer's yeast, *Saccharomyces cerevisiae*, it has reported that sound alter growth and fermentation, significantly [6]. In this study, "*pure tone of low-frequency (100 Hz), and one of high-frequency (10 kHz), both at the same intensity (90 dB at 20 µPa), measured at the location of the fermentation tubes.* [6]"

As is well known in laboratory work, as far as we know, it has not investigated precisely, shaking speed affects the culture rate when shaking E. coli or yeast. The faster the shaking rate, the quicker the culture rate.

2. Result

We have investigated the vibrational and acoustic effects of 40 Hz pure tones. A pure tone is a sinusoidal wave containing no overtones other than 40Hz.

We measured the number of colonies after 24 hours of preincubation and after 24 and 48 hours of incubation.

2.1. Responses to Vibrations

A sinusoidal wave of 40 Hz vibrated the culture petri dish; amplitudes were around 0.2 mm (strong) and approximately 0.1 mm (weak). In each amplitude, the number of colonies decreased compared to silence. The number of colonies decreased as the amplitude increased.

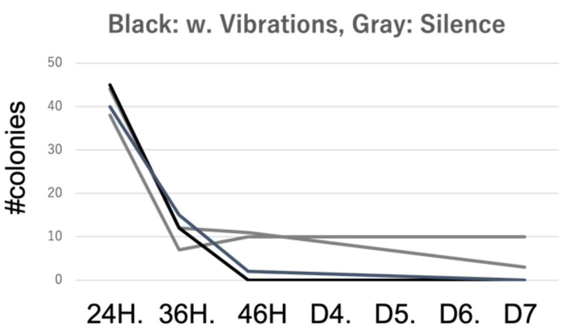


Fig. 1. Comparison of increment of colony numbers: vibrations with 40hz solid sine wave and silence. The horizontal axis is the time in units of 12 hours, and the unit after 48 hours is a day. The vertical axis is the number of colonies increased (increment from the previous time point). The black and grey lines represent the oscillatory and silence stimuli, respectively

Visually, the colony size was more significant when the vibration was more substantial, and the colony number decreased (Fig.1).

This trend was also confirmed when the culture medium was diluted (Fig.2). As the solution was diluted, the decrease in the rate of increase became slower. After 24 h, the number of colonies was higher than that of Silence. 48 h later, Silence continued to increase, while the increase was zero with the addition of vibration.

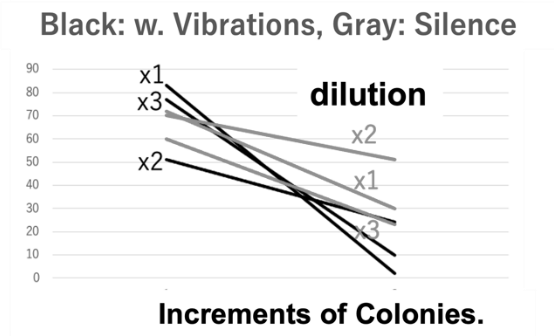


Fig. 2. Comparison of the incremental colony number between 40Hz sine wave intense vibration stimulation and silence in diluted culture medium: The horizontal axis represents 24h and 48h, the vertical axis the number of colonies increased. The black and grey lines represent the oscillatory and silence stimuli, respectively.

2.2. Responses to Sound

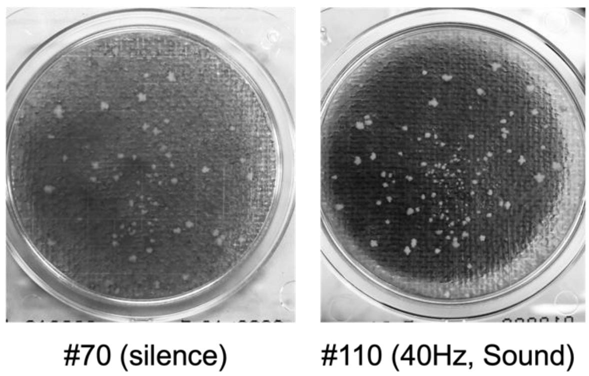


Fig. 3. Left) Silence, right) A typical example of a 50 dB sine wave at 40 Hz exposed for 48 Hours. The number of colonies is high, but the size of the colonies is small.

We exposed 40Hz sinewave in 50 dB to the yeast. And found that the increment number of colonies are more significant than silence and vibrations. We could observe that the size of each colony is smaller (Fig.3).

3. Discussion

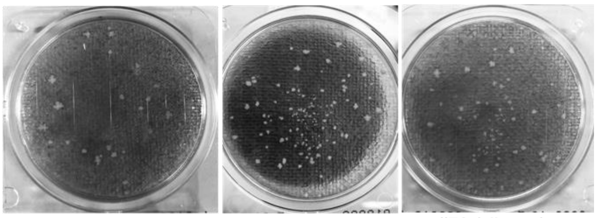


Fig. 4. Typical example of Yest responses: From right to left: Silence, Acoustic stimulus, 40Hz Sinewave with -24 dB FS (weak) and Vibration 40Hz oscillation with acceleration 0.04 m/s^2 , velocity 0.4 m/s, displacement 1mm. The number of colonies in Silence was 70, in 40Hz acoustic stimulation 110, in vibration stimulation 28.

The same 40 Hz sine wave behaved differently when the yeast was exposed to vibration and when it was exposed to sound (Fig.4).

In the case of vibration, the aggregation may be caused by physical vibration. Therefore, it is difficult to discuss the difference only by the number of colonies. Consequently, we cannot conclude an intrinsic difference between sound and vibration.

In this study, commercial baker's yeast was used. Many studies have used yeast with a distinct strain. It is common to pre-culture with shaking for as low as 24 hours in other studies. The yeast was pre-cultured in a thermostatic bath for 24 hours without shaking in this study.

In agar experiments, the number of yeasts is estimated from the number of colonies per square centimetre. In this study we counted the number of colonies.

4. Method and Material

Mix 0.4g of "*Shirakami Kodama*" yeast and 100g of pure water. Drop 0.1ml into the medium and pre-culture for 24 hours in a thermostatic bath at 30°C. YM medium is used (*Nissui* Pharmaceutical, Compact dry, YMR).

We prepare the 40Hz sine wave with Audacity version 3.0.2. Vibrations were generated in a thermostatic chamber at 30°C using a vibration vibrator. The vibrometer was adjusted to an acceleration of 0.04 m/s2, a velocity of 0.3 m/s and a displacement of 1 mm.

The control sample is placed in a thermostatic bath at 30°C. A vibrometer was used to check that the 40 Hz vibration for the vibration stimulus was not transmitted.

The exposure sound stimulus was provided by a subwoofer placed in the 30°C chambers, and the sound level was adjusted to 50 dB using a sound level meter.

Control samples were placed in a 30°C chamber. We confirmed that the 40 Hz tone for the sound stimulus was not transmitted.

Acknowledgement

This work is supported by Grant-in-Aid for Scientific Research No. 21K12108

References

1. Alfred J.Lotoka, Elements of Physical Biology, Williams and Wilkins Company, 1925
2. Chisti, Y. Sonobioreactors: Using ultrasound for enhanced microbial productivity. Trends Biotechnol. 2003, 21, 89-93.
3. Koda, S.; Miyamoto, M.; Toma, M.; Matsuoka, T.; Maebayashi, M. Inactivation of Escherichia coli and Streptococcus mutans by ultrasound at 500kHz. Ultrason. Sonochem. 2009, 16, 655-659.
4. Pitt, W.G.; Ross, S.A. Ultrasound Increases the Rate of Bacterial Cell Growth. Biotechnol. Prog. 2003, 19, 1038-1044.
5. Dang, W.-M.; Wang, S.; Tian, S.-X.; Chen, B.; Sun, F.; Li, W.; Jiao, Y.; He, L.-H. Effects of infrasound on activities of 3beta hydroxysteroid dehydrogenase and acid phosphatase of polygonal cells in adrenal cortex zona fasciculate in mice. Chin. J. Ind.
6. Araújo Alves, J., Neto Paiva, F., Torres Silva, L., & Remoaldo, P. (2020). Low-frequency noise and Its main effects on human health—A review of the literature between 2016 and 2019. Applied Sciences, 10(15), 5205.
7. J. Yuei, "Outline of the 'Guide to dealing with infrasound problems' (in Japanese,「低周波音問題対応の手引書」の概要）,Journal of Enviromental Conservation Engineering, (2004), vol. 33, 10, p. 753-757, Online ISSN 1882-8590, Print ISSN 0388 9459, https://doi.org/10.5956/jriet.33.753, https://www.jstage.jst.go.jp/article/jriet1972/33/10/33_10_753/_article/-char/ja

8. Koch, C. (2017). Hearing beyond the limit: Measurement, perception and impact of infrasound and ultrasonic noise. In *12th ICBEN Congress on Noise as a Public Health Problem, Zurich*.
9. Green, D. N., & Neuberg, J. (2006). Waveform classification of volcanic low-frequency earthquake swarms and its implication at Soufrière Hills Volcano, Montserrat. Journal of Volcanology and Geothermal Research, 153(1-2), 51-63.
10. Le Pichon, A., Ceranna, L., Pilger, C., Mialle, P., Brown, D., Herry, P., & Brachet, N. (2013). The 2013 Russian fireball largest ever detected by CTBTO infrasound sensors. Geophysical Research Letters, 40(14), 3732-3737.
11. Bedard, A. J. (2005). Low-Frequency Atmospheric Acoustic Energy Associated with Vortices Produced by Thunderstorms, Monthly Weather Review, 133(1), 241-263. Retrieved Jul 17, 2021.
12. Feng, H. & Yichun, Yang & Chunchuzov, Igor & Teng, Pengxiao. (2014). Study on Infrasound From a Water Dam. Acta Acustica united with Acustica. 100. 10.3813/AAA.918702.
13. Rob Toulson, Can We Fix It? – The Consequences Of 'Fixing It In The Mix' With Common Equalisation Techniques Are Scientifically Evaluated, Journal on the Art of Record Production, Issue 03, (2008).
14. Metallica Talk Top-Secret Distilling Process Behind New 'Blackened' Whiskey, Rolling Stone, SEPTEMBER 15, 2018

Authors Introduction

Dr. Yasuhiro Suzuki

He is an Associate Professor of Graduate School of Informatics, Nagoya University Japan. graduated from Japan Advanced Institute of Science and Technology in 1995, received D. Informatics from Kyoto University in 2001.

INT8 Activation Ternary or Binary Weights Networks

Ninnart Fuengfusin
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka, 808-0196, Japan
Hakaru Tamukoh
Research Center for Neuromorphic AI Hardware,
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka, 808-0196, Japan
E-mail: fuengfusin.ninnart553@mail.kyutech.jp, tamukoh@brain.kyutech.jp*
https://www.brain.kyutech.ac.jp/~tamukoh/

Abstract

We propose binary or ternary weights with 8-bit integer activation convolutional neural networks. This model is designed as a middle ground between 8-bit integer and 1-bit or 2-bit quantized models. We discover that conventional 1-bit or 2-bit only-weight quantization methods (i.e., BinaryConnect and Ternary weights network) can be utilized jointly with 8-bit integer activation quantization without significant fractions. Based on these two methods, we evaluate our model with a VGG16-like model and CIFAR10 dataset. Our model provides competitive results to a conventional floating-point model.

.*Keywords*: Quantization, Image Recognition, Model Compression

1. Introduction

With the invention of deep learning, the neural network (NN) has achieved a better performance than human experts, especially in the image recognition task [1]. However, to perform well, the NN must have a large number of parameters. Moreover, to deploy this model into mobile or edge devices in real-time processing is a challenge from several constraints: the device's low computational capacity, memory bandwidth, and others.

One of the methods to reduce the computational time is to convert the 32-bit floating-point (FP32) operations into easier to compute formats such as the fixed-point, integer, or other formats. This process is called quantization. Currently, the default quantization technique supported by major deep learning frameworks (i.e., PyTorch [2] and TensorFlow [3]) is an 8-bit integer (INT8) quantization [4]. In general, INT8 quantization transforms almost all FP32 parameters to INT8 via an affine transformation. The conversion allows the model to operate with only INT8 operations faster than FP32 operations. For instance, in NVIDIA Ampere architecture, NVIDIA A100 delivers around 32 times the number of INT8 operations per second compared to the number of FP32 operations that A100 can perform [5]. Furthermore, INT8 also reduces NN's overall memory footprint by factors of four compared with FP32.

To reduce bit-width to less than 8-bit, several researches show that 1 or 2-bit quantization is possible [6, 7] with some degree of loss in the model performance. In this lower-bit width quantization researches, researches can be categorized into quantizing only-weights and quantizing both weights and activation directions. In general, the quantized only-weight model performs better but consumes higher latency and hardware resources if implemented into the hardware.

We propose an INT8 Activation Ternary or Binary Weights Networks (ITBWN) model to lower bundles of the only-weight quantized model. ITBWN is based on a lower-bit width quantization method; however, we also utilize INT8 based quantization method to quantize the model's activations. This reduces overall hardware resource utilization and latency of only-weight quantized model. Furthermore, we show that both INT8 and lower-bit width quantization methods can be utilized jointly with competitive performance to floating-point model

2. Related Works

In this section, we cover two related work sections: INT8 quantization and lower-bit quantization related researches.

2.1. *INT8 Quantization*

INT8 quantization research problem is framed as how to map between floating-point to INT8 variables. In [4], a floating-point tensor is approximated using an affine transform with floating-point scaling factors S, 8-bit unsigned integer (UINT8) zero-points Z, and an INT8 tensor q, as shown in Eq. (1). A number variable of S and Z is either a number of channels of r (per-channel quantization) or a scalar (per-tensor quantization).

$$r = S(q - Z) \tag{1}$$

The optimal S and Z can be found by tracking statistical information (i.e., a minimum and maximum value) during the training or inference. To further reduce more complexity of this approximation, TQT [8] proposed to remove zero-point variables Z and affects Eq. (1) to become Eq. (2). In Eq. (2), TQT designs S as a scaler variable with the power-of-two quantization. This allows for replacing floating-point multiplication with shift-operations between S and q.

$$r = S(q) \tag{2}$$

In contrast, instead of using statistical information, TQT uses a training process with a straight-through estimator [9] to decide on S and t threshold values. t is used to clip minimum and maximum values of r before using Eq. (2). There are several TQT implementation. One of them Xilinx Brevitas [10] provides an easy-to-use PyTorch implementation of TQT.

2.2. *Lower-bit Quantization*

BinaryConnect (BC) [6] and Ternary Weight Networks, both (TWN) [7] are only-weight and lower-bit quantization methods. BC converts weights w into binary weights w^q using Eq. (3).

$$w^q = sign(w) \tag{3}$$

To make this BC trainable with Eq. (3) which discretizes the gradient to w, Eq. (4) transfers the gradient from quantized weights to floating-point weights. Eq. (4) makes Eq. (3) the same as the identity function during back-propagation.

$$\frac{\partial L}{\partial w} = \frac{\partial L}{\partial w^q} \tag{4}$$

TWN quantizes weights w into either $\{-S, 0, S\}$ with Eq. (5) where S and Δ are both positive floating-point variables. TWN applies Eq. (4) to make the model trainable with Eq. (5).

$$w^q = \begin{cases} S: w > \Delta \\ 0: |w| \leq \Delta \\ -S: w < -\Delta \end{cases} \tag{5}$$

Δ can be found as Eq. (6). Where $\boldsymbol{E}$ is an expected value or mean-average of $|w|$.

$$\Delta = 0.7 \times \boldsymbol{E}(|w|) \tag{6}$$

S can be found as Eq. (7). Eq. (7) can be summarized as mean-average of $|w|$ that have values more than Δ.

$$S = \boldsymbol{E}_{i \in \{i||w_i|>\Delta\}}(|w_i|) \tag{7}$$

3. INT8 Activation Ternary or Binary Weights Networks

In only-weight quantized model, the multiplication between floating-point activations and binary {-1, 1} or ternary weights {-1, 0, 1} can be done using only logic gates [11]. However, in [11], the accumulation of feature maps still requires floating-point accumulations. The motivation of ITBWN is to improve [11] by reducing the complexity of this floating-point accumulation. Converting all activation into INT8 reduces the floating-point accumulation into INT8 accumulation instead. A less complex INT8 datatype reduces overall the computational time. Since INT8 methods are based on approximating floating-point variables, well-have INT8 approximation should behave the same as the floating-point variable.

With this motivation in mind, ITBWN jointly utilizes methods from TQT, BC, and TWN together. ITBWN applies TWN or BC to quantize its weights; however, for its activations, ITBWN uses TQT. An overview of the datatypes in a block of ITBWN model is shown in Fig. 1.

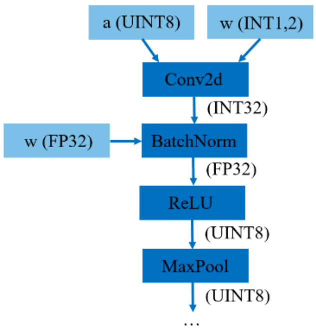


Fig. 1. Overview of datatypes in a block of ITBWN. Where *a* is the activation, *w* is weights or both weights and biases, UINT8 represents 8-bit unsigned-integer.

In this work, there are some modifications to BC, TWN, and TQT. For BC, we do not use clip functions to clip the weights to a range of [-1, 1]. For TWN, we did not use any scaling factors *S*. For TQT, we utilize a default implementation from Xilinx Brevitas, which has minor differences with the TQT setting [12]. Notable differences are Brevitas default settings do not use power-of-two quantization to the scaling factor, and Brevitas applies a different method to initialize the scale factor. At last, we use neither BC, TWN, nor TQT method in the last layer because it may affect the model's performance.

4. Experimental Results and Discussion

In this section, we conducted an experiment with the CIFAR10 dataset. We utilized a model based on VGG-16 [13]. To make VGG-16 operates with the CIFAR10 dataset, we removed the first two fully-connected layers and adjusted a number of input neurons in the last fully-connect layer to 512. We set the hyper-parameters of this experiment as shown in Table 1.

Table 1 Hyper parameters for VGG-16 in CIFAR10 setting.

Hyper Parameters	Value
Epoch	200
Batch size	256
Weight decay	0.0005
Learning rate	0.1

We preprocessed images with mean and standard deviation from each channel of RGB of training dataset. We applied the data augmentation as follows. During the training, the image was padded with zeros to its boundary and randomly cropped back to the original size. Each image was also randomly horizontally flipped. The stochastic gradient descent with momentum is selected to optimize our models. Finally, the learning rate is reduced with Cosine annealing [14].

We set notations as follows: *binary* for BC with TQT, *ternary* for TWN with TQT, *int8* for TQT only, and *float* for all FP32 floating-point model. The test accuracy over training epochs of each model is shown in Fig. 2. The best test accuracy of each model is shown in Table 2.

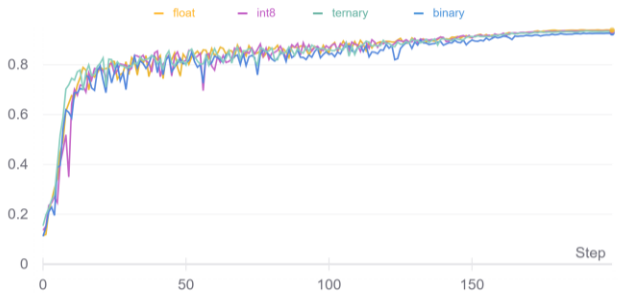


Fig. 2. Test accuracies of different methods per the training epoch.

Table 2 The best accuracy of VGG16 with different quantization methods.

Model	Test Accuracy
binary	0.9269
ternary	0.9351
int8	0.9353
float	0.9389

In this experiment, we found that *float* provides the best test accuracy; however *ternary* and *int8* also provide competitive results to *float*, which is quite surprising. In terms of *binary*, it gives the worst test accuracy. Further analysis of these outcomes, we scope them as future work.

5. Conclusion

We proposed ITBWN or a BC or TWN model with an INT8 quantized activations. Our experiment shows ITBWN with ternary weights provides the competitive result to both INT8 quantized and floating-point models. Converting floating-point activation into INT8 with TQT allows only-weight quantized model to deploy in hardware without worrying about the cost of floating-point accumulation.

For future work, we planned to further analyze why the performance between *ternary* and *int8* is so close to each other. We also would like to find why binary drops a high amount of test accuracy compared to ternary while only 1-bit differs.

Acknowledgements

This research is based on results obtained from a project, JPNP16007, commissioned by the New Energy and Industrial Technology Development Organization (NEDO).

References

1. 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*, pp. 770-778. 2016.
2. Paszke, Adam, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen et al. "Pytorch: An imperative style, high-performance deep learning library." *Advances in neural information processing systems* 32 (2019): 8026-8037.
3. Abadi, Martín. "TensorFlow: learning functions at scale." In *Proceedings of the 21st ACM SIGPLAN International Conference on Functional Programming*, pp. 1-1. 2016.
4. Jacob, Benoit, Skirmantas Kligys, Bo Chen, Menglong Zhu, Matthew Tang, Andrew Howard, Hartwig Adam, and Dmitry Kalenichenko. "Quantization and training of neural networks for efficient integer-arithmetic-only inference." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2704-2713. 2018.
5. NVIDIA A100 Tensor Core GPU Architecture, Accessed December 13, 2021, https://images.nvidia.com/aem-dam/en-zz/Solutions/data-center/nvidia-ampere-architecture-whitepaper.pdf.
6. Courbariaux, Matthieu, Yoshua Bengio, and Jean-Pierre David. "Binaryconnect: Training deep neural networks with binary weights during propagations." In *Advances in neural information processing systems*, pp. 3123-3131. 2015.
7. Li, Fengfu, Bo Zhang, and Bin Liu. "Ternary weight networks." *arXiv preprint arXiv:1605.04711* (2016).
8. Jain, Sambhav R., Albert Gural, Michael Wu, and Chris H. Dick. "Trained quantization thresholds for accurate and efficient fixed-point inference of deep neural networks." *arXiv preprint arXiv:1903.08066* (2019).
9. Bengio, Yoshua, Nicholas Léonard, and Aaron Courville. "Estimating or propagating gradients through stochastic neurons for conditional computation." *arXiv preprint arXiv:1308.3432* (2013)
10. Pappalardo, Alessandro. Xilinx/brevitas. Zenodo, 2021. https://doi.org/10.5281/zenodo.3333552.
11. Fuengfusin, Ninnart, and Hakaru Tamukoh. "Mixed-precision weights network for field-programmable gate array." *PloS one* 16, no. 5 (2021): e0251329.
12. Cite to quantization techniques in QuantIdentity(bit_width=8) and QuantReLU(bit_width=8) Accessed December 13, 2021, https://github.com/Xilinx/brevitas/issues/370.
13. Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.1556* (2014)
14. Loshchilov, Ilya, and Frank Hutter. "Sgdr: Stochastic gradient descent with warm restarts." *arXiv preprint arXiv:1608.03983* (2016).

Authors Introduction

Dr. Ninnart Fuengfusin

He received his B.Eng. degree from King Mongkut's University of Technology Thonburi, Thailand, in 2016. He received his M.Eng. and D.Eng degrees from Kyushu Institute of Technology, Japan, in 2018 and 2021, respectively. Currently, he is a post-doctoral researcher at the Kyushu Institute of Technology, Japan. His research interests include deep learning, efficient neural network design, and digital hardware design.

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.

A parameter tuning method for PQN model

Daimon Sakai
Department of Information Science and Technology, The University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

Takuya Nanami
Institute of Industrial Science, The University of Tokyo,
4-6-1, Komaba, Meguro-ku, Tokyo, 153-8505, Japan

Takashi Kohno
Institute of Industrial Science, The University of Tokyo,
4-6-1, Komaba, Meguro-ku, Tokyo, 153-8505, Japan
E-mail: daimon@iis.u-tokyo.ac.jp, nanami@iis.u-tokyo.ac.jp, kohno@sat.t.u-tokyo.ac.jp
https://www.neumis.iis.u-tokyo.ac.jp/

Abstract

The Piecewise Quadratic Neuron (PQN) model is a spiking neuron model that can be efficiently implemented on digital arithmetic circuits. In addition, this model can reproduce a variety if neuronal activities precisely with optimized parameter sets. In previous studies, we have optimized the parameters using meta-heuristic methods, which required a lot of computational time. In this paper, we proposed an parameter fitting method that takes into account the mathematical structure of the model and reproduces the electrophysiological activities of a target neuron with less computational time. We expect that this method can be used to construct silicon neuronal networks that faithfully replicate the nervous system. This method is expected to applicable to building silicon neuronal networks that faithfully replicate the nervous system.

Keywords: silicon neuronal network, spiking neuron model, PQN model, parameter fitting.

1. Introduction

The nervous system makes it possible for animals to process a variety of complex information. The silicon neuronal network (SNN) aims to achieve this information processing with low power consumption by mimicking the structure of the nervous system. The SNN is composed of the silicon neurons [1][2], which are digital or analog circuit units that simulates neuronal activities by solving the spiking neuron models. A variety of spiking neuron models have been used for the silicon neurons due to a trade-off between the reproducibility of neuronal activity and computational efficiency. For example, previous studies [3][4] used the ionic-conductance models. While these models can reproduce neuronal activities accurately, a large number of resources are required for the circuit implementation. On the other hand, integrate-and-fire (I&F)-based models are also widely used [5][6]. In I&F-based models, the variable corresponding to the membrane potential is reset to emulate the spike process when the neuron fires. This resetting has the advantage of low implementation cost, but makes the reproducibility of neuronal activities less than that of the ionic-conductance models.

We have proposed the Piecewise Quadratic Neuron (PQN) model, which is also known as the digital spiking silicon neuron models [7][8][9]. The feature of this model is that it uses piecewise quadratic functions. This model does not have a cubic term unlike other neuron models e.g. the FitzHugh-Nagumo [10] model and the Hindmarsh-Rose model [11], and can be implemented on the digital circuit with less circuit resources. In previous studies [12][13][14], the network containing the PQN model was constructed on a field-programmable gate array (FPGA) chip and the associative memory tasks were performed. In addition, the parameter fitting methods based on the meta-heuristic approach were proposed in previous studies [15][16], but they required a huge amount of computational time [17].

In this work, we propose the parameter fitting method focusing on the mathematical structure of the PQN model. And this method is applied to electrophysiological experimental data publicated in [18]. While the previous study [17] took several hours to determine the parameters, this method obtains the parameters within one minute.

The remainder of this paper is organized as follows. Section 2 explains about equations and the parameter fitting method for the PQN model. The results of parameter fitting are shown in Section 3 and the work is concluded in Section 4.

2. Methods

2.1. *PQN model*

Equations of the PQN model are given by

$$
\begin{aligned}
\frac{dv}{dt} &= \frac{\phi}{\tau}(f(v) - n - q + I_0 + I_{\mathrm{stim}}),\\
\frac{dn}{dt} &= \frac{1}{\tau}(g(v) - n),\\
f(v) &= \begin{cases} a_{fn}(v-b_{fn})^2 + c_{fn} & (v<0)\\ a_{fp}(v-b_{fp})^2 + c_{fp} & (v \geq 0)\end{cases},\\
g(v) &= \begin{cases} a_{gn}(v-b_{gn})^2 + c_{gn} & (v<r_g)\\ a_{gp}(v-b_{gp})^2 + c_{gp} & (v \geq r_g)\end{cases},\\
b_{fp} &= \frac{a_{fn}b_{fn}}{a_{fp}},\ c_{fp} = -\frac{a_{fn}^2 b_{fn}^2}{a_{fp}} + a_{fn}b_{fn}^2 + c_{fn},\\
b_{gp} &= -\frac{a_{gn}(-b_{gn}+r_g)}{a_{gp}} + r_g,\\
c_{gp} &= -\frac{a_{gn}^2(b_{gn}-r_g)^2}{a_{gp}} + a_{gn}(b_{fn}-r_g)^2 + c_{gn},
\end{aligned}
$$

where *v* is a state variable corresponding to the neuronal membrane potential. *n* is a recovery variable. Parameters φ and τ control time constants of variables. I_{stim} represents the input stimulus and parameter I_0 is a bias current. Parameters a_x, b_x, c_x and r_g, where *x* is *fp*, *fn*, *gp*, or *gn*, control the shapes of the *v*- and *n*-nullclines. Parameters b_{fp}, c_{fp}, b_{gp} and c_{gp} are determined so that the nullclines are continuous and smooth.

2.2. *Parameter fitting*

Figure 1 shows an example of neuronal activities simulated by the PQN model in response to a sustained current stimulus. In this study, for the efficient parameter fitting, we focus on these three features of the spike waveform: the distance from the minimum value to the threshold, the distance from the maximum value to the threshold, and the inter-spike interval. We defined the threshold as the value of the point at which the slope changes most rapidly before and after. The values of these three features are dependent on these three parameters, a_{fn}, ϕ, and I_0, respectively (Fig. 2). The a_{fn} controls a slope of the *v*-nullcline where *v* is less than 0. For example, when a_{fn} is increased, the slope becomes larger and the distance from the minimum value to the

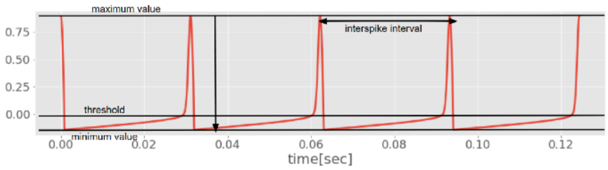


Figure 1: Typical spikes in the PQN model.

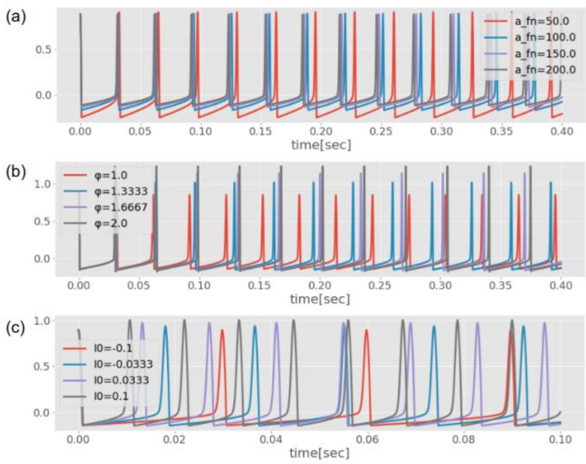


Figure 2: Effects of each parameter in the PQN model.

threshold becomes shorter (Fig. 2(a)). Note that when a_{fn} is changed, a_{gn}, b_{fn}, b_{gn}, c_{fn}, and c_{gn} are modified according to the following equation in order to minimize the effect on other features as much as possible.

$$a'_{fn} = ra_{fn},\ a'_{gn} = ra_{gn},\ b'_{fn} = b_{fn}/r,$$
$$b'_{gn} = r_g - (r_g - b_{gn})/r,\ c'_{fn} = a_{fn}b_{fn}^2 + c_{fn} - a'_{fn}b'^2_{fn},$$
$$c'_{gn} = a_{gn}(r_g - b_{gn})^2 + c_{gn} - a'_{gn}(r_g - b'_{gn})^2,$$

where r is the rate of change in a_{fn} and x', where x is a_{fn}, a_{gn}, b_{fn}, b_{gn}, c_{fn}, or c_{gn}, means the value of x after this modification. The distance from the maximum value tothe threshold is dependent on ϕ. Figure 2(b) shows that as ϕ is increased the amplitude is also increased. In addition, by increasing I_0 , the inter-spike interval becomes smaller (Fig. 2(c)).

In the process of the parameter fitting, we firstly tuned a_{fn} by comparing the difference from the minimum value to the threshold. Then, ϕ is tuned in order to adjust the distance from the maximum value to the threshold. Finally, the inter-spike interval is fit by changing I_0. We repeated this procedure three times, and totally required 55.4 seconds in average. In this experiment, we used the computer with Intel Core i7-8700 CPU. And cython library (version 0.29.24) was used in the source code.

3. Results

Figure 3 compares the initial waveform and the waveform simulated by obtained parameters. Table 1 shows the mean squared error (MSE) between electrophysiological data and simulated data for three neurons in [20]. MSE is given by: where T is the number of time

$$\mathrm{MSE} = \sum_{t=1}^{T}(V_{\mathrm{measured}}(t) - V_{\mathrm{PQN}}(t))^2,$$

steps. $V_{\mathrm{measured}}(t)$ and $V_{\mathrm{PQN}}(t)$ are values of measurement and simulated data at time t, respectively. In all three data, the value of MSE became smaller than before the parameter fitting.

4. Conclusion

In this work, we proposed the parameter fitting method focusing on the mathematical structure of the PQN model. And we used this method to reproduce the electrophysiological data of three different neurons. The results showed that the MSE were reduced through the fitting. In our future work, we will study how the parameters effect on the time from the start of the spike until the spike reaches the highest point and the time from the spike reaching its highest point until it reaches the minimum value. By fitting these times, MSE is expected to be more reduced.

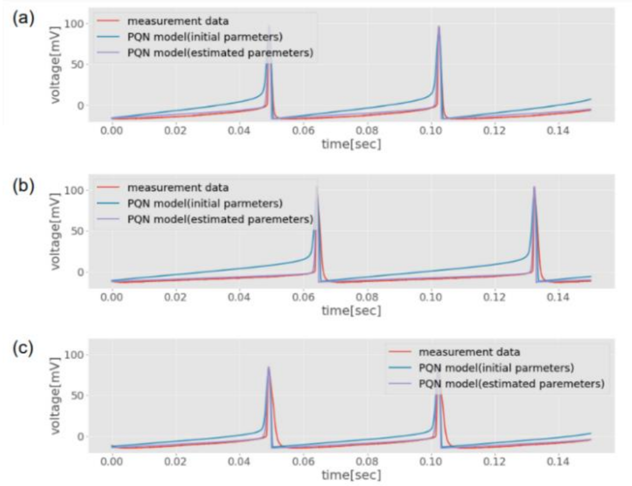


Figure 3: Results of parameter fitting. The specimen ids of the data (a), (b) and (c) are 613438337, 614635228, 614726150, respectively [20].

Table 1: Mean squared error.

specimen id	initial parameters	estimated parameters
613438337	0.006368	0.003469
614635228	0.009431	0.005571
614726150	0.011699	0.009618

5. Appendix

Table 2: initial parameters

a_{fp}	-2	b_{fp}	1	c_{fp}	1.48	r_g	0
a_{fn}	50	b_{fn}	-0.04	c_{fn}	-0.6	ϕ	0.8
a_{gp}	2	b_{gp}	-0.98	c_{gp}	-2.44	τ	0.001
a_{gn}	49	b_{gn}	-0.04	c_{gn}	-0.6	I_0	-

Table 3: estimated parameters

id	613438337	614635228	614726150
a_{fp}	-2	-2	-2
a_{fn}	121.13	238.99	140.87
a_{gp}	2	2	2
a_{gn}	118.71	234.21	138.05
b_{fn}	-0.016511	-0.0083686	-0.014198
b_{gn}	-0.016511	-0.0083686	-0.014198
c_{fn}	-0.55302	-0.53674	-0.54840
c_{gn}	-0.55396	-0.53800	-0.54943
ϕ	1.2637	1.4695	1.0491
I_0	-0.059050	-0.10996	-0.14165

References

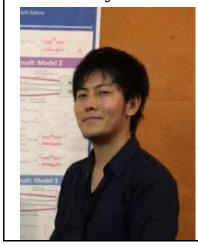
1. Misha Mahowald and Rodney Douglas. A silicon neuron. Nature, 354(6354):515–518, 1991.
2. Arindam Basu and Paul E Hasler. Nullclinebased design of a silicon neuron. IEEE Transactions on Circuits and Systems I: Regular Papers, 57(11):2938–2947, 2010.
3. Safa Yaghini Bonabi, Hassan Asgharian, Saeed Safari, and Majid Nili Ahmadabadi. Fpga implementation of a biological neural network based on the hodgkin-huxley neuron model. Frontiers in neuroscience, 8:379, 2014.
4. Xiaoyu Hu and Chongxin Liu. Dynamic property analysis and circuit implementation of simplified memristive hodgkin–huxley neuron model. Nonlinear Dynamics, 97(2):1721–1733, 2019.
5. Paul A Merolla, John V Arthur, Rodrigo Alvarez-Icaza, Andrew S Cassidy, Jun Sawada, Filipp Akopyan, Bryan L Jackson, Nabil Imam, Chen Guo, Yutaka Nakamura, et al. A million spiking-neuron integrated circuit with a scalable communication network and interface. Science, 345(6197):668–673, 2014.
6. Alexander D Rast, Francesco Galluppi, Xin Jin, and Steve B Furber. The leaky integrate-andfire neuron: A platform for synaptic model exploration on the spinnaker chip. In The 2010 International Joint Conference on Neural Networks (IJCNN), pages 1–8. IEEE, 2010.
7. Takashi Kohno and Kazuyuki Aihara. Digital spiking silicon neuron: concept and behaviors in gj-coupled network. In Proceedings of International Symposium on Artificial Life and Robotics, volume 2007, 2007.
8. Takuya Nanami and Takashi Kohno. Simple cortical and thalamic neuron models for digital arithmetic circuit implementation. Frontiers in neuroscience, 10:181, 2016.
9. Takuya Nanami, Azusa Kamikouchi, and Takashi Kohno. Modeling primary neurons in drosophila auditory system. In The 2018 International Symposium on Nonlinear Theory and Its Applications, pages 670–673.
10. Richard FitzHugh. Mathematical models of threshold phenomena in the nerve membrane. The bulletin of mathematical biophysics, 17(4):257–278, 1955.
11. James L Hindmarsh and RM Rose. A model of neuronal bursting using three coupled first order differential equations. Proceedings of the Royal society of London. Series B. Biological sciences, 221(1222):87–102, 1984.
12. Takashi Kohno, Munehisa Sekikawa, Jing Li, Takuya Nanami, and Kazuyuki Aihara. Qualitative-modeling-based silicon neurons and their networks. Frontiers in neuroscience, 10:273, 2016.
13. Takano Naruaki and Kohno Takashi. Towards modeling cholinergic modulation for neuromorphic computing. 人工生命とロボットに関する国際会議予稿集 25:159–162, 2020.
14. Jing Li, Yuichi Katori, and Takashi Kohno. An fpga-based silicon neuronal network with selectable excitability silicon neurons. Frontiers in neuroscience, 6:183, 2012.
15. Takuya Nanami, Filippo Grassia, and Takashi Kohno. A metaheuristic approach for parameter fitting in digital spiking silicon neuron model. Journal of Robotics, Networking and Artificial Life, 5(1):32–36, 2018.
16. Takuya Nanami and Takashi Kohno. An fpgabased cortical and thalamic silicon neuronal network. J. Robotics Netw. Artif. Life, 2(4):238–242, 2016.
17. Takuya Nanami. Qualitative-modeling-based digital silicon neuronal network models. PhD thesis, The University of Tokyo, 3 2019.
18. The Allen Mouse and Human Brain Atlases. https://celltypes.brainmap.org/data.

Authors Introduction

Mr. Daimon Sakai

He received his bachelor's degree from the Department of Mathematical Engineering and Information Physics, University of Tokyo, Japan in 2020. He is currently a Master course student in University of Tokyo, Japan.

Dr. Takuya Namami



He is a research associate in the Institute of Industrial Science at the University of Tokyo. He received the PhD degree in Department of Electrical Engineering and Information Systems in 2019 from the University of Tokyo,

Dr. Takashi Kohno

Dr. Kohno 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, Japan.

Hardware Development of Edge-Preserving Bubble Image Conversion in High-level Synthesis

Jiang Qin

Akira Yamawaki

Kyushu Institute of Technology
1-1 Sensui, Tobata, Kitakyushu 804-8550, Japan
E-mail: qin.jiang816@mail.kyutech.jp, yama@ecs.kyutech.ac.jp

Abstract

The non-photorealistic rendering, NPR, is widely used in social networking service on the mobile device. To realize a real time NPR with low power making battery life of mobile device longer, we attempt to develop hardware module by using high-level synthesis, HLS, converting software to hardware automatically. This research focuses on edge-preserving bubble image, EPB, converting photos into image like filled with bubbles. We proposed a software description method for EPB algorithm so that HLS can generate a high-performance and low-power hardware module. Through the practical experiments, we show that our proposed description method can make HLS generate good hardware module improving the performance and power efficiency compared with the conventional method.

Keywords: Non-photorealistic, rendering, hardware, high-level synthesis, buffer, FPGA.

1. Introduction

In recent years the opportunity to use non-photorealistic rendering into a wide variety of images in social networking services has been increasing. The global market for embedded image processing systems has been expanding rapidly due to the development of such social networking services and measures taken by the Corona disaster to remotely control them and is expected to continue to increase in the future. In order to gain a large market share for embedded image processing systems, it is necessary to develop high-performance and power-efficient image processing system products and introduce them to the market as swiftly as possible.

There are two methods of realizing image processing: hardware and software. Hardware-based processing is desirable because of its higher performance and lower power consumption. However, hardware processing has the disadvantage of a long development period and a large design burden.

In order to solve this problem, it is essential to develop a software description method that can generate efficient hardware by using a technology called High-Level Synthesis (HLS)[1-4], which automatically converts software into hardware.

In this work, we focus on the edge-preserving bubble image, EPB, converting photos into image like filled with bubbles. The EPB is one of non-photorealistic image rendering processes[5], which are conventionally processed by software. We aim to develop a hardware image processing system that is faster than conventional methods by using HLS. This paper develops a software description method so that HLS can generate good hardware module.

2. Edge-preserving Bubble Image Conversion

Edge-preserving bubble transformation is a method to generate an edge-preserving bubble image called EPB (Edge-Preserving Bubble). Fig. 1 shows the overview of this algorithm.

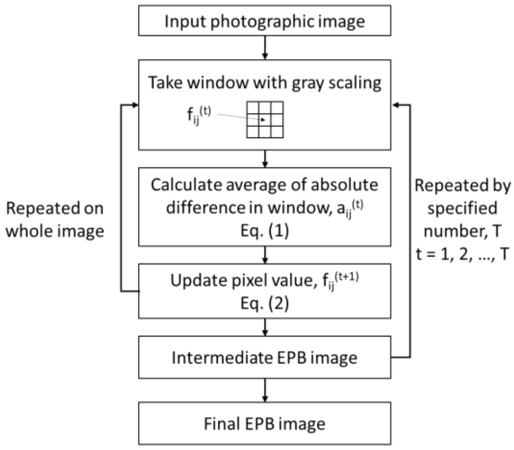


Fig. 1 EPB Algorithm Overview

This algorithm takes a window on the input image as well as spatial filters like mean, sobel, gaussian, and so on. The average absolute brightness difference between the opposite pixels of the processed image corresponds to the edge of the target window, and the transformation of the bubble image can be realized by repeatedly increasing or decreasing the brightness value of the edge equivalent part.

The EPB algorithm is implemented by an iterative process consisting of two steps. The first step is to compute the average absolute difference $a_{i,j}^{(t)}$ in the window by applying the target window of the processed image to Eq. (1).

$$a_{i,j}^{(t)} = \frac{\sum_{k=-1}^{1}\sum_{l=-1}^{1}|f_{i+k,j+l}^{(t)} - f_{i-k,j-l}^{(t)}|}{(2W+1)^2 - 1} \qquad (1)$$

Where, all images to be processed in this study are grayscale images. The $f_{i,j}^{(t)}$ is the luminance value of pixel (i,j), $t(=1,2\ldots)$ is the number of iterations. The k and l are the positions in the window. Then in step 2, all pixels in the process image are update according to Eq. (2).

$$f_{i,j}^{(t+1)} = \begin{cases} f_{i,j} - ba_{i,j}^{(t)} & (t\%2 = 0) \\ f_{i,j} - ba_{i,j}^{(t)} & (t\%2 = 1) \end{cases} \qquad (2)$$

Where, the b is a positive constant, and % is the remainder operator.

If $f_{i,j}^{(t+1)}$ is less than 0, $f_{i,j}^{(t+1)}$ must be set to $-f_{i,j}^{(t+1)}$, and furthermore, the pixel luminance value must be set to 255. If $f_{i,j}^{(t+1)}$ is greater than 255, then $f_{i,j}^{(t+1)}$ must be set to $255 + \frac{255+(255-f_{i,j}^{(t+1)})}{b}$.

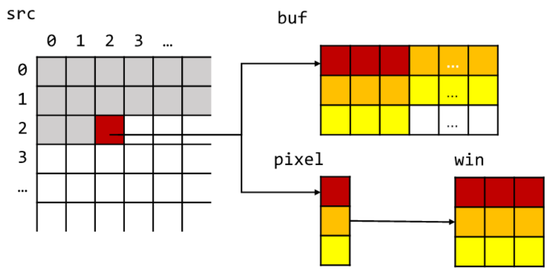


Fig. 2 Buffer Usage for Window Processing

The process of steps 1 and 2 is to be repeated T times for all pixels in the photographic image, and the image composed of the new pixels is the edge-preserving bubble image.

3. Develop of Software Program for HLS

To create efficient hardware, each pixel of the processed image should be read and processed one by one, so that data is continuously fed into the system at each clock cycle, and high-speed processing is achieved through pipelining. To achieve this performing a window-based process, we use a buffer as shown in Fig. 2, and generate the ideal hardware using high-level synthesis. By using

```
for( i = 0 ; i < h; i++ ){
  for( j = 0; j < w; j++ ){
    pix = GetGray( src[i * w + j] ); // Gray scaling
    // Line buffer update
    for( k = 0; k < WIN_N - 1; k++ ){
      buf[k][j] = buf[k + 1][j];
      pixel [k] = buf[k    ][j];
    }
    pixel[WIN_N - 1] = buf[WIN_N - 1][j] = pix;

    // Window update
    for( k = 0; k < WIN_N; k ++ )
      for( l = 0; l < WIN_N - 1; l++ )
        win[k][l] = win[k][l+1];
    for( k = 0 ; k < WIN_N; k++ )
      win[k][WIN_N-1] = pixel[k];

    // Update brightness by Eq.(1) and Eq.(2)
    f = UpdateGrayFixed( win, t );

    // Output new brightness
    dst[i * w + j] = f << 16 ¦ f << 8 ¦ f;
  }
}
```

Fig. 3 Pseudo Source Code for HLS

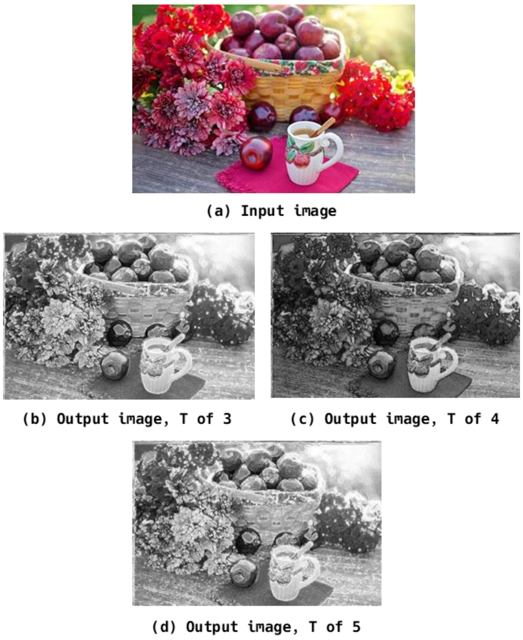


Fig. 4 Input and output Image

three buffers, the memory accesses are performed one by one pixel continuously while performing window processing. Fig. 3 shows the overview of the software programming code we developed for HLS.

4. Experiment and Discussion

4.1. Experimental Setup

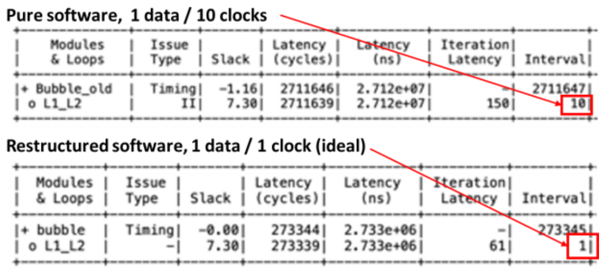
We done the experiments on the real prototype system built on FPGA. We used HLS tool is Xilinx Vitis HLS 2021.1. The hardware generated is implemented on the ZYNQ FPGA. We measured the performance on ZYBO which is an FPGA board with ZYNQ FPGA.

The images used in the experiment are shown in Fig 4 (a). It is a BMP file image of size 427 in height and 460 in width. The output images are also shown in Fig. 4(b)-(d) with the iteration number of 3, 4 and 5.

As comparable evaluations, we prepared three versions: software execution on PC, software execution on embedded processor in FPGA, hardware execution in FPGA. The processor of PC is Core i5 at 3.7GHz. The embedded processor is Cortex A9 at 650MHz. The HLS hardware modules run at 100MHz in FPGA.

4.2. Effect of Software Restructuring

To clarify the effect of software restructure introduced to make memory access perform one pixel by one pixel, we


Pure software, 1 data / 10 clocks

Modules & Loops	Issue Type	Slack	Latency (cycles)	Latency (ns)	Iteration Latency	Interval
+ Bubble_old	Timing	-1.16	2711646	2.712e+07	-	2711647
o L1_L2	II	7.30	2711639	2.712e+07	150	10

Restructured software, 1 data / 1 clock (ideal)

Modules & Loops	Issue Type	Slack	Latency (cycles)	Latency (ns)	Iteration Latency	Interval
+ bubble	Timing	-0.00	273344	2.733e+06	-	273345
o L1_L2	-	7.30	273339	2.733e+06	61	1

Fig. 5 HLS Result

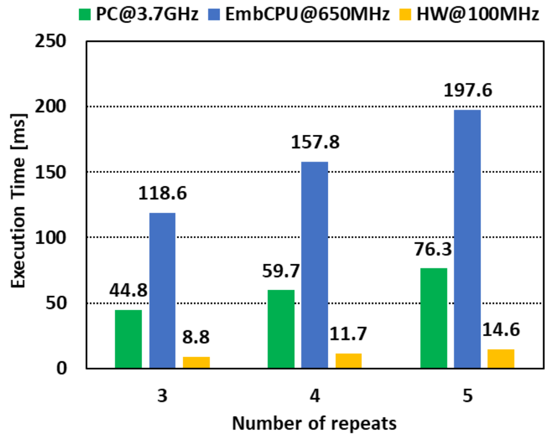


Fig. 6 Execution Time

compare the reports output by the HLS tools. Fig. 5 shows the results that HLS tool converted the pure software intuitively implementing the algorithm and the restructured software considering hardware characteristic shown in Fig. 3.

For pure software, the HLS generated a poor hardware module just produces 1 data per 10 clocks. It is not optimum pipelined data path. In contrast, the HLS can generate the well-organized hardware module with optimum pipelined data path outputting 1 data per 1 clock from the restructured software programming code.

4.3. Performance Evaluation

Fig. 6 shows the measured execution time of PC, embedded processor and FPGA respectively.

Based on Fig. 6, we confirmed that the hardware execution time was reduced by a factor of about 5 and 13 compared to the software execution time on the computer and that on the embedded processor respectively. Our hardware module generated by HLS from software code restricted can achieve a significant performance improvement compared with the software execution.

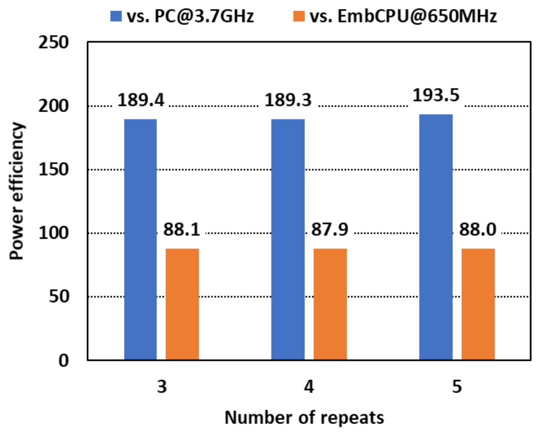


Fig. 7 Power Efficiency

4.4. Power Efficiency Estimation

To investigate a power efficiency about hardware module compared to the software execution, we estimate a operating power efficiency by Eq. (3).

Fig.7 shows the power efficiency estimated by the measured values shown in Fig. 6 and Eq. (3). This result shows that the hardware module generated by HLS from out restricted software can achieve a significant power efficiency compared to the software execution on PC and

$$\text{Power Efficiency} = \frac{\text{Software operating power}}{\text{Hardware operating power}} = \frac{\text{SW exec. time} \times F_{CPU}}{\text{HW exec. time} \times F_{FPGA}} \quad (3)$$

$$F_{CPU} = 3.7\text{GHz, PC} \qquad F_{FPGA} = 100\text{MHz}$$

$$= 650\text{MHz, embedded CPU}$$

embedded processor.

5. Conclusion

In this paper, we focused on speeding up the EPB bubble image conversion system and developed a software description method to generate efficient hardware by high-level synthesis. In this study, we developed a software description method to generate efficient hardware by high-level synthesis. When the hardware generated by high-level synthesis was compared with the software, it was confirmed that the power efficiency was also greatly improved. Furthermore, when compared to the execution on an embedded processor, the performance was also improved. Considering these points, we can conclude that the proposed method is an effective method for embedded image processing systems that require high-speed processing with low power consumption.

References

1. Shinya TAKAMAEDA, "Hardware Design with High Level Synthesis : Custom FPGA Computer Development Is Available to Everyone," Journal of the Institute of Electronics, Information and Communication Engineers, Vol.100, No.2, pp. 103-108 (2017) in Japanese.
2. Nane R, Sima VM, Olivier B, Meeuws R, Yankova Y, and Bertels K, "DWARV 2.0: a CoSy-based C-to-VHDL hardware compiler," In Proceedings of 22nd international conference on field programmable logic and applications, Oslo, Norway, 29–31 August 2012, pp 619‑622.
3. Pilato C and Ferrandi F, "Bambu: a modular framework for the high-level synthesis of memory-intensive applications," In: Proceedings of 23rd international conference on field programmable logic and applications, Porto, Portugal, 2–4 September 2013, pp. 1–4.
4. Canis A, Choi J, Aldham M, Zhang V, Kammoona A, Anderson JH, Brown S, and Czajkowski T, "LegUp: high-level synthesis for FPGA-based processor/accelerator systems," In: Proceedings of the 19th ACM/SIGDA international symposium on field programmable gate Arrays, Monterey, CA, USA, 27 February–1 March 2011, pp 33–36.
5. Toru HIRAOKA, "A High-Speed Method for Generating Edge-Preserving Bubble Images" IEICE TRANS. INF & SYST. VOL.E103-D, NO.3 MARCH 2020.

Authors Introduction

Ms. Jiang Qin

She entered the department of Electrical and Electronic Engineering, Faculty of Engineering, Kyushu Institute of Technology in 2018. She is now pursuing her B.E. study at the same university. Her current research interests are in Hardware development of image conversion.

Dr. Akira Yamawaki

He is an Associate Professor of Faculty of Engineering at Kyushyu Institute of Technology in Japan. He received his PhD. in Electrical Engineering from the Graduate School of Engineering, Kyushu Institute of Technology, in 2006. His reserch interest in Digital Circuit Systems.

Development of Haze Removing Hardware Using High-Level Synthesis

Daiki Shirai, Akira Yamawaki
[1,2]*Department of Engineering, Kyushu Institute of Technology, 1-1, Sensui-cho, Tobata-ku Kitakyushu-shi, Fukuoka, Japan*

E-mail: shirai.daiki885@mail.kyutech.jp, yama@ecs.kyutech.jp

Abstract

We have developed a hardware for low-power and high-speed haze removing to sharpen hazy images in battery-powered embedded image processing systems such as drones. For development, we used HLS (High-Level Synthesis), which automatically converts software into hardware. In this paper, we evaluate the performance and power efficiency of the hardware obtained by high-level synthesis based on the haze removing software developed considering the hardware specifications and show its effectiveness.

Keywords: Haze Removing, High-Level Synthesis, HLS, FPGA

1. Introduction

In recent years, the research and development of embedded devices using image recognition has been very active, and the market scale is rapidly expanding.

Among them, the use of aerial images taken by cameras mounted on drones and other UAV equipment is attracting a lot of attention, and research and projects related to it are being actively conducted. In Japan, where infrastructure facilities are aging rapidly and the workforce is shrinking, more and more facilities are being inspected using aerial images from drones.

In such cases, the aerial images are blurred by weather and shooting conditions, making it impossible to perform accurate inspections. In order to solve this problem, we decided to develop an image processing module that can be installed in the drone to remove the haze generated in the captured images in real time.

The module to be developed in this paper is a hardware that executes the haze removing based on the darkest channel priority method,[1, 2] and high-level synthesis is used for its development. In addition, to ensure that the hardware to be developed has high performance, we will also develop software programs that describe the processing contents and modify the algorithms in consideration of high-level synthesis.

Finally, the effectiveness of the developed hardware is demonstrated by comparing the hardware obtained by high-level synthesis with pure software following the original algorithm and software considering hardware.

2. Haze Removing

2.1. *Dark channel priority method*

The haze removing algorithm in this paper is based on the dark channel priority method. Normally, when taking pictures with a camera, skylight[a] hits an object, and the light bounced off the object, called direct light, reaches the camera to capture the image. On the other hand, if haze or mist is present at the time of shooting due to

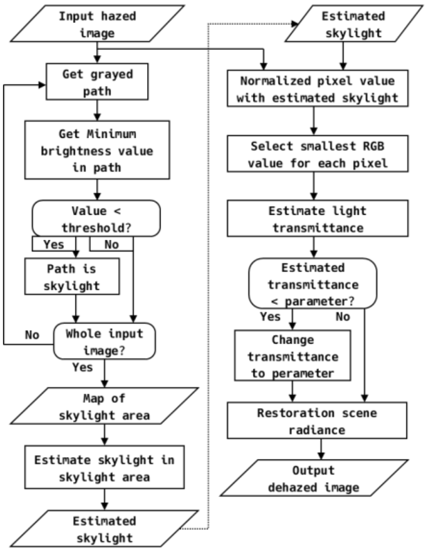


Fig. 1. Flowchart of haze removing.

weather or other factors, the skylight will be scattered by the haze, creating what is called airglow. This atmospheric light mixes with the direct light before it reaches the camera, creating a haze in the captured image.

Based on these facts, the mechanism that causes haze can be modeled by the following equation

$$I(x) = J(x)t(x) + A(1 - t(x)). \qquad (1)$$

Here, x is the entire shooting scene, J is the direct light from the object (the original clear image), A is the airglow (assumed to be constant throughout the shooting scene), and $t(0 < t \leq 1)$ is the transmittance (mixing of direct light and skylight), $A(1 - t(x)$ is airglow, and I is the captured image (hazed image).

In order to obtain the original image, the above equation is transformed into the following equation.

$$J(x) \approx \frac{I(x) - \tilde{A}(x)}{\tilde{t}(x)} + \tilde{A}(x). \qquad (2)$$

Here, $\tilde{A}$ is an estimated value of skylight, and $\tilde{t}$ is an estimated value of transmittance. By obtaining $\tilde{A}$ and $\tilde{t}$, direct light from objects, that is, the original clear image can be obtained.

2.2. *Flowchart of haze removing*

The haze removing is divided into four stages: extraction of skylight region, calculation of the skylight estimate, calculation of the transmittance, and restoration of direct light. In this section, these four processes will be explained in broad terms (Fig. 1).

The first and second steps will be described together as the skylight estimation process. Assuming that the influence of skylight in the image is reflected in the luminance, the luminance value is calculated by the RGB value of each pixel. This is then subjected to a minimization filter for each local region on the image called a patch. The minimum value for each patch is compared with the set threshold value, and if it exceeds the threshold value, the entire patch is set as the skylight region.

Since the next calculation of the skylight estimate assumes that the value of skylight in the image is constant, it is obtained by taking the average of the RGB values of the estimated skylight region.

The third stage of transmittance estimation is calculated using the following equation, which is obtained by transforming equation (1)

$$\tilde{t}(x) \approx 1 - \alpha \min_{w \in \Omega(x)} \left[\min_{C \in \{R,G,B\}} \left\{ \frac{I^C(x)}{\tilde{A}^C} \right\} \right]. \qquad (3)$$

Here, $\tilde{t}$ is the estimated transmittance and $\alpha(0 < \alpha \leq 1)$. From equation (3), the transmittance can be calculated by implementing the darkest channel local minimization filter on the normalized haze image and finding the value adjusted by α to prevent it from becoming extremely small.

In the final process of restoring scene radiance, we use equation (2) based on the estimated skylight value and transmittance. In this process, the denominator value is adjusted so that the first term on the right side of equation (2) does not become extremely large.

This is an overview of the haze removing based on the darkest channel priority method, and the software created based on this method is called pure software.

3. Haze Removing Software for HLS

3.1. *HLS (High-level synthesis)*

HLS automatically generates circuit data described by HDL (Hardware Description Language) such as Verilog from software program. While there are advantages to using HLS, such as greatly reducing the development time and the burden on developers, there are several things to keep in mind.

For example, if the software input to the HLS uses recursive functions or floating point, the circuit size of the resulting hardware will be unnecessarily large, and the performance will be low. Therefore, we design a software program considering HLS based on the flowchart described in Section 2.2, taking the above points into account.

3.2. *Software design considering HLS*

For the skylight estimation process, which is the first and second steps of haze removing, a software program for HLS has been developed in a previous study and its effectiveness has been confirmed[3]. Briefly there are three major changes in the skylight estimation process for high level synthesis: converting patch processing to pixel processing, removal of the skylight region extraction map, and finally, pipelining of the process. The patch processing has been converted to pixel processing and the skylight region extraction map has been removed. This strategy gives a significant improvement in terms of memory access for hardware. The pipelining of processing plays a major role in improving the processing performance of the hardware being generated.

For the third step, the transmittance estimation process, we changed the method to perform the process for each pixel as in the skylight estimation process described above, as opposed to using the local minimum value filter in the conventional method. In the case of the conventional method using the local minimum filter, the banding[b] caused by this effect is repaired by applying a smoothing process called soft matching. However, this soft matching is complex and computationally expensive, so implementing it in hardware would unnecessarily increase the circuit size. In the proposed method, the transmittance is estimated for each pixel instead of the local minimum filter, and the restoration process can be performed without smoothing and with maintaining the hardware performance.

In the last step of restoring the scene radiance, the calculation process was changed from the floating-point method used in the conventional method to a fixed-point method using unsigned 32-bit integers that can maintain the minimum necessary accuracy. This calculation method was applied in the same way in the transmittance estimation process. This will reduce the circuit size of the development hardware.

Based on the above, we created a haze removing software program considering HLS (hereafter referred to as HLS software).

4. Experiments and Discussions

4.1. *Experimental environment*

The HLS software program created in Section 3.2 was input to Xilinx's high-level synthesis tool Vivado HLS 2018.3 to perform high-level synthesis. Then, the HDL program generated by high-level synthesis was loaded on ZYBO (Zynq-7000 Development Board), an FPGA board manufactured by DIGILENT, by using Vivado 2018.3, a development environment software for Xilinx FPGAs, to verify the operation. The operating frequency for each device is 2.9GHz for the PC, 650MHz for the embedded CPU on the ZYBO board (hereafter referred to as ZYBO CPU), and 100MHz for the developed hardware, and the size of the input image used in the experiment was 1024 x 768 pixels.

4.2. *Experimental method*

The performance of the HLS hardware was evaluated in terms of the processing performance improvement ratio and the runtime power improvement ratio for each of the following four patterns, (i) pure software on PC, (ii) HLS software on PC, (iii) pure software on ZYBO CPU, (iv) HLS software on ZYBO CPU.

The processing performance improvement ratio is calculated by determining the ratio of the processing execution time in each case to the time required by the HLS hardware for processing, and the runtime power improvement ratio is obtained by multiplying the power improvement ratio [c] by the processing performance improvement ratio.[4, 5]

4.3. *Experimental results and discussions*

A graph summarizing the time required for the HLS hardware and each of (i) through (iv) listed in Section 4.2 to perform the haze removing is shown in Fig. 3.

As can be seen from Figure 3, the HLS hardware outperforms its execution time for all cases except (ii). The processing time of the HLS hardware is slower for

[b] Noise caused by discontinuities in transmittance at the borders of local regions

[c] Ratio of the respective operating frequencies of the HLS hardware and the case under comparison

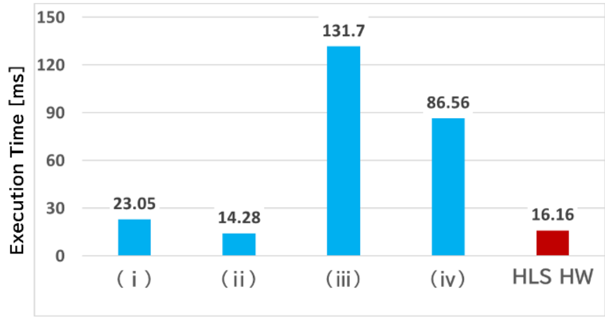


Fig. 3. Execution time of the haze removing in each case

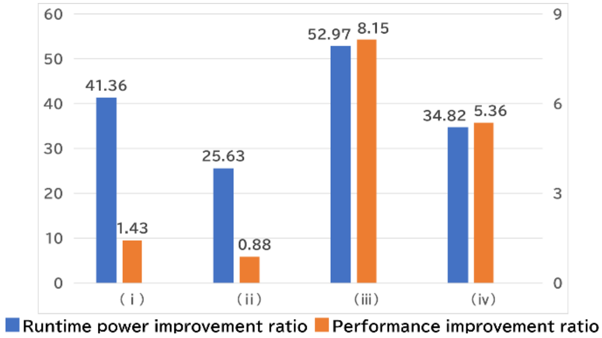


Fig. 4. Runtime power and performance improvement ratio

(ii), but this is probably due to the large difference in operating frequency.

Fig. 4 shows the runtime power and performance improvement ratios of the HLS hardware for each case.

As for the performance improvement ratio, since it is a ratio of execution time, we were able to achieve performance improvement for all cases except (ii). As for the runtime power improvement ratio, even in the case of (ii), where no performance improvement was achieved, we were able to achieve a power improvement of approximately 25.6 times.

5. Conclusion

We have developed a low-power and high-performance hardware of the haze removal process using high-level synthesis by restricting algorithm.

Although the HLS hardware could not even surpass the processing time of the software used in HLS when run on a PC, it was able to improve the runtime power with almost no loss in processing performance.

In the future, we will actually mount the hardware developed in this paper on a drone to verify its usefulness.

References

1. K. He et al: "Single Image Haze Removal Using Dark Channel Prior. Proc.", IEEE CVPR, Vol.1, pp.1956-1963, 2009
2. Hiroaki Kotera: "A color Correction for Degraded Scenes by Air Pollution", Journal of the Color Science Association of Japan, Vol.40, No.2, pp. 49-59, 2016
3. D. Shirai, A. Yamawaki: "Hardware Development of Skylight Estimation Processing in Haze Removing Using High-level Synthesis", ICIAE2021, pp107-111, 2021
4. M. Yamasaki et al: "Effect of Redundant Function Execution to Reduce Memory Access on High-level Synthesis", Processing of the 6th IIAE International Conference on Industrial Application Engineering 2018, pp.206-209, 2018
5. M. Yamasaki, A. Yamawaki: "Description Method for High-level Synthesis of Histogram Generation and Their Evaluation", IEIE Transactions on Smart Processing and Computing, Vol. 8, No. 3, pp.178-185, 2019

Authors Introduction

Mr. Daiki Shirai

He received his bachelor's degree from Faculty of Engineering, Department of Electrical and Electronic Engineering, electronic Systems Engineering Course, Kyushu Institute of Technology, Japan in 2020. He is currently a master's course student in Kyushu Institute of Technology, Japan.

Dr. Akira Yamawaki

He is an Associate Professor of Faculty of Engineering at Kyushyu Institute of Technology in Japan. He received his PhD. in Electrical Engineering from the Graduate School of Engineering, Kyushu Institute of Technology, in 2006. His reserch interest in Digital Circuit Systems.

Automatic Approximation of Primitive Shapes using Point Clouds

Yuma Yoshimoto, 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: yoshimoto@brain.kyutech.ac.jp, tamukoh@brain.kyutech.ac.jp
https://www.kyutech.ac.jp

Abstract

This paper proposes a method to estimate appropriate primitive shapes by automatically using a point cloud of objects. The process is as follows. First is to estimate a rotation angle of the object and, then place a primitive shape (i.e., cylinder or sphere) in the center of the object. The primitive shape is either stretched or compressed to fit the object. The distance between all the point cloud of the object and the primitive shape are measured. We apply these methods with various primitive shapes and find the most suitable primitive shape. We show that objects like apples and chikuwa can be recognized with primitive objects, such as spheres and cylinders.

Keywords: Primitive Shape, Point Cloud, Bulk Picking

1. Introduction

Bulk picking is one of the fundamental tasks of industrial robots[1]. In this task, the robot is tasked to picks up an object from a case containing several objects. To achieve this task, the robot required to estimate a grasping point of the object. However, this is a difficult task for the robot to estimate the proper grasping point without knowing the hiding point or occlusion of the object. To understand the occlusion of the object, it is necessary to teach shapes of the object to the robot with CAD data. However, with unknown objects, the robot must investigate the shape of the object by itself. One of the methods to determine the object shape is to approximate it with primitive shapes. The primitive shapes are defined simple shapes, for instance a sphere or a rectangle.

In this study, we propose a method for estimating primitive shapes in a situation where a robot observes a part taken out of an enclosed case. Our method is to calculate the likelihood of shapes by stretching and shrinking several kinds of primitive shapes and compare to the object. With this we able to determines the closest primitive shape to the object. In this research, we evaluate our method with two types of objects: apples and bananas. Our result shows that we were able to find an appropriate primitive shape for each object.

2. Importance of Understanding Occlusion

Occlusion is a part of an object that is not visible with the observer point of view. The example of occlusion is shown in Fig. 1. Figure 1 displays a case where an understanding of the shape occlusion is necessary to pick the object. In this situation, a robot arm is tasked to carry fried food from the enclosed box. If the robot does not understand the object's occlusion, the robot may fail to find an optimized location (center of gravity) of the object to grasp and fail the task. To prevent this situation, the robot must understand the shape of the object.

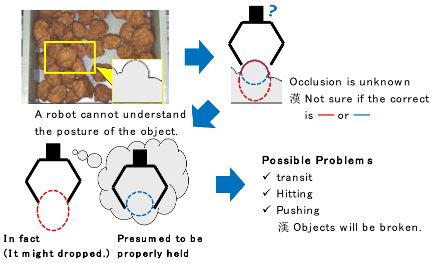


Fig. 1. Importance of understanding occlusion

3. Proposed Methods

The overview flow of our proposed method is shown in Fig. 2. In general, the system works as follows:

1. Capture the object point cloud from an interested object with the RGB-D camera.
2. Restrict the area.
3. Detect and remove the planes.
4. Estimate the orientation of the object.
5. Place each primitive shape at the estimated orientation and Cartesian coordinate of the object, and scale each shape according to the size of the object point cloud.
6. Measure the distance between the point cloud and the primitive shapes, and detect the object with the shortest distance.
7. Steps 4 to 6 are applied several times to decide the most appropriate primitive shape.

The object with the shortest distance is considered the object with the best match. In following sections, we further describe important steps in our proposed method.

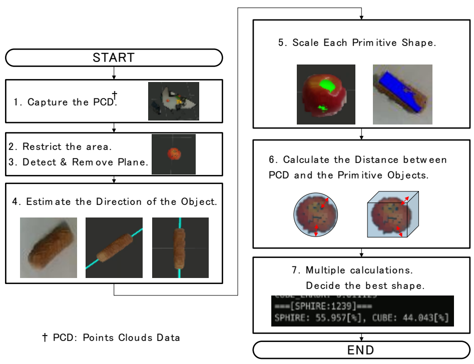


Fig. 2. Overview pipeline of our proposed method

3.1. *Step 4. Estimate the direction of the object.*

In step 3, we divide the process into XY and XZ planes. In the XY, the yaw angle of the object is calculated, and in the XZ, the pitch angle of the object is calculated. These are calculated using the least-squares method.

3.2. *Step 5. Scale each primitive shapes*

In step 5, primitive shapes are placed in the point cloud of the object and stretch. First, the system finds a center of gravity of the object point cloud. Next, it places a primitive shape at the location of center of gravity. Finally, the primitive shape is stretched and contracted.

3.3. *Step 7. Decide the best shape*

These process from step 4 is repeated several times, in each time, voting for each object that is estimated to be appropriate. After a predefined threshold number of votes, the shape with the highest vote rate is decided.

4. Experimental results

The experiment was conducted to evaluate our method using objects like apples and chikuwa. For apples, the voting score for spheres was 57%, and for chikuwa, the voting rate for cylinders was 62%. This indicated that appropriate primitive shapes can be voted.

5. Conclusion and Future Work

We proposed a method for estimating primitive shapes that adequately approximate objects for bulk picking industrial robots. We conducted an experiment with apples and chikuwa. We found that our method can recognize the approximate primitive shape objects properly. In the future, we will increase the number of primitive shapes and construct a system to apply the approximated primitive shapes to objects stacked together. In addition, we will integrate with a system that performs highly accurate object recognition using RGB and Depth images[2].

References

1. D. Buchholz, M. Futterlieb, S. Winkelbach, F. M. Wahl, *Efficient Bin-Picking and Grasp Planning Based on Depth Data*, IEEE International Conference on Robotics and Automation (ICRA) 2013, pp. 3245-3250, 2013.
2. Y. Yoshimoto, 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.

Authors Introduction

Dr. Yuma Yoshimoto

He received his B.Eng. degree from National Institute of Technology (KOSEN) Maizuru College, Japan, in 2016. He received his M.Eng. and D.Eng degrees from Kyushu Institute of Technology, Japan, in 2018 and 2021, respectively. And, he was JSPS researcher, in 2019-2021. Currently, he is a post-doctoral researcher at the Kyushu Institute of Technology, Japan. His research interests include deep learning, robot vision and digital hardware design.

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.

An EtherCAT Based Delta Robot Synchronous Control Application

Chung-Wen Hung*, Yu-Hsuan Tseng, Guan-Yu Jiang
Department of Electrical Engineering, National Yunlin University of Science & Technology, 123 University Road, Section 3, Douliou, Yunlin 64002, Taiwan, R.O.C.

Chau-Chung Song
Department of Aeronautical Engineering, National Formosa University, No.64, Wenhua Rd., Huwei Township, Yunlin County 632, Taiwan
Email: wenhung@yuntech.edu.tw

Abstract

The delta robot synchronous control based on the Ethernet Control Automation Technology (EtherCAT) protocol is proposed in this paper. Personal Computer (PC) is used as master and the delta robot motor drivers are used as slaves in this work. The Master sends command to slave base on the motion control profile CAN in Automation 402(CiA402). Subsequently, the program in C# perform the user's interface and EtherCAT communication. And the system is not only easy use, but also quickly high-precision. A complex painting application is proposed to show this system workable.

Keywords: delta robot, kinematics, path planning, EtherCAT

1. Introduction

In recent years, the problem of labor shortage is getting worse. [1][2] proposed to reduce the labor costs, improve production efficiency, and stabilize production quality, the company has successively introduced automation technology or replace manpower with a robotic arm.

[3][4] knows that controlling each axis of the robotic arm needs to perform coordinate movement tasks at the same time, so accuracy and synchronization are important for control the robotic arm. Using Ethernet to realize synchronous control of each axis is low-stability and take lots of time for the response and cyclic task. Therefore, in order to achieve the precise synchronous control of each axis, this system uses the industrial communication EtherCAT as the communication protocol between the robotic arm and the PC, which has the advantages of processing on the fly and good synchronization[5]-[9].

[10][11] proposed that using EtherCAT high speed communication to improve accuracy and to reach higher control frequency between the motor drivers and master. [12] used EtherCAT to communicate with multi-axis motors. The results show that EtherCAT solves the problem of transmitting large amounts of data in real-time and realizes high-precision synchronous control with high-efficiency algorithms. [13] compared EtherCAT and other common industrial protocols on multi-axis motor control. The result shows that EtherCAT control is accurate and synchronize.

2. Kinematics of the Delta Robot

The motion trajectory of the delta robot is based on a parallelogram as a track, moving on the X-axis, Y-axis, or Z-axis, and can rotate around the vertical axis of the platform center.

If the rotation angle of each axis motor ($\theta_1,\ \theta_2,\ \theta_3$) is known, we can use vector loop equations proposed by [14][15] to calculate the coordinate vector ($A_{1v'}, A_{2v'}, A_{3v'}$) via the forward kinematics. Finally, the above coordinate vectors are substituted into the sphere equations and quadratic formula to derive the final target coordinates P(*x, y, z*) of the end effector.

$$\boldsymbol{A_{1v'}} = \begin{bmatrix} \mathbf{0} \\ \frac{-\sqrt{3}}{3}\boldsymbol{s_P} + \frac{-\sqrt{3}}{6}\boldsymbol{s_B} - \boldsymbol{L}\cos\boldsymbol{\theta_1} \\ -\boldsymbol{L}\sin\boldsymbol{\theta_1} \end{bmatrix}$$

$$A_{2v'} = \begin{bmatrix} \frac{\sqrt{3}}{2}\left(\frac{\sqrt{3}}{6}s_B + L\cos\theta_2\right) - \frac{\sqrt{3}}{3}s_P\cos\frac{\pi}{6} \\ \frac{1}{2}\left(\frac{\sqrt{3}}{6}s_B + L\cos\theta_2\right) - \frac{\sqrt{3}}{3}s_P\sin\frac{\pi}{6} \\ -L\sin\theta_2 \end{bmatrix}$$

$$\boldsymbol{A_{3v'}} = \begin{bmatrix} \frac{-\sqrt{3}}{2}\left(\frac{\sqrt{3}}{6}\boldsymbol{s_B} + \boldsymbol{L}\cos\boldsymbol{\theta_3}\right) + \frac{\sqrt{3}}{3}\boldsymbol{s_P}\cos\frac{\boldsymbol{\pi}}{6} \\ \frac{1}{2}\left(\frac{\sqrt{3}}{6}\boldsymbol{s_B} + \boldsymbol{L}\cos\boldsymbol{\theta_3}\right) - \frac{\sqrt{3}}{3}\boldsymbol{s_P}\sin\frac{\boldsymbol{\pi}}{6} \\ -\boldsymbol{L}\sin\boldsymbol{\theta_3} \end{bmatrix}$$

If the final target coordinates P(x, y, z) is known, we can use vector loop equations proposed by [14][15] to calculate the rotation angle of each axis motor ($\boldsymbol{\theta_1}$, $\boldsymbol{\theta_2}$, $\boldsymbol{\theta_3}$) via the inverse kinematics. The relevant equation is such as Eq.(1)-Eq.(3).

$$\boldsymbol{\theta_1} = 2\tan^{-1}\left(\frac{-z \pm \sqrt{4z^2 - 4F_1(F_1 - y - a)}}{F_1}\right) \tag{1}$$

$$\boldsymbol{\theta_2} = 2\tan^{-1}\left(\frac{2z \pm \sqrt{16z^2 - 4F_2\left(F_2 + 2\sqrt{3(x+b)} + y + c\right)}}{F_2}\right) \tag{2}$$

$$\boldsymbol{\theta_3} = \mathbf{2}\, \boldsymbol{tan}^{-1}\left(\frac{\mathbf{2z} \pm \sqrt{\mathbf{16z^2} - \mathbf{4F_3(F_3} - \mathbf{2}\sqrt{\mathbf{3(x-b)}} - \mathbf{y} - \mathbf{c)}}}{\boldsymbol{F_3}}\right) \quad (3)$$

where:

$$F_1 = (y+a) - \frac{1}{2L}(x^2+y^2+z^2+a^2+L^2+2ya-l^2)$$

$$F_2 = -\left(\sqrt{3(x+b)}+y+c\right) - \frac{1}{L}(x^2+y^2+z^2+b^2+c^2+L^2 + 2(xb+yc)-l^2)$$

$$F_3 = \left(\sqrt{3(x-b)}-y-c\right) - \frac{1}{L}(x^2+y^2+z^2+b^2+c^2+L^2 - 2(xb-yc)-l^2)$$

The parameters are defined as follows:

$$\mathbf{a} = \frac{-\sqrt{\mathbf{3}}}{\mathbf{6}} s_P\,; \quad \mathbf{b} = \frac{s_P}{\mathbf{2}} - \frac{\mathbf{1}}{\mathbf{4}} s_B\,; \quad \mathbf{c} = \frac{\sqrt{\mathbf{3}}}{\mathbf{6}} s_P - \frac{\sqrt{\mathbf{3}}}{\mathbf{12}} s_B\,;$$

name	meaning	value (mm)
s_B	base equilateral triangle side	865.9388
s_P	platform equilateral triangle side	118.5589
L	upper legs length	336.1
l	lower legs parallelogram length	1022.4

To plan the trajectory in the joint space, we translate the current coordinates to each axis angle from the inverse kinematics, to calculate the following error in the cartesian space, the current axis was translated to coordinates by inverse kinematics. Fig.1 and Fig.2 are the kinematics analysis diagrams of the delta robot.

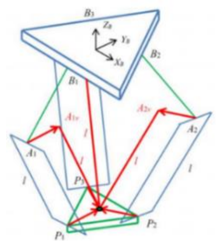


Fig.1 Delta robot FK diagram [15]

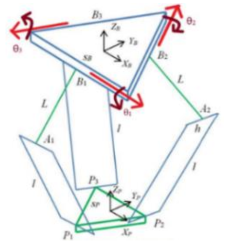


Fig.2 Delta robot IK diagram [15]

3. Path Planning

In order to complete the path planning of the image. First, upload the image from the user and convert the picture to a binary image. Then analyze and sort the coordinate pixels of the neighboring to connect them as paths. At last, we can solve the point to point velocity by planning maximum movement distance between each coordinate point in the cartesian space.

The path planned by adjacent coordinate make the robot moves continuously during the drawing process and accurately shows the image. After test, we use the greedy algorithm to find the adjacent coordinate.

The greedy algorithm is a fast-iterative method to find successive regional optimal solutions. Each regional optimal solution must be proved the final problem obtain the overall optimal solution and can be reduced the problem range to get the overall optimal solution [16], the design steps are as follows:

(1) Store the path coordinate in the array, and then define the thresholds of adjacent points to complete the establishment of the problem model.

(2) Divide the problem into n.

(3) Define the purpose (Find the coordinates nearest to the target).

(4) Find the best solution to n small problems based on the greedy strategy.

(5) Combine the best solutions of n small problems to get the overall best solution.

4. System Architecture

Fig.3 and Fig.4 show the system architecture and actual configuration diagrams developed in this paper. In order to complete the communication of the hardware, image processing, path planning, calculation of the forward and inverse kinematics, and the delta robot control algorithm, the Windows operating system as the computing platform used in this paper.

The Automation Device Specification(ADS) communication protocol from Beckhoff is used to exchange data between TwinCAT, the EtherCAT master station, and the user's human-machine interface based on Visual Studio C#.

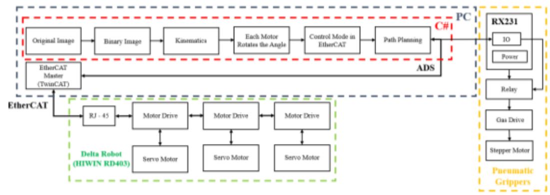


Fig.3 System architecture diagram

Fig.4 Actual configuration diagram

To complete the motor control of each axis of the robotic arm, the system uses the CiA402 sub-protocol based on the CANopen over EtherCAT (CoE) protocol that conforms to the EtherCAT application layer.

In addition, the user could plan the robotic arm's trajectory by their motion and control algorithm, so the Cyclic Sync Position mode (CSP mode) and Cyclic Sync Velocity mode (CSV mode) in the CiA402 sub-protocol are more suitable for multi-axis synchronous control.

To complete the purpose of multi-axis synchronization control, the system chooses CSP and CSV modes with 4ms as the communication cycle time. The following will introduce the control methods based on these two control modes in this paper:

(1) Use CSP mode, set the command of motor rotation angle distance to the Target Position object on the driver.

(2) Use CSP mode, set the command of motor rotation angle distance to the Target Position object on the driver, as well as calculate the following error in real-time, and set to velocity offset object value to improve velocity references.

(3) Use CSV mode, convert the command of motor rotation angle distance to speed, and set it to the Target Velocity object on the driver.

(4) Use CSV mode, convert the command motor rotation angle distance to speed, and set it to the Target Velocity object on the driver as well as calculate the following error in real-time, and set it to the velocity offset object value to compensate error.

(5) Use CSV mode, convert the command of motor rotation angle distance to speed, and set it to the Target Velocity object on the driver, as well as calculate the following error in real-time. Compensate the error according to the proportional-derivative controller (PD controller) shown in Fig.5.

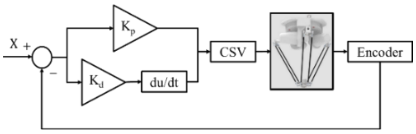


Fig.5 PD controller system block diagram

5. System Testing Result

To discuss the method proposed in this paper, we compared the accuracy with different maximum movement distances (such as: 0.1mm, 0.01mm, and 0.005mm) with each method.

The root-mean-square error (RMSE) can be a simple standard to make comparison of the data in different numerical range. The equation is $\sqrt{\frac{\sum_{t=1}^{n}(\hat{y}_t-y_t)^2}{n}}$ and the correlation results are shown in Fig.6 to Fig.8 and Tab.1 to Tab.2.

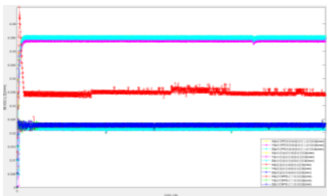


Fig.6 Movement error diagram of five control method with a maximum movement distance of 0.005mm

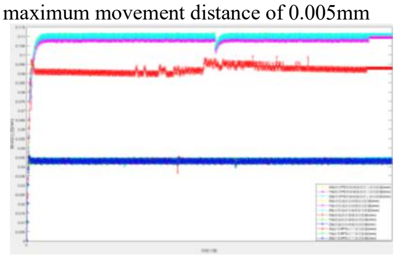


Fig.7 Movement error diagram of five control method with a maximum movement distance of 0.01mm

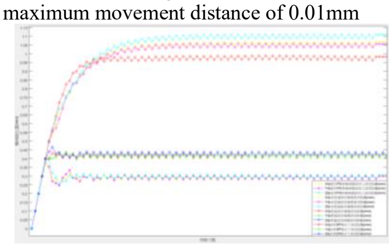


Fig.8 Movement error diagram of five control method with a maximum movement distance of 0.1mm

Tab.1 Different control methods and motor moving distance in the specified path of the RMSE

Control Method / Maximum moving distance of motor(mm)	CSP mode	CSP mode, compensate error to Velocity Offset	CSV mode	CSV mode, compensate error to Velocity Offset	CSV mode with PD controller to compensate error
0.005	8.554E-02	8.872E-02	9.424E-02	3.952E-02	4.009E-02
0.01	1.601E-01	1.656E-01	1.882E-01	7.410E-02	7.502E-02
0.1	1.673E+00	1.705E+00	1.731E+00	7.162E-01	5.077E-01

Tab.2 Different control methods and motor moving distance in the specified path of the final coordinate error

Control Method / Maximum moving distance of motor(mm)	CSP mode	CSP mode, compensate error to Velocity Offset	CSV mode	CSV mode, compensate error to Velocity Offset	CSV mode with PD controller to compensate error
0.005	8.689E-02	8.689E-02	9.540E-02	3.880E-02	4.089E-02
0.01	1.587E-01	1.620E-01	1.912E-01	7.260E-02	7.693E-02
0.1	1.770E+00	1.770E+00	1.867E+00	7.116E-01	5.261E-01

The more distance we move means the more angle that motor rotate, will cause bigger moving error. To discuss the relationship between moving distance and error, there are three moving distance commands are tested, they are 0.005, 0.01 and 0.1 mm. It can be concluded from Tab.1 and Tab.2. The better choice between two kinds of CSP method is the simple CSP control. On the other hand, CSV mode with PD control is the better one in CSV control.

To completely reproduce the human painting, it is necessary to compensate the errors instantly when drawing the path. In this case, we choose CSV mode with PD control to be our control method. It will compute the error distance in real-time, next convert the error distance to velocity, then use the PD controller to speed up or slow down to complete compensation. We take Fig. 9 as the example. Use 0.1mm as the maximum moving distance to compare simple CSP mode with CSV mode with PD controller. Fig.10 and Fig.11 are the drawing results. Tab.3 shows the RMSE of two control mode. According the result of RMSE, it is easy to find that the CSV mode with PD controller is a better choice.

(a) Original image (b) Drawing path

Fig.9 The command drawing path

Fig.10 The drawing result in the CSP mode

Fig.11 The drawing result in the CSV mode with PD controller

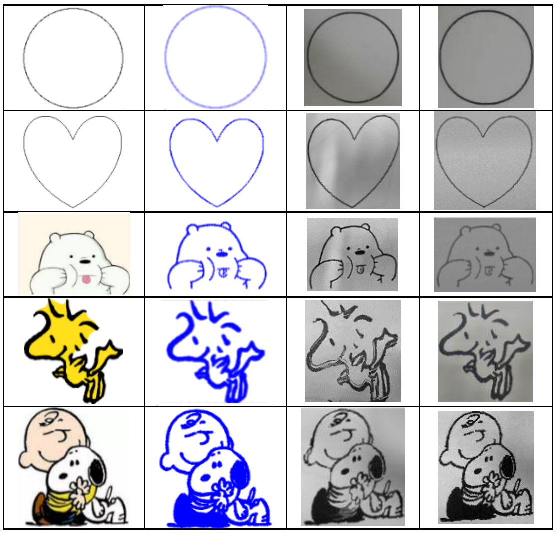
Tab.3 The RMSE of five control method with a maximum movement distance of 0.1mm

	CSP mode	CSV mode with PD controller to compensate error
RMSE(mm)	9.1214E-01	4.3303E-01

We also use these two control methods with the maximum moving distance 0.1mm to draw other figures. The results are shown in Tab.4.

Tab.4 The drawing result in the CSP mode and CSV mode with PD controller to compensate error

Original image	Drawing path	CSP mode	CSV mode with PD controller to compensate error



6. Conclusion

For the image processing, it converts the image to a binary image, analyze and sort the coordinate pixels of the neighboring to connect them as path. After the test, the analyzed image command coordinate points are enough to present the original image information.

For the robotic arm control, if you want to use CSP mode and CSV mode, you need to adjust the proportional coefficient and derivative coefficient of the motor drivers to reduce the error. After the actual adjustment, the error was not reduced due to incorrect adjustment methods or mechanical problems. In order to fully reproduce human paintings, we used in the CSV mode with PD control. This method can via calculating the following error distance of the end effector in real-time, convert the following error distance to velocity, and use the PD controller to improve the velocity to complete compensation.

Experiments have confirmed that in the CSV mode with PD control, the error is less than 0.5mm which is the smallest error in these control methods. What's more, this control method compensates error in real-time, that can precisely control the delta robot to complete the path planning and fully reproduce human paintings.

References

[1] A. Nayyar and A. Kumar, "A Roadmap to Industry 4.0 : Smart Production, Sharp Business and Sustainable Development," Springer, Cham,2020, pp 157-169

[2] H. A. F. Almurib, H. F. Qrimli and N. Kumar, "A review of application industrial robotic design," 2011 Ninth International Conference on ICT and Knowledge Engineering, February 2012

[3] I. K. Jung and S. Lim "An EtherCAT based real-time centralized soft robot motion controller" 2012 International Symposium on Instrumentation & Measurement, Sensor Network and Automation (IMSNA), October 2012

[4] C. Y. Lin, T. T. Mac and L. W. Chuang, "Real-time artistic human face portrait by humanoid robot," 2009 IEEE Control Applications, (CCA) & Intelligent Control, (ISIC), October 2009

[5] J. H. Kim, J. W. Jeon, J. H. Lee and B. S. Yeom, "Migration from RS-485 to EtherCAT for closed loop step motor drive," 2014 IEEE International Conference on Industrial Technology, September 2014

[6] G. Shuai, S. Zhuoyuan, W. Zhiyong and M. M. Islam, "Beckhoff based arm control system design for Elderly Assisting Robot," 2011 IEEE International Conference on Automation and Logistics (ICAL), Chongqing, 2011

[7] G. Zhang, F. Ni, Z. Li and H. Liu, "A Control System Design for 7-DoF Light-weight Robot based on EtherCAT Bus," 2018 IEEE International Conference on Mechatronics and Automation (ICMA), Changchun, 2018, pp. 2169-2174.

[8] S. Lim, H. S. Jung and I. K. Jung, "Real time PC based control of parallel robot," Intelligent Robotics Research Center, Korea Electronics Technology Institute, Bucheon-si, South of Korea, April 2015

[9] G. Cheng and X. Wang, "Design of tri-axial motion servo control systems based on EtherCAT," 2012 IEEE International Conference on Information Science and Technology, June 2012

[10] G. Zhang, F. Ni, Z. Li and H. Liu, "A Control System Design for 7-DoF Light-weight Robot based on EtherCAT Bus," 2018 IEEE International Conference on Mechatronics and Automation (ICMA), Changchun, 2018, pp. 2169-2174.

[11] S. Lim, H. S. Jung and I. K. Jung, "Real time PC based control of parallel robot," Intelligent Robotics Research Center, Korea Electronics Technology Institute, Bucheon-si, South of Korea, April 2015

[12] G. Cheng and X. Wang, "Design of tri-axial motion servo control systems based on EtherCAT," 2012 IEEE International Conference on Information Science and Technology, June 2012

[13] L. Wang, H. Jia, J. Qi, and B. Fang, "The construction of soft servo networked motion control system based on ethercat," In International Conference on Environmental Science and Information Application Technology (ESIAT), July 2010

[14] L. Zhang and Y. Song "Optimal design of the Delta robot based on dynamics" 2011 IEEE International Conference on Robotics and Automation, May 2011

[15] L. Williams II, Ph. D, "The Delta Parallel Robot: Kinematics Solutions Robert" Mechanical Engineering, Ohio University, October 2016

[16] J. D. Gorman, "The application of optimization algorithms in antenna array beampattern synthesis" 2015 26th Irish Signals and Systems Conference (ISSC), July 2015

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 an Associate Professor in National Yunlin University of Science &Technology.

Prof. Chau-Chung Song

He received the M.S. degree in Control Engineering in 1994 and the PhD degree in Electrical and Control Engineering from National Chiao Tung University in 2001 at Taiwan. He is currently a distinguished professor in Department of Aeronautical Engineering, National Formosa University. His research interests include the analysis and control of nonlinear systems the integration application of networking information system, the UAV technology, and the cloud network and Cyber-Physical system (CPS).

Mr. Guan-Yu Jiang

He is a graduate student at department of Electrical Engineering, National Yunlin University of Science & Technology.

Web-based SCADA using MQTT protocol And AES

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 †

Tzu-Yuan Lin
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
Department of Electrical Engineering, 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, m10812021@yuntech.edu.tw, khhsia@yuntech.edu.tw

(leave space here should be 4 lines between *www.university_name.edu* and **Abstract***)*

Abstract

Internet of Thing (IoT) technology is a very popular research topic. Especially in the application of Industry 4.0, it is the most basic and important part. However, related security and application system development often have many problems. For example, the integration of various communication interfaces and data formats, and the rapid establishment of application systems. Therefore, this paper uses MQTT protocol and AES encryption technology to develop a Web-based SCADA system. This Web-based SCADA uses drag-and-drop operation, and users can quickly build a WYSIWYG (What You See Is What You Get) application system. Moreover, with the characteristics of MQTT message transmission, users can choose to transmit the device's message as it is or to process the data before transmitting when creating the SCADA system. And these messages will be encrypted by AES, making the whole system safer and more efficient. Finally, by collecting different communication protocols and integrating multiple communication interfaces, this SCADA can be connected to many PLCs and other equipment, so that the industry can build application systems quickly and at low cost.

Keywords: Internet of Things (IoT), MQ Telemetry Transport (MQTT), AES, Supervisory Control And Data Acquisition (SCADA).

1. Introduction

Internet of Thing (IOT) technology is a very popular research topic recently, especially in the application of Industry 4.0, it is the most basic and most important part. There are many production equipment may be designed without taking into account the function of automatic data collection. Or the equipment has a data collection function, but the equipment manufacturer is unwilling to disclose the communication protocol or data format. All these have caused the industry to be hindered in the progress of Industry 4.0. Therefore, many industries or government agencies have put forward related plans one after another, hoping to quickly solve the problem of equipment networking communication protocols and data format standards. For example, the international automation company Rockwell also provided a complete

.

set of software and hardware solutions in 2019 [1]. The internationally renowned IBM also released the open source visual interface development tool Node-RED [2]. Provides various APIs, including Internet services, and calls using various communication protocols, such as MQTT, MODBUS, OPC-UA, etc. On embedded systems, Node-RED provides the function of controlling GPIO, and uses MQTT or HTTP protocols to communicate with the cloud to build an Internet of Things system.

After solving the communication protocol, how to build this information into an application system is another complicated problem. In the past, it was usually necessary to write code, set up the host, manage the database system, etc., which usually required a lot of manpower and expense. Therefore, how to quickly build an application system with a lower technical threshold and cost has become a research topic for many scholars. For example, NDUKWE, Cherechi [3]; etc., which use LoRa to develop a low-cost SCADA monitoring system for small-scale renewable energy. SINGH [4] and others have built an energy-saving smart home system based on IoT. JAYASHREE [5] and others introduced in detail the relevant knowledge of the enterprise to build the Internet of Things system. From the above examples, we can know that the Internet of Things is an important link in Industry 4.0 and many automation applications.

However, due to the inability to integrate communication protocols and data formats, and the complex construction of visual monitoring systems, Industry 4.0 and even the application of artificial intelligence will be greatly hindered. Therefore, this paper is expected to integrate IoT modules, and through common and open software and hardware, various applications can be quickly built. In addition, the web version of the SCADA system is also used to allow users to quickly build the system. Moreover, the IoT module and SCADA system of this study will also integrate artificial intelligence algorithms, as well as various database connections, various industrial communication standards and other protocols. After users build their own systems, they can also extend the connection with ERP, MES, APS and other systems. So that the application of Industry 4.0 and artificial intelligence can be implemented quickly and at low cost to the industry.

2. System Architecture

In the application of the IoTs, because the communication interface and data format of each IoTs device are different, it is very troublesome to collect these data, and to effectively apply and store these data. Therefore, this paper selects the MQTT protocol as the entire Internet of Things data structure. When the data transmission content of the IoTs module can clearly know the corresponding I/O or sensor data, we can clearly distinguish these data, and can effectively publish and apply. If the data transmission content of the IoT module cannot clearly distinguish the corresponding I/O or sensor, the data transmitted by the IoT device can still be completely transmitted. And use artificial intelligence algorithms to parse out the corresponding I/O or sensor data.

MQTT is a lightweight protocol and a protocol designed for the Internet of Things. Therefore, the network bandwidth it requires is very low, and the hardware resources required are also low. It is very suitable for IoT environments with low power consumption and limited network bandwidth, such as smart home appliances or medical devices. MQTT uses the Publish/Subscribe mechanism to transmit data, which contains 4 main elements, Publisher, Subscriber, Topic, and Broker. Among them, the Internet of Things module is the Publisher, and the information that the Internet of Things module sends out is the Topic. These information Topic are not sent directly to the demand-side Subscriber, but are sent through the forwarding station Broker. Therefore, the Publisher and the Subscriber are not directly connected. Therefore, the information and communication security of the entire Internet of Things system must be controlled through a Broker. At present, most brokers have SSL (Secure Sockets Layer) encryption mechanism when transmitting data. But SSL is too complicated for IoT devices. Therefore, this paper uses AES (Advanced Encryption Standard) technology to ensure the security of all communication and data in the communication of the IoT module and the communication of the relay station (Broker).

In this architecture, we use the WebSocket[6] network transmission protocol, which is often used in Internet applications. WebSocket makes the data exchange between the client application or browser and the server easier, especially allowing the server to actively push data to the client. In the WebSocket API, the browser and the server only need to complete a one-time handshake operation, and a sustainable connection can be established between the two, and two-way data transmission can be carried out. WebSocket and HTTP are different communication protocols. Although both belong to the application layer of the OSI model, the transport layer is also a TCP protocol. But WebSocket and HTTP are not the same communication protocol. HTTP uses the "request and response" mode for communication, while WebSocket uses two-way communication. But in order to make it compatible with the HTTP protocol, WebSocket works through HTTP ports 80 and 443, and supports HTTP proxy and middleware. The WebSocket protocol supports the interaction between a Web browser or other client applications and a Web server, and has better communication efficiency. In order to facilitate real-time data transmission between the client and the server. The server can be implemented in a standardized way without the client requesting first. And to allow messages to pass back and forth to each other while remaining connected. In this way, a two-way continuous conversation can be carried out between the client and the server. Communication can be done through TCP port 80 or 443. This is beneficial in an environment where a firewall prevents non-Web network connections, with better flexibility, higher security, and efficiency. The entire communication architecture is shown in Fig. 1.

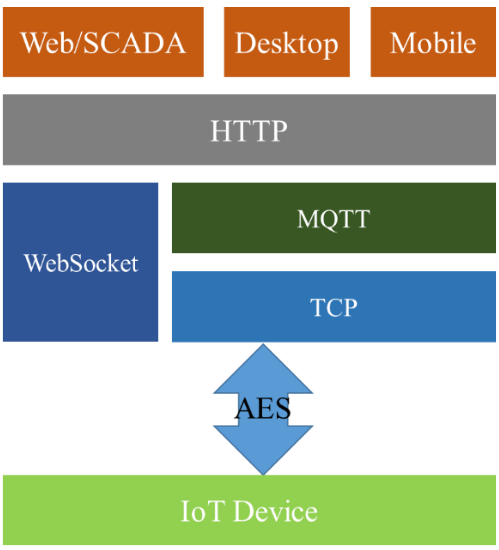


Fig. 1. Block diagram of the communication architecture between IoT devices and application systems.

2.1. *MQTT data structure*

MQTT allows devices and devices to exchange data with each other through topics, so theme design is very important. Since this system will be connected to many controllers, sensors, and actuators, in the design principle of Topic, consider the use of a four-level design, the format is:

Table 1. MQTT communication data structure.

MQTT serial number/IoT device name/I/O name/data type.

The description of this structure is as follows.

- MQTT sequence number layer: The first layer is named after the device ID, called MQTT ID, starting from 01.
- Device name layer: Since there are many Internet of Things devices or data sources of other devices in the system, this layer can be used to distinguish different devices or data sources. For example, this paper uses the Internet of Things module to include An STM32 terminal module and an ESP32CAM can be named stm32 and esp32 respectively, as shown in Fig. 2.
- Input and output layer: This layer can be used to represent data output or input, and are named after control and data respectively.
- Data type: The last layer refers to the format of the subscription data, including sensors, connection status, video streaming, etc., which are respectively sensor, connection, and stream.

The complete MQTT message of this IoT module example is shown in Table 2 to Table 3.

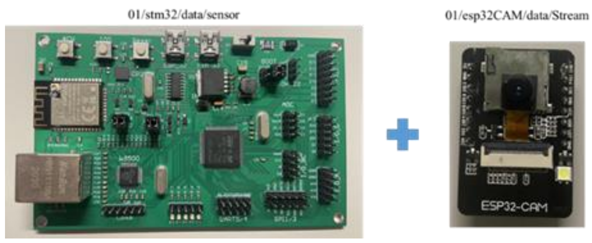


Fig. 2. Examples of IoT modules (STM32 control board and ESP32 imaging device).

Table 2. MQTT communication data structure_data.

Topic Name	content

01/stm32/data/sensor	6-channel ADC, temperature and humidity and I/O status on the STM32 module.
01/stm32/data/connection	Wi-Fi and LAN connection information, such as Wi-Fi SSID and password.
01/esp32CAM/data/stream	ESP32CAM image

Table 2. MQTT communication data structure _control.

Topic Name	Payload	Function
01/stm32/control	setpagedata	Turn on sending Wi-Fi data.
01/stm32/control	statusON	Turn on sending sensor data.
01/stm32/control	statusOFF	Turn off sending all data.
01/stm32/control	reset_io	Reset all I/O.
01/stm32/control	OUTPUT1_ON~ OUTPUT5_ON	Set I/O to high level.
01/stm32/control	OUTPUT1_OFF~ OUTPUT5_OFF	Set I/O to low level

Table 3. MQTT communication data structure_Broker.

Topic Name	content
$SYS/broker/load/messages/received/1min	The amount of messages received by the Broker per minute.
$SYS/broker/load/messages/sent/1min	The amount of messages sent by Broker per minute.
$SYS/broker/clients/total	The total number of clients.
$SYS/broker/clients/connected	The number of currently connected clients

In the design of the SCADA system, we use Web-base and drag-and-drop operation. And in the design of the screen components, try to adopt the design close to the actual object, and through the parameter setting method, to make the operation simple. After the user's screen is designed, what is presented is the final used screen. Fig. 3 is the screen of SCADA in the design mode of this thesis. The screen is mainly divided into three blocks, namely:

(A) Instrument icon area: This area provides icons of common industrial instruments and equipment provided by the system.

(B) Parameter setting area: This area is for setting the communication parameters, display parameters, MQTT related settings, etc. of the instrument icon.

(C) SCADA design area: Using the drag-and-drop method, place the icons to be displayed and controlled in this area, and you can quickly build the WYSIWYG monitoring system.

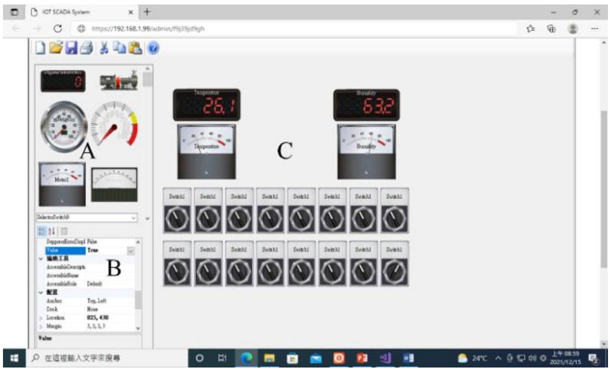


Fig. 3. SCADA design mode screen.

3. Experimental Results

This paper has developed a Web-based SCADA system that allows users to quickly establish a SCADA monitoring system by drag-and-drop operation and parameter setting. We actually apply this system to industrial production lines. Fig. 4 shows that the original production equipment does not have the function of network communication, and needs to monitor the temperature and humidity of the environment. Therefore, we use a self-developed IoT module to collect the signal with an external sensor, and then send it back to the large-scale LED signage and SCADA system on site. Fig. 5 shows the results of the SCADA implementation developed in this paper. Considering that the entire screen information will not be too confusing, the data displayed in the center of the screen can be selected by the user. Or when there is a problem with the field device, the data of the problematic device will be displayed automatically.

Fig. 4. SCADA system application environment and IoT installation status.

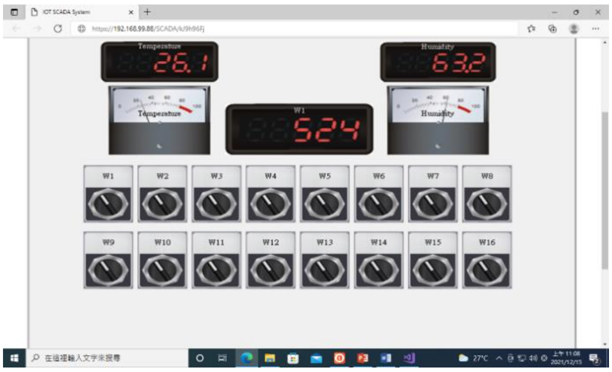


Fig. 5. SCADA screen.

4. Conclusions

This thesis uses the MQTT protocol to develop a Web-base SCADA system, which has built many instruments and devices commonly used in industry. And it allows users to quickly build a SCADA system by drag-and-drop and parameter setting. In the connection with IoT devices, we support multiple communication protocols. If the device data can be clearly distinguished and distinguished, these data will be processed into MQTT messages and used on the SCADA system we designed. If the content of the device's data cannot be clearly known, the SCADA system will still store it in the database and try to parse the content. Finally, we actually applied this SCADA system to the production line, and the overall operation was in good condition. In the future, we will add image streaming and recognition functions. In addition to providing general data monitoring, this system can also perform large-scale image streaming data transmission and image recognition functions. Let this system be applied to more industries.

References

1. Rockwell also provided a complete set of software and hardware solutions
2. IBM "Nodered"
3. NDUKWE, Cherechi; IQBAL, M. Tariq; KHAN, Jahangir. Development of a Low-cost LoRa based SCADA system for Monitoring and Supervisory Control of Small Renewable Energy Generation Systems. In: 2020 11th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON). IEEE, 2020. p. 0479-0484.
4. SINGH, Prithvi Pal; KHOSLA, Praveen Kumar; MITTAL, Mamta. Energy conservation in IoT-based smart home and its automation. Energy conservation for IoT devices, 2019, 155-177.
5. JAYASHREE, L. S.; SELVAKUMAR, G. Getting Started with Enterprise Internet of Things: Design Approaches and Software Architecture Models. Springer Nature, 2020.
6. WebSocket,2021/12/08

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

Mr. Tzu-Yuan Lin

He was born in 1997. He received the B.S. degrees in Automation Engineering from National Formosa University of Science & Technology Yunlin, Taiwan, in 2019, and M.S. degrees in Electronic Engineering from National Yunlin University of Science & Technology, Yunlin, Taiwan, in 2021. Currently, he is a software engineer in Smart Creating Co., Ltd. His research interests include Artificial Intelligence, Internet of Things, SCADA, Embedded Systems and Cloud Computing.

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.

Smart Identification System of Teaching-type Autonomous Vehicles

Chun-Chieh Wang

Department of Automation Engineering and Institute of Mechatronoptic Systems, Chienkuo Technology University, Taiwan
(Tel: 886-4-7111111ext3350, Fax: 886-4-7111164)
jasonccw@ctu.edu.tw / bigjasonwang@gmail.com

Abstract

To improve the teaching efficiency of autonomous driving image recognition technology, a self-driving car intelligent identification system suitable for teaching has been developed in this article, so that tea chers and students can easily use this system for experiments. As for the main controller of the car body is the Raspberry Pi microcomputer processor. It works with Python for image processing. Image processing techniques include grayscale, binarization, morphology, image cutting, etc. To realize the fu nction of self-driving cars, there are four main functional tests in the scenario setting. It includes road identification, conversion of lane turning arc into front wheel turning angle, intersection identific ation, and traffic light identification. The experimental results confirm that the developed smart identif ication system and the experimental environment planning are helpful for autonomous driving related t eaching and can enhance students' willingness to learn.

Keywords: Smart Identification System (SIS), Raspberry Pi, Python, Teaching-type Autonomous Vehicles (TAV)

1 INTRODUCTION

Today's image recognition technology has been used in many engineering fields. Especially the image recognition combined with the automatic driving system can bring greater traffic convenience to people. Related research has also shown good results in the past 10 years.

In 2010, Lin B.Z. [1] proposed high speed detection and recognition of traffic light. The main purpose of this thesis is to quickly detect and identify traffic lights to obtain real-time traffic information, which can effectively assist and remind drivers. In 2011, Chen K.M. [2] proposed image processing technique for road detection with depth information. First, the parallel dual-lens architecture is used to synchronize the image processing of the cameras on both sides, and to identify the road surface and lane markings in front of the car. Second, the effective feature information after image processing and identification is converted from a two-dimensional image plane to a three-dimensional spatial coordinate distribution. Finally, the road feature and depth information of the image in front of the car is provided to the driver or the driving assistance control system. In 2016, Chen J.H. [3] proposed street sign detection on Google street view images. This research uses Google Street View services and route planning to build a web platform with JavaScript, CSS and HTML to obtain street view image materials. Python is used to implement algorithms in filtering, edge detection, morphological image processing, and parallel computing (multiprocessing) in order to develop a set of automated street sign detection algorithm. In the same year, Felipe Jiménez et al. (2016) proposed an algorithm based on vehicle dynamics mathematical model to improving the lane reference detection for autonomous road vehicle control.

Based on the above, in order to develop teaching-type autonomous vehicles, this article will use smart identification technology to complete road identification, conversion of lane turning arc into front wheel turning angle, intersection identification, and traffic light identification.

2 TEACHING-TYPE AUTONOMOUS VEHICL ES (TAV)

2.1 TAV Platform

The car body in this article uses the steering gear to drive the direction of rotation of the vehicle. The way the differential gear is driven can prevent the tires from being able to drive continuously in the stuck state. In addition, a 360-degree small servo motor is used to control the forward and backward functions. In order to stabilize the power supply, this article will extend a power cord from the car body in Fig. 1 so that the test process can be kept in the best condition. A 5V, 2.5A transformer can provide stable voltage and current to the Raspberry Pi. The two upper and lower lenses on the front of the car are used to identify the road and identify the traffic lights, as shown in Fig. 2. In addition, there are parts structure of the differential and the steering gear at the chassis.

Fig.1. Power cable on the side of the car

Fig.2. Two cameras on the front of the car

In this paper, the rear drive method is used to make the car body generate thrust from the rear. The pushed body controls the direction of rotation, as shown in Fig. 3. The advantage of rear drive is that the body's center of gravity is stable, and there will be no internal wheel difference due to the body's interrogation. When the turning speed is slightly higher, over-turning will occur (the red range in Fig. 4). Therefore, this article uses a servo motor that can rotate 360 degrees to control the rotation speed. We use the feature of adjustable angle to control the speed of forward and turning.

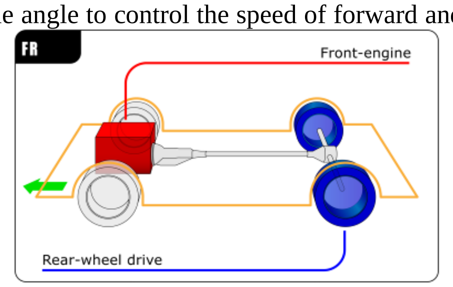


Fig.3. Schematic diagram of rear-drive driving

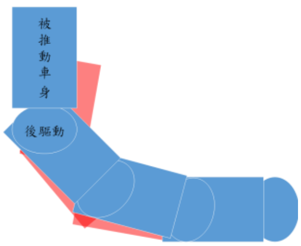


Fig.4. Excessive turning condition of rear drive

2.2 Controller and Peripheral Equipments

The main controller in this article is Raspberry Pi, which is a single-chip computer based on the Linux operating system, as shown in Fig. 5.

Fig.5. Raspberry Pi3 Model B

Fig.6. 360-degree rotating servo motor-SG90

Fig7. 180 degree rotation high torque servo motor-MG996R

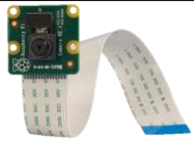


Fig.8. Raspberry Pi Camera V1.3

Fig.9. Logitech C920

The steering of the car is controlled by the 360-degree rotating servo motor-SG90 and the 180-degree high torque servo motor-MG996R, as shown in Fig. 6 and Fig. 7. As for road recognition and traffic light recognition, Raspberry Pi Camera V1.3 and Logitech C920 cameras are used respectively, as shown in Fig. 8 and Fig. 9.

3 SMART IDENTIFICATION SYSTEM (SIS)

3.1 Setting Method of Road Coordinate Points

In order to keep the car at the center of the road, this article sets the coordinate position on both sides of the road. We use the coordinate position of the road width to obtain the center point to ensure that the route of the car is in the center of the road. When encountering a turn on the way, the center point position will be obtained according to the road width in the lens to control the turning range of the steering gear.

Get the coordinate points x_1 and x_3 on both sides of the track to calculate the center point (Q) of the road. Apply formula (1) to obtain the Q point coordinates of Fig. 20. Lock the coordinate points on both sides of the road, which are $A(x_0,y_0)$, $B(x_1,y_1)$, $C(x_2,y_2)$, $D(x_3,y_3)$ four coordinate positions.

In order to prevent the vehicle from exceeding the driving track, the controller needs to determine the turning range early, so this article additionally sets the coordinates A and C. Coordinates B and D are the positions where the vehicle needs to obtain the center point. When the subtracted distance of the two points B and D cannot be directly expressed as the coordinates of the center point, the x_1 coordinate needs to be added to the correct position of the Q point.

When the vehicle encounters a situation that needs to turn while driving, use the arc sine theorem (Fig.21) to control the servo motor to adjust the turning wind direction of the steering gear. The green area in Fig.22 is the appropriate turning range. Use formula 2 to obtain the angle value of θ^o to adjust the amplitude of the steering gear. According to Fig.22, $\theta_L{}^o$is the amplitude of the left wheel steering and $\theta_R{}^o$ is the amplitude of the right wheel steering. The θ^o generated by the coordinate changes on both sides may not be the same. Therefore, in order to obtain the correct $\theta_0{}^o$, formula 3 needs to be applied to ensure the correct position of $\theta_0{}^o$. Fig.23 is the coordinate position on the road. According to the curvature and direction of the road, an appropriate θ^o angle is generated to control the vehicle to drive along the road.

$$Q = x_1 + \frac{x_3 - x_1}{2} \tag{1}$$

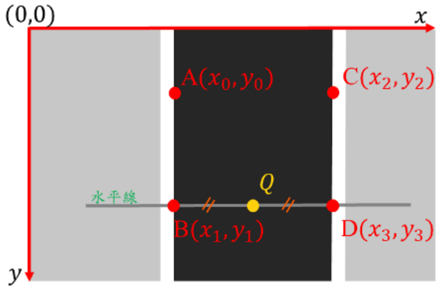


Fig.20. The central position of road recognition

$$\theta = tan^{-1}\frac{|x_n - x_{n+1}|}{|y_n - y_{n+1}|} \tag{2}$$

$$\theta_0 = \frac{\theta_L - \theta_R}{2} \tag{3}$$

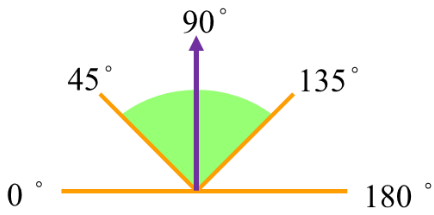


Fig.21. Servo motor and the range of arc sine theorem

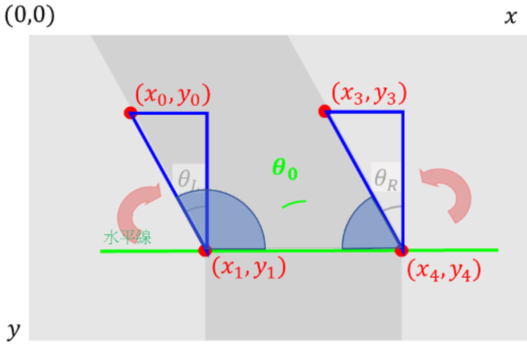


Fig.22. Angle range of θ_L° and θ_R°

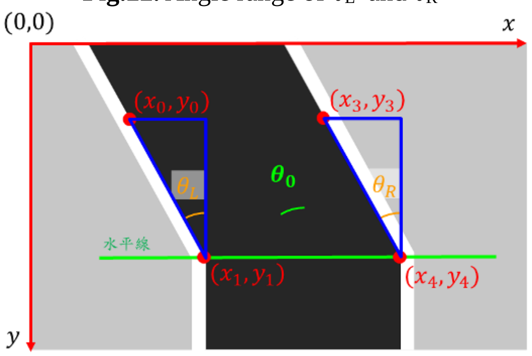


Fig.23. Schematic diagram of the coordinates of the road turning

3.2 Control Flow Design

In order to identify the image results of roads and traffic signs, turn on the lens for image recognition. First determine whether the image of the lens is the entrance. Fig. 24 is the system flow chart.

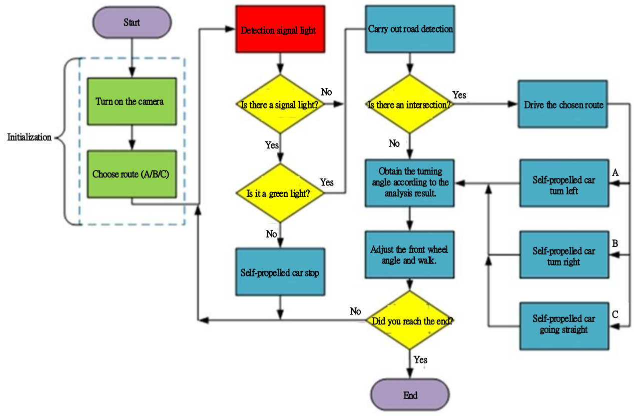


Fig.24. System flow chart

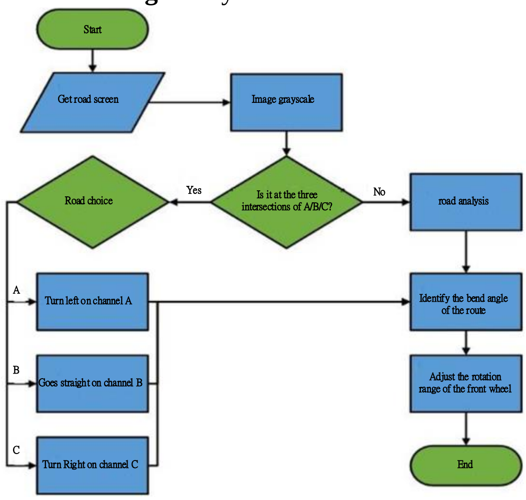


Fig.25. Road identification flow chart

When the road recognition lens is turned on, the road image is first taken and the image is grayed out. Make sure that the driving route is correct if there are three intersections A, B, and C. Fig. 25 is a flowchart of road identification. Fig.26 is the flow chart of traffic light identification.

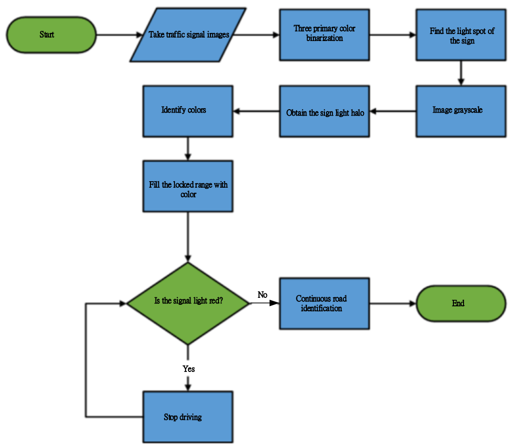


Fig.26. Traffic light identification flow chart

3.3 Road Boundary Detection

In order to make self-driving road identification more accurate, the following first plans for the judgment conditions of road identification. First of all, the judged turning method is to transfer the image of the lens to the Raspberry Pi to obtain the boundary points of the road and then calculate the road width. The judgment can be divided into the calculation formula of the right boundary (line_R) and the left boundary (line_L). Taking Fig.27 as an example, if the X coordinate of line_R minus line_L is less than 120, it means that the captured boundary area appears as a straight line. If the X coordinate position of line_R is less than 10, it means that the track has deviated and needs to be turned right. If the X coordinate of line_L is greater than 120, it means that the car is already too right, and it must turn left to lead back to the center of the road.

When encountering fork roads such as Fig.28, Fig.29, Fig.30, add 1 from 0 through the counter. For example, assuming that the route the car is going to take is A2, in Fig.31, it can be judged whether to turn or go straight at the next intersection.

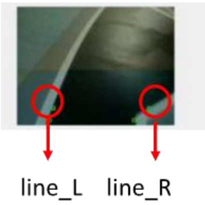


Fig.27. Road boundary detection point

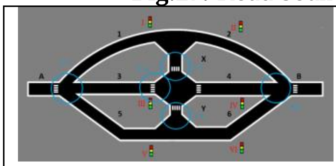


Fig.28. Straight curve detection position

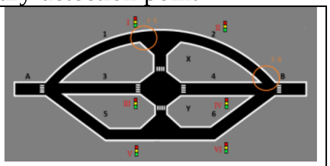


Fig.29. Curved corner detection position

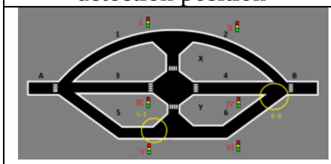


Fig.30. T-shaped corner detection position

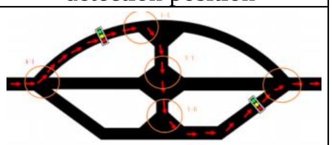


Fig.31. A2 route curve detection position

4 EXPERIMENTAL SCENARIO DESIGN

The position and number of traffic lights in this article are as shown in Fig.32.

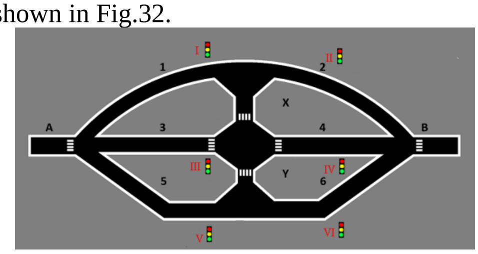


Fig.32. Road site simulation map (6 traffic lights)

5 EXPERIMENTAL RESULTS

Please refer to the following URL directly for the experimental results. https://youtu.be/3TKsgnAGx_4. In order to prevent the road recognition function from being affected by the lack of power, this article directly uses the mains connection transformer to provide power to the car (the line connected to the car in the film is the power line, not a remote control line). Fig.33-Fig.50 are screenshots (including schematic diagrams) of self-driving walking pictures of 9 routes.

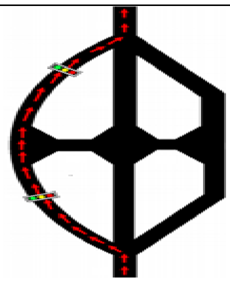


Fig.33. Schematic diagram of route A

Fig.34. Actual walking screen of Route A

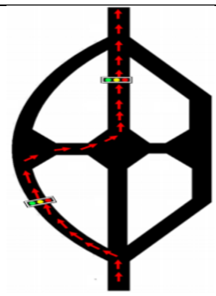


Fig.35. Schematic diagram of route A1

Fig.36. Actual walking screen of Route A1

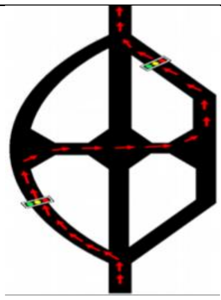


Fig.37. Schematic diagram of route A2

Fig.38. Actual walking screen of Route A2

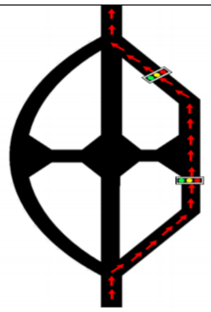


Fig.39. Schematic diagram of route D

Fig.40. Actual walking screen of Route D

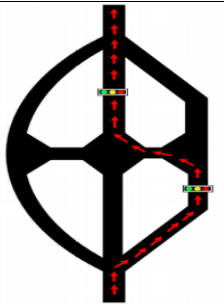


Fig.41. Schematic diagram of route D1

Fig.42. Actual walking screen of Route D1

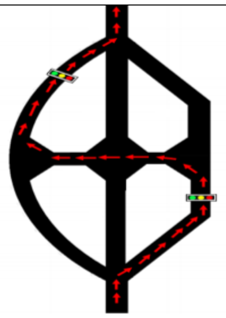


Fig.43. Schematic diagram of route D2

Fig.44. Actual walking screen of Route D2

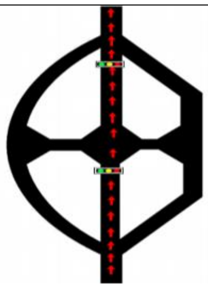


Fig.45. Schematic diagram of route S

Fig.46. Actual walking screen of Route S

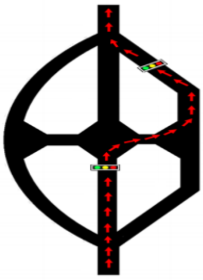


Fig.47. Schematic diagram of route S1

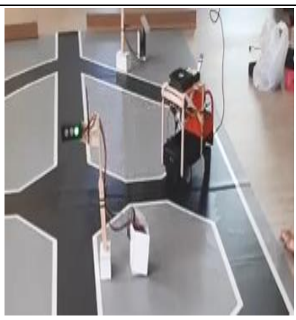


Fig.48. Actual walking screen of Route S1

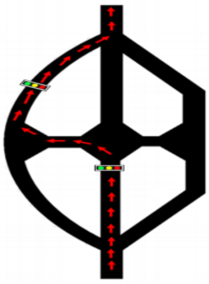


Fig.49. Schematic diagram of route S2

Fig.50. Actual walking screen of Route S2

6 CONCLUSION

This article successfully developed a teaching-type self-driving vehicle with smart identification systems based on the structure of the physical vehicle. At the same time, the experimental field of the teaching version is planned, in which 9 different paths are designed to allow self-driving cars to perform visual identification based on different paths. Moreover, we use the smart visual recognition system to complete the autonomous driving of the self-driving car on the simulated road. Not only that, it can also successfully recognize the current state of the road based on the intelligent visual recognition system, and perform the functions of going straight and turning.

To realize the function of self-driving cars, there are four main functional tests in the context setting. It includes road identification, conversion of lane turning arc into front wheel turning angle, intersection identification, and traffic light identification. The experimental results confirm that the developed smart identification system is helpful for teaching and can enhance students' willingness to learn.

REFERENCES

[1] Lin B.Z. (2010), High Speed Detection and Recognition of Traffic Light, Master Thesis of National Sun Yat-sen University.

[2] Chen K.M. (2011), Image Processing Technique for Road Detection with Depth Information, Master Thesis of National Sun Yat-sen University.

[3] Chen J.H. (2016), Street Sign Detection On Google Street View Images, Master Thesis of National Chung Hsing University.

[4] Felipe Jiménez, Miguel Clavijo, José Eugenio Naranjo, and Óscar Gómez (2016), Improving the Lane Reference Detection for Autonomous Road Vehicle Control, Journal of Sensors, Vol. 2016, Article ID 9497524.

[5] OpenCV modules

Authors Introduction

Prof. Chun-Chieh Wang

He is is a professor of Department of Automation Engineering and Institute of Mechatronoptic Systems of Chienkuo Technology University. 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.

Automatic Anti-Lock Brake System for Anti-Rollover Control of Autonomous Heavy-Duty Truck

Chian C. Ho, Riki Umami Sanaz Ulfitria
Department of Electrical Engineering, National Yunlin University of Science and Technology
Douliou, Yunlin County 64002, Taiwan
E-mail: futureho@yuntech.edu.tw, rikiumamisanazulfitria@gmail.com
www.yuntech.edu.tw

Abstract

In recent years, there are more and more rollover accidents about autonomous heavy-duty trucks or autonomous heavy-duty forklifts in intelligent airports or seaports. These accidents lead to the hot research field about the prevention of rollovers in advance for autonomous vehicles, especially in autonomous heavy-duty trucks or forklifts. This paper develops an automatic Anti-lock Brake System (ABS) as one way to stabilize the autonomous vehicles. Then, through monitoring both vehicle's and wheel's speeds, this paper helps the autonomous vehicle keep stability control with automatic ABS even when the wheels halt or slip on the road. This paper adopts TruckSim to model the vehicle safety dynamics and MATLAB/Simulink to simulate the vehicle stability control. Experimental results show that the elaborate automatic ABS proposed by this paper can smoothly keep the vehicle stable and tractive even under dangerous road conditions of sharp corner or hairpin turn

Keywords: Anti-lock Brake System (ABS), anti-rollover, autonomous truck.

1. Introduction

In the past few years, autonomous heavy-duty trucks or heavy-duty forklifts have been more and more popular in intelligent airports or seaports, but rollover accidents also get more and more frequent. Research demonstrates that autonomous heavy-duty trucks are likely harmful to road safety during logistics, such as rollovers, shimmy, or jackknives due to the high mass center, comparatively big size, or relatively small base [1]. The most significant hazard to road safety is the rollover of autonomous heavy-duty trucks, which can lead to catastrophic results and significant losses [2].

A vehicle rollover is a phenomenon of vehicle instability caused by extreme steering or external excitation of the road. In the United States in 2016, there are over 6 million vehicle crashes. Among these crashes, 17.9% are vehicle roll-off deadly incidents, 8.5% are from massive lorries and buses [3].

Unconsidered rollover events are typically caused by high speed or extreme roll due to unexpected track modifications. If a vehicle crosses a curved road because of its centrifugal force, it leans outside the curve. The roll-up is possible if the autonomous vehicle misjudges the curve sharpness and keeps excessively fast because of his lateral acceleration tolerance [4].

The chance of a rollover accident might be reduced by a combined rollover risk assessment and active roll control under difficult driving situations. When rollover dangers can be predicted, information may be given to the autonomous vehicle or the stability controller. Proper reaction from the autonomous vehicle or the stability controller can prevent the rollover accident effectively and efficiently. The active rotating control is the best solution if the autonomous vehicle does not react appropriately [5]. Many researchers are analyzing these stability issues and improving the autonomous heavy-duty trucks or forklifts to lessen the occurrence probability of the rollover.

Because automatic Anti-lock Brake System (ABS) is based on slip control, the slip ratio of the wheel's angular speed to the vehicle speed is determined. The slip ratio may be adapted to supply the wheels with braking power using other methods, for example, the slip ratio may be estimated by a specific speed sensor. This paper proposes to adopt automatic ABS to keep stability control of autonomous heavy-duty trucks or forklifts. So the autonomous vehicles can always run stable and tractive even in the double lane change scenario or circle lane scenario.

This paper is organized as follows. Section 2 explains the definition of autonomous heavy-duty vehicle modeling. Then, in Section 3, it explains the simulation and result in MATLAB/Simulink and TruckSim. Section 4 discusses the simulation result, and the last section remarks the conclusion and future direction of the research.

2. Vehicle Dynamics Modeling

The modeling of significant rolling action in the middle of the mass in autonomous heavy-duty trucks has special restrictions because of its lengthy wheel lowering, dividing the spring weight between the front and back. Considering the front-to-back connective bobbin, the optimum torsional bar works with the stiffness of the torsion and has no mass for the front and rear spring systems [5], as indicated in Figure 1.

In view of the equation of different degree of freedom, the principle can be defined and derived as follows:

Longitudinal motion:

$$m(\dot{u} - vr) = F_{xT} - F_r - 2F_{Y1}\sin\delta_f \qquad (1)$$

Lateral motion:

$$m(\dot{v} + ur) - m_{sf}h_f\ddot{\varphi}_{sf} - m_{sr}h_r\ddot{\varphi}_{sr} = 2F_{Y1}\cos\delta_f + 2F_{Y2} \qquad (2)$$

Yaw motion:

$$I_z\dot{r} = 2aF_{Y1}\cos\delta_f + 2bF_{Y2} \qquad (3)$$

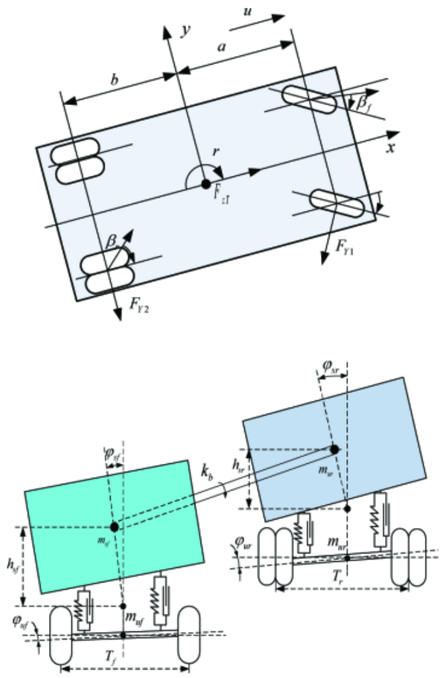


Fig. 1. Dynamics model of the autonomous truck.

Roll motion of front sprung mass:

$$I_{xf}\ddot{\varphi}_{sf} = -k_f(\varphi_{sf} - \varphi_{uf}) - l_f(\dot{\varphi}_{sf} - \dot{\varphi}_{uf}) + m_{sf}h_f a_y + m_{sf}h_f\varphi_{sf} + k_b(\varphi_{sr} - \varphi_{sf}) \qquad (4)$$

Roll motion of rear sprung mass：

$$I_{xf}\ddot{\varphi}_{sf} = -k_r(\varphi_{sr} - \varphi_{ur}) - l_r(\dot{\varphi}_{sr} - \dot{\varphi}_{ur}) + m_{sr}h_r a_y + m_{sr}h_r\varphi_{sr} + k_b(\varphi_{sr} - \varphi_{sf}) \qquad (5)$$

Roll motion of front unsprung mass：

$$2F_{Y2}h_c + m_{uf}(h_{uf} - h_{cf})a_y = -m_{uf}g(h_{uf} - h_{cf})\varphi_{uf} + k_{uf}\varphi_{uf} + k_f(\varphi_{sf} - \varphi_{uf}) - l_f(\dot{\varphi}_{sf} - \dot{\varphi}_{uf}) \qquad (6)$$

Roll motion of rear unsprung mass：

$$2F_{Y2}h_c + m_{ur}(h_{ur} - h_{cr})a_y = -m_{ur}g(h_{ur} - h_{cr})\varphi_{ur} + k_{ur}\varphi_{ur} + k_r(\varphi_{sr} - \varphi_{ur}) - l_r(\dot{\varphi}_{sr} - \dot{\varphi}_{ur}) \qquad (7)$$

Lateral acceleration at the center of mass of the vehicle:

$$a_y = (\dot{v} + ur) \tag{8}$$

Longitudinal displacement of the vehicle：

$$\dot{X} = u\cos\psi - v\sin\psi \tag{9}$$

Lateral acceleration at the center of mass of the vehicle:

$$\dot{Y} = u\sin\psi - v\cos\psi \tag{10}$$

Where *m* refers to the total mass of the autonomous heavy-duty vehicle, the front and rear axles are indicated by *{f,r}*. m_s and m_u denote the equivalent sprung and unsprung masses, respectively, the longitudinal distances between mass-center-to-front-axle and mass-center-to-rear-axle are a and *b*, correspondingly. *h* implies the length between the sprung mass center and the rolling center, so h_u and h_c are measured upwardly of the center of rolls and the center of the unsprung mass, respectively, from the road. F_{xT} represents the longitudinal force of the tires and F_r means the wheel rolling resistance, then for lateral forces of the front and rear axles are the pneumatics, F_{Y1} and F_{Y2}. g is the gravitational acceleration, I_z is the yaw inertia of the heavy-duty vehicles, and I_x is the roll inertia of the sprung mass. *u* is the longitudinal speed, and *v* denotes the lateral velocity. k and k_u are the suspension and the unsprung mass of the equivalent roll stiffness coefficients, respectively. *l* is the equivalent roll damping coefficient of the suspension, *kb* is the torsion stiffness coefficient of the vehicle frame, *r* denotes the yaw rate of the sprung mass, and δ_f is the front-wheel steering angle. φ_u and φ_s are the roll angles of the sprung and the unsprung masses, respectively. *X* is the longitudinal displacement, *Y* is the lateral displacement, and *ψ* denotes the heading angle.

2.1. *Static Rollover Threshold*

Static Rollover Threshold (SRT) is a maximum lateral acceleration criterion to prevent the vehicle from rolling out. The autonomous truck is at the initial positioned horizontally. The platform is then slowly rotated till the tires become lost on the platform surface. Between the inclining angle between the platform and the ground, the tangent formed is the SRT value. The official definition of SRT is:

$$SRT = \frac{a_t}{g} \tag{11}$$

$$SRT = \frac{T}{2h} - \frac{\Delta_y}{h} \tag{12}$$

2.2. *ABS control*

ABS is a system for electronic safety scheme which checks and monitors the speed of the wheel during braking. The relative sliding between the wheel and road surface may be assessed in the braking process by the slip ratio [6]:

$$S = \frac{V - \omega.\ \mathrm{r}}{V} \text{ x } 100\% \tag{13}$$

Where:
ω = angular velocity of the wheel
r = the effective wheel radius.

With the increasing of the brake force, the tire slip ratio raises the braking force coefficient and finally reduces the slip ratio. At the same time, the coefficient of lateral force gradually diminishes. With the glitch ratio reaching S_c, peaks of braking and lateral strength are relatively high. It also ensures the ideal braking performance, and avoids sideslip. By controlling the hydraulic pressure in brake lines, ABS may keep the ratio of slip around S_c [6].

2.3. *Lateral Load Transfer Ratio (LTR)*

A typical LTR is estimated using the data received at a particular moment. It looks like capturing a snapshot of a dynamic system. Analytical analyses and experimental data help define the rollover threshold based on the predicted LTR values. When the threshold is set to be sensitively low, the LTR will alert or trigger a rollover prevention system even during regular and safe driving period. When the threshold is set to be slightly high, prevention measures might activate too late to prohibit a rollover of the vehicle [7].

$$LTR = \frac{F_{zR}}{F_{zL}} - \frac{F_{zL}}{F_{zR}} \tag{14}$$

This index in (14) uses vertical pneumatic forces, F_{zL} and F_{zR}. LTR describes vehicle rollover when lifting from the ground as either on the left or right side of the vehicle. LTR varies from –1 to 1, where –1 and 1 refer to either the left or right vehicle tires losing contact with the ground, and 0 refers to equal vertical forces on both sides of the vehicle (zero rolls) [7].

3. Simulation Results

Based on the co-simulation of MATLAB/Simulink and TruckSim, modeling the truck dynamics in the TruckSim simulator with several significant parameters, as shown in Figure 2, can make it easier for the developer to perform the evaluation and analysis. Each parameter in TruckSim simulator can be captured and calculated by MATLAB/Simulink.

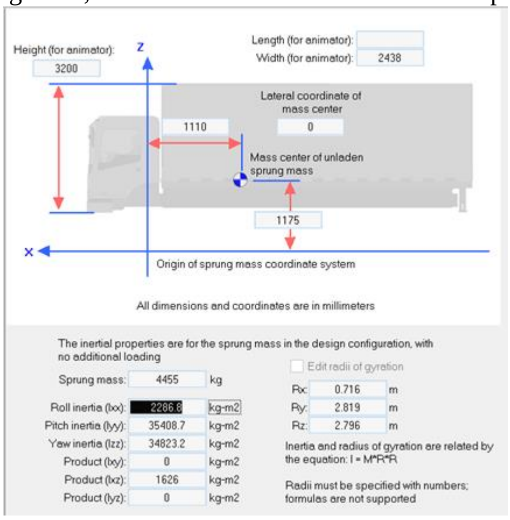


Fig. 2. Vehicle dynamics modelling by TruckSim.

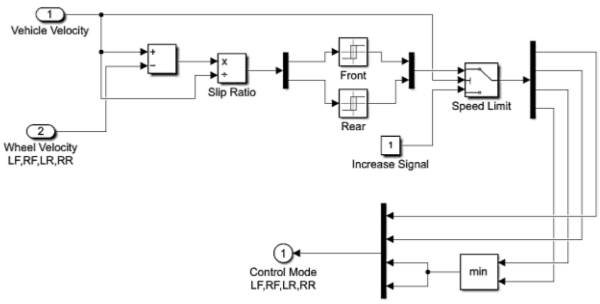


Fig. 3. ABS control model in MATLAB/Simulink.

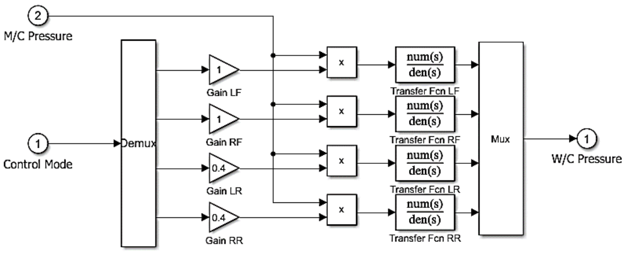


Fig. 4. Brake actuator model in MATLAB/Simulink.

Brake actuators are the instruments that turn a compressed air force into a mechanical force within the vehicle or air reservoir of the truck that triggers the cadence brake. Figure 4 shows MATLAB/Simulink model for brake actuators.

Table 1: Vehicle dynamics model of vehicle 1

Properties	Value	Unit
Unsprung mass	570	kg
Axle roll & yaw inertia	350	kg-m^2
Left spin inertia	10	kg-m^2
Right spin inertia	10	kg-m^2
Wheel center height	510	mm
Center of gravity	2030	mm
Sprung mass origin	510	mm

Table 2: Vehicle dynamics model of vehicle 2

Properties	Value	Unit
Unsprung mass	735	kg
Axle roll & yaw inertia	285	kg-m^2
Left spin inertia	20	kg-m^2
Right spin inertia	20	kg-m^2
Wheel center height	530	mm
Center of gravity	1863	mm
Sprung mass origin	530	mm

For comprehensive comparison, this experiment designs two models of two different vehicle dynamics for autonomous heavy-duty truck simulation, which are listed in Table 1 and Table 2, respectively.

A series of simulation results are performed and shown in Figures 5-10. Figure 5 shows the lateral tracking of the vehicle running on some straight or curved road situations. It means that the vehicle may track the trajectory to recover the vehicle stability and then follow some curved road situation with elaborate longitudinal speed control, as dynamically depicted in Figure 6. In Figure 7, the vehicle's pitch angle begins oscillating at the second of 13 while the variation of lateral tracking becomes larger, as shown in Figure 5. This is because the vehicle expects to decrease its longitudinal speed by proposed automatic ABS, as shown in Figure 6, and rotate its steering angle to follow the curved road situation, as shown in Figure 5. In Figure 7, the vehicle's pitch angle has been oscillating during the second of 13 to 22, and its ripple drops from the degree of -0.540 to the degree of-0.590 because of the vehicle speed’s boost-up since the second of 22. Then the yaw angle of sprung mass in figure 8 has kept constant until the second of 13. Afterward, the yaw angle of sprung mass has been increasing up until the second 23.6 to make the vehicle follow the curved road situation, as shown in Figure 5. Finally, Figures 9 and 10 verify that the vertical tire forces of vehicle 1 and vehicle 2 has kept

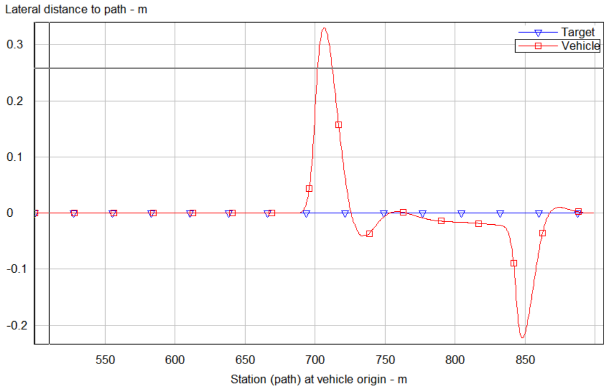


Fig. 5. Vehicle’s lateral tracking under some road situation.

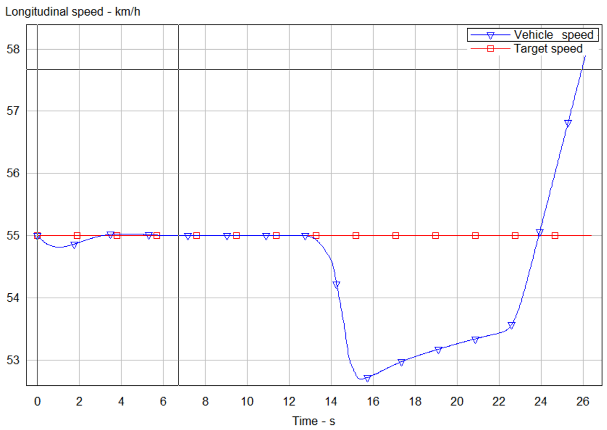


Fig. 6. Longitudinal speed and mass center.

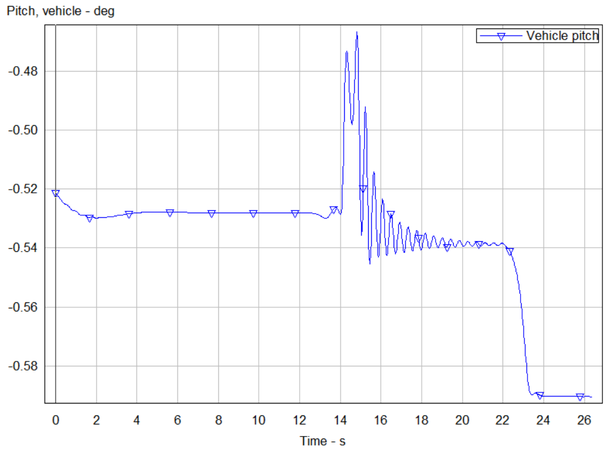


Fig. 7. Pitch angle of sprung masses.

the symmetry variation between the vertical tire force of L and R to ensure the vehicle’s four tires to contact closely with the road surface all the time, through the proposed automatic ABS with confined slip ratio.

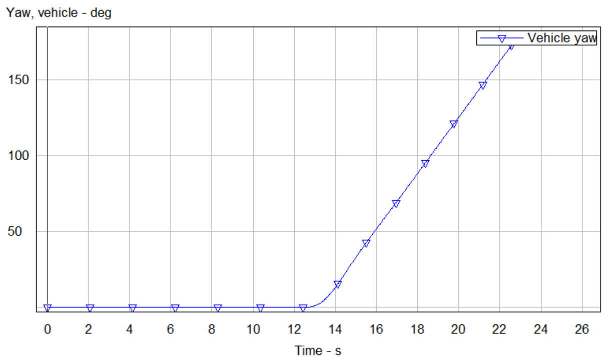


Fig. 8. Yaw angle of sprung masses.

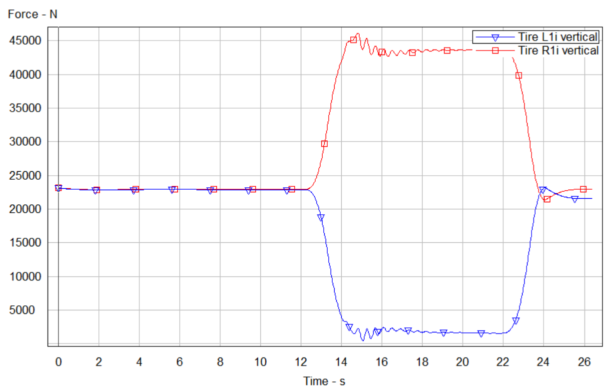


Fig. 9. Vertical tire forces of vehicle 1.

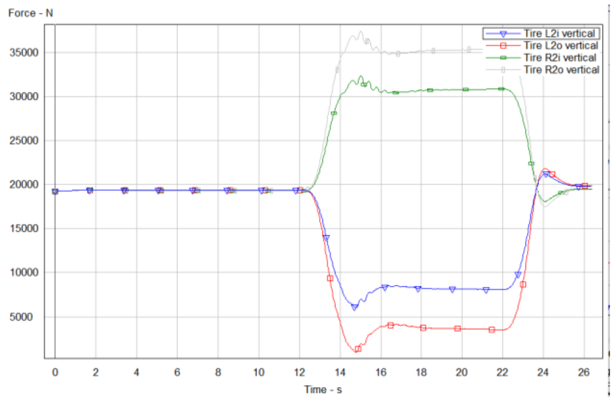


Fig. 10. Vertical tire forces of vehicle 2.

4. Conclusions

In this paper, the integrated stability control scheme of autonomous heavy-duty trucks or forklifts increases the stability by proposed automatic ABS. By increasing the brake force elaborately, the tire slip ratio raises the braking force coefficient and ultimately reduces the slip ratio. ABS plays a crucial part in regulating the speed of the wheel on slick and loose surfaces. Most autonomous heavy-duty trucks in such conditions tend to lose control of stability. The ABS has four sensors of wheel speed, vehicle speed, electrical and hydraulic valves. The simulation result shows that the proposed stability control by automatic ABS can work well on two models of two different vehicle dynamics even under dangerous road conditions of sharp corner or hairpin turn.

References

1. Z. Tianjun, "Research on improved TTR rollover warning and multi-objective stability control algorithm of a heavy-duty vehicle," *Ph.D. Thesis of Jilin University*, 2010.
2. J. Hilton and U. Shankar, "Motor vehicle traffic crash injury and fatality estimates," *NHTSA Technical Report*, 2003.
3. National Highway Traffic Safety Administration, "A compilation of motor vehicle crash data from the fatality analysis reporting system and the general estimates system," *Traffic safety facts 2016*, 2018.
4. H. Yu, L. Gäuvenc, and Ü. Özgüner, "Heavy-duty vehicle rollover detection and active roll control," *IFAC Proceedings Volume*, vol. 16, pp. 128–133, 2005.
5. Y. He, X. Yan, X. Y. Lu, D. Chu, and C. Wu, "Rollover risk assessment and automated control for heavy-duty vehicles based on vehicle-to-infrastructure information," *IET Intelligent Transport Systems*, vol. 13, no. 6, pp. 1001–1010, 2019.
6. X. Lu, Z. Zhou, and Y. Wang, "Design and Comparative Study of ABS Control Strategies Based on Co-Simulation," *Proceedings of IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER)*, no. 1, pp. 1665–1670, 2015.
7. V. Tsourapas, D. Piyabongkarn, A. C. Williams, and R. Rajamani, "New method of identifying real-time predictive lateral load transfer ratio for rollover prevention systems," *Proceedings of American Control Conference*, pp. 439–444, 2009.

Authors Introduction

Prof. Chian C. Ho

He received his B.S. degree in electrical engineering from National Chung Cheng University in Chiayi, Taiwan, in 1996. After four-year pursuit of M.S. and Ph.D. degrees, he received his Ph.D. degree in electrical engineering still from National Chung Cheng University in 2000. Just before his graduation in 2000, he went to Department of Electrical Engineering, University of California at Los Angles, as a visiting Ph.D. student for three months. From 2000 to 2005, he worked at ChungHwa Telecom Labs for VoIP equipment development in Taoyuan, Taiwan. Since 2005, he has joined the faculty of Department of Electrical Engineering, National Yunlin University of Science and Technology in Yunlin, Taiwan. He serves as an associate professor now.

Ms. Riki Umami Sanaz Ulfitria

She received her Bachelor's degree from State polytechnic of Bengkalis, Riau, Indonesian, in 2019. She is currently a graduate student in Department of Electrical Engineering, National Yunlin University of Science and Technology, in Yunlin, Taiwan.

Development of Intelligent Beehive and Network Monitoring System for Bee Ecology

Chau-Chung Song[a,b*], Geng-Yi Lin[a], Chi-Chung Peng[c] and Chung-Wen Hung[d]
[a]*Department of Aeronautical Engineering, National Formosa University, Yulin, Taiwan*
[b]*Smart Machine and Intelligent Manufacturing Research Center, National Formosa University, Yulin, Taiwan*
[c]*Department of Biotechnology, National Formosa University, Yulin, Taiwan*
[d]*Department of Electrical Engineering, National Yunlin University of Science & Technology*

[*]*E-mail: ccsong@nfu.edu.tw*
www.nfu.edu.tw

Abstract

In this paper, development of intelligent beehive and network monitoring system for bee ecology is focused on data acquis such as temperature, humidity, weight, and GPS positioning, combining the beehive with MCU, sensors, and ZigBee to implement front end sensing nodes to build a bee ecological network monitoring information system for real-time remote network monitoring, and assist beekeepers to establish a cloud-based real-time monitoring system and history traceability Bee product management system to enhance the convenience of beekeeper management and risk control, thereby effectively improving the efficiency of labor utilization. Also, clear production history information can increase consumers' trust in products and enhance the overall bee-related industry Economic benefits. This paper cooperated with beekeepers in Gukeng, Yunlin County, set up an Intelligent beehive system in the bee farm. Observations in the past month have shown that the activity and number of adult bees and larvae grown in the intelligent beehive are in good condition, user can also connect to the Intelligent beehive monitoring website through their mobile phone, tablet, or computer to analyze and monitor the status of each beehive.

Keywords: Intelligent beehive, Bee ecological information system, Real-time monitoring system.

1. Introduction

B Bees play an important role in agriculture, such as pollinating crops, producing a variety of bee products, and creating high agricultural economic output. The overall output value of bee products in Taiwan has reached NT$2.3 billion. The current intelligent beekeeping products use intelligent systems to grasp the status of the colony in real-time, such as monitoring the weight of the colony to know the amount of honey and determining when to harvest. Also reflect the health status of the bee colony by monitoring the data to prevent the problem from worsening or spreading. However, beekeepers in Taiwan still use traditional methods for nurturing and management, which are not only laborious and time-consuming but also disturb the bee colony. Furthermore, with increases in labor costs in rural areas and the general shortage of beekeeping manpower, the beekeeping industry is facing development difficulties and serious management challenges. [1-2]

2. Methods

2.1. *Environmental sensor module design*

The intelligent beehive detecting ecology of bee colony has 4 main functions.

2.1.1 *Temperature and humidity sensing*

Temperature and humidity are highly related to honey quality and bee colony conditions. [3]

2.1.2 *Weight sensing*

Monitoring the weight of honey can grasp the progress of honey collection by bees, and store and record daily beekeeping data. The maximum measuring weight of the weighing platform is 80kg, the resolution is 0.1kg, and the error range is ±5%.

2.1.3 *Beehive positioning*

With accurate positioning information, productivity records, and beehive environment data of deployment spots in seasons, beehive transfer plans can be made more cost-effectively.

2.1.4 *Wireless sensor network*

With wireless sensor networks and Internet of Things technology, beehive and bee colony status can be collected in real-time, also overcome the data collection difficulties of smart beehives deployed outdoors and achieve the function of collective remote beehive monitoring. More, a large number of beehives can be monitored at the same time.

3. System Design and Architecture

In the system, wireless sensor network of beehive environment sensor groups containing several sensor nodes is connected to an online database and system server by heterogeneous network gateway to provide remote monitoring of system status and create production records which help to build both the production decision system and production and sales resume system. Big data analysis system for detecting and diagnosing the ecology and health of bee colonies can be built with huge amounts of data. The system architecture is shown in Fig.1.

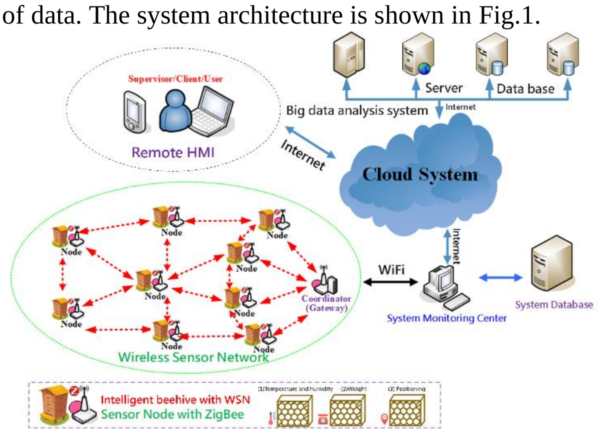


Fig. 1. The system architecture of Intelligent beehive and network monitoring system for bee ecology

3.1. *Wireless sensor network*

The wireless sensor network architecture is shown in Fig.2. When a beehive that has a Zigbee coordinator receives a packet from other node beehives, it repacks the data and uploads it to the cloud via Wi-Fi in JSON for further usages like production management or bee ecology and health of bee colony, etc. [4-5]

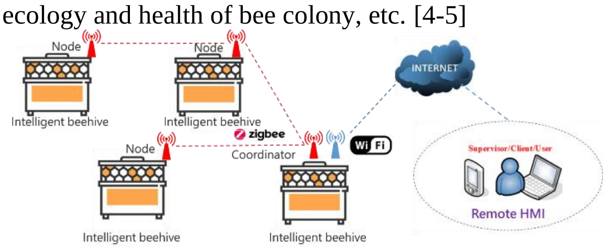


Fig. 2. The wireless sensor network architecture

3.2. *Cloud management system*

The webpage interface and the database server are based on WampServer which is an open-source integrated installation environment containing Apache Web Server, MySQL, and PHP. The Cloud network management system architecture is shown in Fig.3. [6]

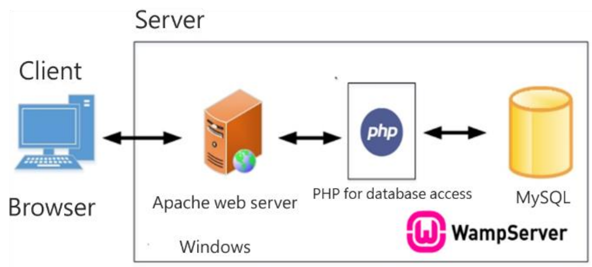


Fig. 3. The Cloud management system architecture

3.3. *Remote monitoring system*

The remote monitoring interface is a cross-platform browser-based application developed with HTML for the webpage, JavaScript for Interactive behavior, and CSS for layout. Data transfer via Asynchronous JavaScript and XML(AJAX) and HTTP POST between client and server. The server access database via PHP. The remote monitoring system architecture is shown in Fig.4.

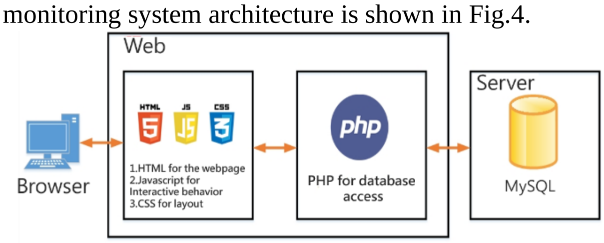


Fig. 4. The remote monitoring system architecture

4. Experiment results

4.1. *Environmental sensor module components*

Environmental sensor module components two main parts, "intelligent beehive main unit" and "weighing platform". In order to maximize the accessibility of the system to beekeepers, parts selections are optimized to meet the minimum system requirement and low cost. The components list is shown in Table 1.

Table 1. Components list [7]

Model	Parameter	Specification
Si7021	Humidity	0~100%, resolution 0.01%
	Temperature	-10℃~80℃, resolution 0.01
AT8502	Weight	The weighing platform can measure up to 80kg with a resolution of 0.01kg.
NEO-6M	Positioning	GPS latitude and longitude, the error distance is less than 100 meters.
-	Input Voltage	110VAC/60Hz
-	Spare Battery	Lithium Battery 3.6V (3250mAh), 3S

4.1.1 *Main Unit of Intelligent Beehive*

Main unit of Intelligent beehive as shown in fig.5 uses 110V/60Hz AC city power as the main power source. An internal backup lithium battery provides about 20 hours of backup time without the main power source. The positioning module and ZigBee module are installed in a dust/waterproof box to increase system durability in tough environments.

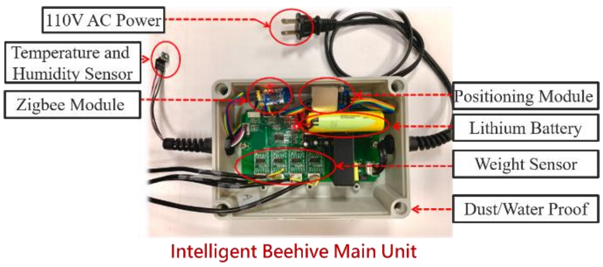


Fig. 5. Main Unit of Intelligent Beehive

4.1.2 *Weighing platform*

The weighing platform as shown in fig.6 is made of painted stainless steel to provide rust resistance for the outdoor environment.500*640mm platform size fits the general size of most beehives.

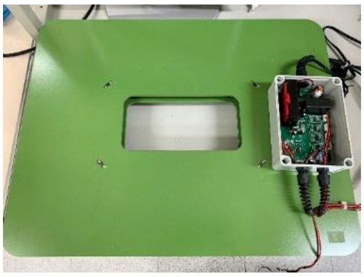


Fig. 6. Weighing platform

4.1.3 *Sensor module component configuration*

The positioning module and communication module are located in the dust/waterproof box of the beehive main unit. A beehive is placed on the top of the weighing platform. A humidity sensor is installed inside the beehive at 2/3 interior height to avoid stagnant water Influence. Component configuration as shown in Fig.7.

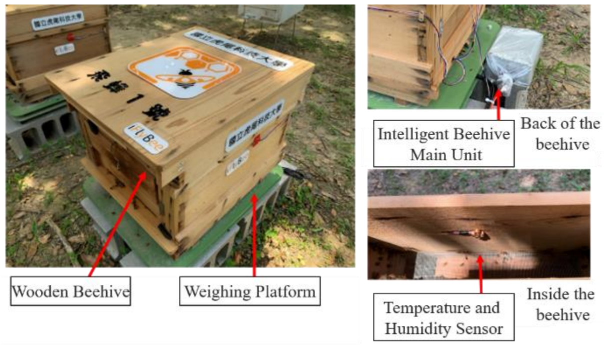


Fig. 7. Component configuration

4.2. *Monitoring and recording system*

The web page interface shown in Fig.8 provides visualized real-time temperature, humidity, weight information of each beehive set to users using a mobile phone, tablets, computers, or any web browsing device.

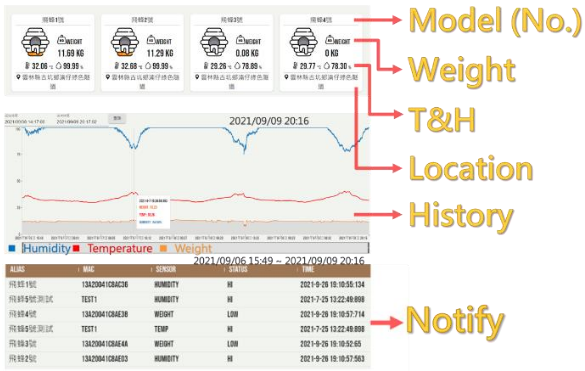


Fig. 8. Web page interface

5. Conclusion

In this paper, a series of sensor modules, wireless sensing network systems, cloud network management systems, remote monitoring systems, etc. are designed based on the bee ecological remote network monitoring and recording system, integrated and applied to bee ecological monitoring and recording. And cooperated with beekeepers in Gukeng, Yunlin County, set up an Intelligent beehive system in the bee farm, assist beekeepers to establish a bee product management system with real-time monitoring system, digitization, and traceability of histories to improve the convenience of beekeeper management and controlling the risk, thereby effectively improving the efficiency of labor utilization. Clear production history information can increase consumers' trust in products, improve the overall economic benefits of bee-related industries, and accelerate domestic research and development of intelligent beekeeping and related industries. The key application technology promotes the transformation and upgrading of the domestic beekeeping and bee products industry.

Acknowledgements

This research was supported by the Miaoli District Agricultural Research and Extension Station, Council of Agriculture, Executive Yuan, R.O.C., Research project: "Development of Intelligent beekeeping monitoring system," under Grant: 109I065-04.

References

1. Chen, Y. T.& Zhang, Y.X., *"Global bee industry issues and the current development of Taiwan's bee industry"*, Agricultural Biotechnology Industry Quarterly, Issue 52 (2017/12/31), P53-57.
2. Huang, W. Z., "High-tech Save the Bees", Commercial Times, 2016/10/30.
3. H. F. Abou-Shaara1, A. A. Owayss2,5, Y. Y. Ibrahim3, N. K. Basuny (2017). A review of impacts of temperature and relative humidity on various activities of honey bees. Insectes Soc., 64, 455-463.
4. ZigBee Alliance.
5. ZigBee Alliance, "ZigBee-Specification-2006-Download.pdf", http://www.ZigBee.org/.
6. WampServer official website.
7. Silicon Laboratories, "AN607: Si70xx Humidity Sensor Designer's Guide", 2016.

Authors Introduction

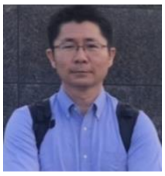
Chau-Chung Song

Chau-Chung Song is with the department of Aeronautical Engineering, National Formosa University, Taiwan. He received the M.S. degree in Control Engineering in 1994 and the Ph.D. degree in Electrical and Control Engineering from National Chiao Tung University in 2001 in Taiwan. He is currently a professor in the Institute of Aeronautical and Electronic Technology, and Department of Aeronautical Engineering, National Formosa University since 2006. His research interests include the analysis and control of nonlinear systems, power electronics applications, integration application of networking information systems, UAV technology and applications, cloud networks, and Cyber-Physical system (CPS).

Geng-Yi Lin

Geng-Yi Lin received his Bachelor's degree from the Department of Aeronautical Engineering, National Formosa University, Taiwan in 2021. He is currently a Master's course student in National Formosa University, Taiwan.

Chung-Wen Hung



Chung-Wen Hung received the Ph.D. degrees in Electrical Engineering from National Taiwan University in 2006. Currently he is an associate professor in National Yunlin University of Science & Technology. His research interests include IoT, IIoT, Power Electronics, Motor Control, and AI Application.

Data Balanced Algorithm Based on Generative Adversarial Network

I-Hsien Liu, Cheng-En Hsieh, Wei-Min Lin, 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: {ihliu, cehsieh, wmlin}@cans.ee.ncku.edu.tw, jsli@mail.ncku.edu.tw, chufenli@gmail.com*
www.ncku.edu.tw, www.nfu.edu.tw

Abstract

In order to defend against malicious attacks, intrusion detection systems have introduced machine learning as a protection strategy. However, machine learning algorithms and datasets have a great influence on the effectiveness of the machine learning model. This study uses five algorithms which are Naïve Bayes, CNN, LSTM, BAT, and SVM to train the IDS machine learning model. We design a data-balanced method based on the GAN algorithm to improve the data imbalance problem of the IDS dataset.

Keywords: Anomaly Traffic Detection, Machine Learning, IDS Dataset, GAN, Performance Analytics

1. Introduction

As more data is transmitted on the Internet, the greater the risk of malicious network attacks, and even national security issues may arise. With the increasing popularity of artificial intelligence and the maturity of big data collection technology, machine learning has also been applied to intrusion detection systems. How to convert the huge amount of network traffic data into clean and highly recognizable machine learning datasets is a very important issue. There are many types of public IDS datasets. Some datasets are not compatible with current network attacks because of outdated attack methods and insufficient diversity. The content is not suitable for the research of modern artificial intelligence intrusion detection systems. Our research selects three representative IDS datasets as research objects based on attack diversity and data integrity, namely NSL-KDD [1], UNSW-NB15 [2] and CICIDS 2017 [3]. At the same time, we found through sensitivity analysis that traffic datasets with too large a gap between malicious traffic data and normal traffic data performed poorly. In order to overcome this problem, this research designed a data-balanced algorithm based on a generative adversarial network for the data distribution of the dataset, called GAN-BAL. The GAN-BAL algorithm is used to generate traffic datasets for training and evaluation. This dataset satisfies the diversity of attacks and data integrity and retains the heterogeneity of the data, which can achieve better results in model training.

2. Related Work

At present, most intrusion detection system researchers use artificial intelligence to defend against constantly changing and evolving malicious attack methods. The below explains common machine learning algorithms used in malicious traffic detection.

2.1. *Naïve Bayes classifier (NB)*

Naïve Bayes classifier (Naïve Bayes, NB) is a simple method for constructing classifiers. The theoretical basis is to assume that each feature is conditionally independent. According to Bayes' theorem, we can classify conditionally independent features easily. The Naïve Bayes classifier will calculate each data the conditional probability of each category, and then use the maximum a posteriori (MAP) estimation to determine the best classification method.

2.2. *Support vector machine (SVM)*

The support vector machine originated from the algorithm proposed by Vapnic et al. based on statistical mathematics in 1963 and it is designed to solve the problems related to regression analysis and statistical classification [4]. SVM is better than many traditional machine learning methods in classifying nonlinear and high-dimensional data. The core concept of the support vector machine is to classify data by finding a hyperplane. SVM searches for the closest data point to the hyperplane. The distance between the hyperplane and the point is called the support vector.

2.3. *Convolution neural network (CNN)*

Convolutional neural network is a combination of the convolutional layer and deep neural network. Convolutional neural network uses the convolutional layer and the pooling layer to achieve a method that can reduce data without losing too much information. Convolutional neural networks have achieved good results in image processing and speech analysis.

2.4. *Long short-term memory (LSTM)*

The short-term memory model is an improved model of the recurrent neural network. Hochreiter et al. published the paper for the first time in 1999 [5], which mainly improved the defects of the recurrent neural network for time series. LSTM is composed of four units: Input Gate), Output Gate, Forget Gate, and Memory Cell. It is mainly used to alleviate the problem of the disappearance of the gradient. The memory unit will record the data of the last state, and use 4 valves to determine whether the input or output data needs to be stored or output.

2.5. *BAT*

The BAT algorithm is a deep learning method proposed by Su et al. for network intrusion detection systems [6]. This algorithm is designed based on a two-way long and short-term memory model and attention mechanism and is used in the data processing. The convolutional layer is used for processing, and the two-way long and short-term memory model is used to learn the characteristics of each flow and obtain the vector corresponding to each flow. Then use the attention mechanism to perform feature learning on the sequence data composed of traffic vectors to obtain subtle features to achieve a better classification effect.

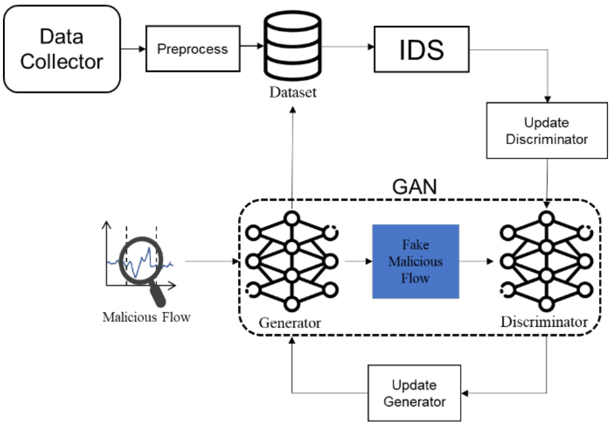


Fig. 1. System architecture.

3. System Structure

The structure of this research is introduced in this chapter. There are two stages: data collector and generative adversarial network data balance algorithm. The system architecture of the experiment is shown in Fig. 1.

3.1. *Data collector*

We collect the traffic in the real network set in our Lab and then use CICFlowmeter to transform the traffic into the statistics files. The CICFlowmeter is an Ethernet traffic Bi-flow generator and analyzer for anomaly detection, it can transform the pcap to the csv file. We call the traffic collected from the real network the true flow dataset.

We use the Cuckoo sandbox system to trigger malicious programs to create malicious traffic datasets. The malicious program samples are provided by the National High-Speed Network and Computing Center. The Cuckoo triggers malicious program samples through the network on the client-side, and Agent.py on the client-side will record various behaviors of the malicious program samples and send them back to the host on the user side. The main purpose of the above behaviors is to record the behavior of the malicious program samples without contacting malicious programs. For program samples, we analyzed the traffic generated by malicious samples and find the background traffic in the entire recording traffic is very small. The reason is that the Cuckoo system starts recording after the samples are triggered. The background traffic is so small that it can be ignored. Therefore, we call the traffic recorded by the Cuckoo system the malicious flow.

3.2. *GAN-BAL algorithm*

The core of the generative confrontation network is mainly divided into two parts: the generative model and the discriminative model. The generative model will generate fake traffic to the discriminant model during the training process. After the discriminating model discriminates the fake traffic true or false, the identification result will be fed back to the former, so that the generative model can improve the strategy of making malicious traffic. Then, the generative model will give the traffic generated after the modified strategy to the intrusion detection system which judges the authenticity of the forged traffic, and the discriminant model is used to improve the strategy of judging the traffic. After many times of updating, the generative model can generate fake traffic that is almost the same as the real traffic. The discriminant model will train a neural network that can identify the authenticity of the traffic, which is used to test the traffic generated by each generation model. On the other hand, the generation model will repeatedly improve its ability to "fake" from the feedback of each discriminant model. To fight against it, this is the concept of generating a confrontational network. We refer to the balanced data algorithm for generating adversarial networks designed by Huang et al. [7]. This algorithm uses Gaussian Noise as the training data for the generative model. In order to reduce training time and forge malicious traffic to be closer to the real malicious traffic. We change the input data of the generative model. We use the malicious traffic that triggers the malicious program sample to generate the malicious traffic training data. The flow chart of the algorithm is as follows Fig. 2. We call the improved algorithm the GAN-BAL algorithm.

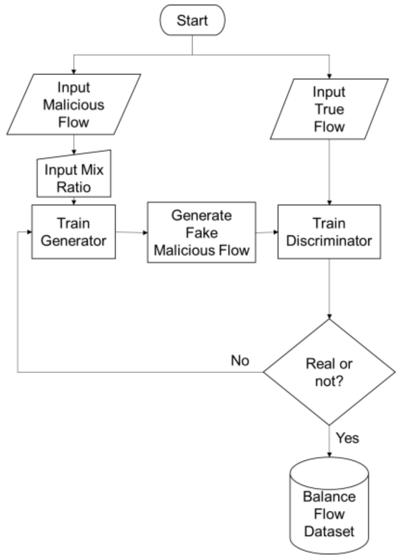


Fig. 2. The flow chart of the GAN-BAL algorithm.

4. Experimental Result

In order to verify whether the data balancing algorithm of this study improves the model results, we use supervised machine learning algorithms to evaluate its effects. We apply the GAN-BAL algorithm to the CICIDS 2017 dataset. We use five algorithms such as CNN, LSTM, BAT, SVM, and NB to perform model training on the original CICIDS 2017 dataset and the CICIDS 2017 dataset processed by the GAN-BAL algorithm to verify whether the algorithm improves the model results. The comparison results are shown in Table 1. As a result, the recall rate of CNN has increased by 20%, and the accuracy rate has increased by 4%; the recall rate of LSTM Increased by 10%, accuracy rate increased by 2%; BAT recall rate increased by 16%, accuracy rate increased by 3%; SVM recall rate increased by 15%, accuracy rate increased by 4%. Based on the above experimental results, although the use of the GAN-BAL algorithm proposed in this research will cause a loss of precision, the recall rate and accuracy rate are improved. The experimental results verify the

Table 1. Comparison of CNN, LSTM, BAT, SVM, and NB model training results

		Recall	Precision	F1-score	Accuracy
CNN	With GAN-BAL	0.991210	0.937769	0.963749	0.989934
	Without GAN-BAL	0.769745	0.998967	0.869503	0.957217
LSTM	With GAN-BAL	0.900657	0.947915	0.875400	0.918632
	Without GAN-BAL	0.800050	0.996801	0.887654	0.932155
BAT	With GAN-BAL	0.994952	0.953778	0.973930	0.992715
	Without GAN-BAL	0.824089	0.998934	0.903127	0.968218
SVM	With GAN-BAL	0.999636	0.992143	0.995876	0.998828
	Without GAN-BAL	0.816055	0.998819	0.898234	0.967708
NB	With GAN-BAL	0.165259	0.366583	0.227816	0.380103
	Without GAN-BAL	0.157882	0.368219	0.221004	0.352480

improvement of malicious intrusion detection using GAN-BAL algorithm.

5. Conclusion

This study analyzes the data sensitivity of the algorithm by evaluating the performance of each algorithm in the evaluation index of different datasets and proposes a data set balance algorithm for the unbalanced defects of the dataset. Compared with the research of intrusion detection system based on generative confrontation network proposed by Shahriar et al. [8], this research is developed for modern malicious attack behavior, and Huang et al. developed IGAN-IDS. The system uses noise as the input of the generative model [7]. This research uses the malicious traffic actually induced by the malicious sample as the input to reduce the training time and the basis for generating the malicious traffic.

Acknowledgements

This work was supported by the Ministry of Science and Technology (MOST) in Taiwan under contract numbers MOST 108-2221-E-006-110-MY3 and MOST 110-2218-E-006-013-MBK.

References

1. 1. The UCI KDD Archive, "KDD Cup 1999 Data," 28 10 1999. [Online]. Available: https://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html. [Accessed 1 6 2021].
2. 2. N. Moustafa and J. Slay, "UNSW-NB15: A Comprehensive Data set for Network Intrusion Detection systems," 2015 Military Communications and Information Systems Conference (MilCIS), Canberra, ACT, Australia, 10-12 Nov., 2015.
3. 3. University of New Brunswick, "CICIDS 2017 dataset," University of New Brunswick, 1 7 2017. [Online]. Available: https://www.unb.ca/cic/datasets/ids-2017.html. [Accessed 15 6 2021].
4. 4. V. N. Vapnik, "An overview of statistical learning theory," IEEE Transactions on Neural Networks, vol. 10, no. 5, pp. 988 - 999, 1999.
5. 5. S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation (1997), vol. 9, no. 8, pp. 1735-1780, 1997.
6. T. Su, H. Sun, J. Zhu, S. Wang and Y. Li, "BAT: Deep Learning Methods on Network Intrusion Detection Using NSL-KDD Dataset," IEEE Access, vol. 8, pp. 29575 - 29585, 2020.
7. S. Huang and K. Lei, "IGAN-IDS: An imbalanced generative adversarial network towards intrusion detection system in ad-hoc networks," Ad Hoc Networks, vol. 105, p. 102177, 2020.
8. M. H. Shahriar, N. I. Haque, M. A. Rahman, and M. Alonso, "G-IDS: Generative Adversarial Networks Assisted Intrusion Detection System," 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC), Madrid, Spain, 13-17 July, 2020.

Authors Introduction

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. Cheng-En Hsieh

He received his B.S. degree from the Department of Communication Engineering, National Central University, Taiwan in 2019. He got the M.S. degree in National Cheng Kung University in Taiwan. His research focuses on network communication and cyber security.

Ms. Wei-Min Lin

She received her B.S. degree from the Department of Electrical Engineering, Yuan Ze University, Taiwan in 2020. She is acquiring the master's degree in Department of Electrical Engineering / Institute of Computer and Communication Engineering, National Cheng Kung University in Taiwan.

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.

s

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.

Fault-Tolerant Control System Design for Nonlinear System with Actuator Faults

Ho-Nien Shou*
Dep. of Aviation & Communication Electronics,
Air Force Institute of Technology, No.198, Jieshou W. Rd., Gangshan Dist., Kaohsiung City 820009, Taiwan (R.O.C.).

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 (R.O.C.)
E-mail: honien.shou@gmail.com, hylai@cans.ee.ncku.edu.tw,
www.afats.khc.edu.tw, www.ncku.edu.tw

Abstract

This article deals with observer-based integrated robust fault estimation and accommodation design problems for nonlinear system. The environmental disturbance torque, actuator faults, sensor faults and model uncertainties are considered. Firstly, we propose the augmented fault estimation observer (AFEO) to guarantee the convergence of $\boldsymbol{H}_\infty$ performance index of fault estimation and to restrict the influence of uncertainties with respect to the fault estimation error as well. We then design the fault accommodation which is based on the dynamic output feedback to keep the stability of the closed-loop system while malfunctioning. AFEO and dynamic output feedback fault tolerant controller (DOFFTC) are separately designed. Their performances are separately considered. Finally, we propose a simulation result of the micro-satellite attitude control system to demonstrate the effectiveness of the presenting method.

Keywords: augmented fault estimation observer, $\boldsymbol{H}_\infty$ performance index, dynamic output feedback fault tolerant controller.

1. Introduction

The main missions of the satellite include communication, navigation, earth observation, and so on. The attitude control system is one of the most important subsystems of the satellite which requires high performance and excellent functionality. Different kinds of unavoidable malfunctions will occur in a harsh mission environment such that some performances will become poorer and even shut down the whole system. The fault diagnosis theory might be one of the most efficient methodologies among all attitude control design theories. It can provide reliable attitude control when component faults or malfunctions occur. Therefore, many contributions are made to the satellite attitude control problem on the basis of fault diagnosis theory [1-5].

There are three types of fault diagnosis (FD) method: signal-based, knowledge-based, and model-based. Three essential tasks are included in the basic FD strategy: fault detection, fault isolation, and fault estimation. Due to the increased requirements of performance, safety, and reliability, the model-based fault detection and isolation are getting concerned in the last two decades [1-8].

Fault estimation is the most common method among these model-based methods [1]. It uses filters to generate residuals and then to set a threshold according to these residuals to detect the faults. Recently, with the developments of robust control theory and $H\infty$ optimal control theory, more methods have been proposed to solve the robust fault detection and isolation (RFDI) problems. The goal of the robust fault detection is to distinguish faults effects and effects of uncertain signals and perturbations, which is different from the robust control. Therefore, the performance of the RFDI system is measured by the appropriate trade-off between robustness and sensitivity.

This article is divided into two parts. The microsatellite H∞ robust fault detection filter (RFDF) design problem is introduced in the first part. The second part discusses the microsatellite fault-tolerant control problem. In the first part, the microsatellite is considered as a nonlinear system with external disturbance input. The aim of this part is to design a stable RFDF system with $\boldsymbol{H}_\infty$ performance for the microsatellite. The system is described by a state space model formed by the linear terms combining with the nonlinear terms which satisfy the Lipschitz condition. The residual model is used to simplify the RFDF design problem to the standard H∞ model matching problem in this article. The performance index used to design the reference residual model considers the robustness to the disturbances and the sensitivity to the faults. In [3], the RFDF system was designed with a similar performance index without considering the failure of the sensor, hence, it cannot deal with the condition when both actuators and sensors malfunctions. We apply a linear matrix inequalities (LMIs) method in $\boldsymbol{H}_\infty$ optimization technique to the RFDF design problem in this article.

In the second part, fault detection and isolation are considered to be the most important tasks and also are the main concern in references. The accurate and immediate fault estimation helps to reconstruct fault parameters (and/or) signals, hence, it plays a critical role in the fault-tolerant control system. Even though the fault effects can be accommodated in the corresponding control reconfiguration in the fault-tolerant control system, the existences of disturbances, noises, and model uncertainties make the fault estimation design becomes more difficult.

There are three main contributions of this paper. Firstly, we propose a multi-objective augmented fault estimation observer (AFEO). This observer contains exponential stability and $\boldsymbol{H}_\infty$ performance, which not only guarantees the convergence speed of fault estimation but also limits the influence of uncertainty as much as possible. Secondly, it is well known that traditional multi-objective design methods generate stability for the system when faults occur, and the detailed design steps of the strategy are given based on LMIs. The third is to consider uncorrelated actuator failure and sensor failure issues at the same time. It is worth noting that AFEO and DOFTC are independently designed from the perspective of the entire design process. To avoid the coupling problem caused by observer state feedback and to help the estimation of design parameters, the performances of AFEO and DOFTC are simultaneously considered. As far as we know, no reference about fault estimation and accommodation control problem has been reported.

This article is organized as follows. The description of the system refers to Section 2. A new type of augmented fault estimation observer (AFEO) is designed for a class of uncertain Lipschitz nonlinear systems in Section 3, which can provide an H_∞ performance index to estimate the fault vector. We propose a new fault accommodation design method based on dynamic output feedback and present a result of an actual system in Section 4. At the end, the conclusion of this article.

2. Micro-satellite Attitude Control System Model

By using the well-known Euler's moment equation, we can describe the dynamic characteristics of a rigid microsatellite in a circular orbit in terms of a nonlinear equation of motion [1]:

$$J\dot{\boldsymbol{\omega}} = -\boldsymbol{\omega}^{\times} J\boldsymbol{\omega} + \boldsymbol{\tau}_w + \boldsymbol{\tau}_u \tag{1}$$

where $\boldsymbol{J} = \operatorname{diag}\left\{J_x, J_y, J_z\right\}$ are inertia moments of the satellite along principal axes; $\boldsymbol{\omega}$ are the angular velocity of the body-fixed reference frame. $\boldsymbol{\tau}_w$ are space disturbance torques and $\boldsymbol{\tau}_u$ are the control torques along principal axes. Denotes a skew-symmetric matrix which is given by

$$\boldsymbol{\omega}^{\times} = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix} \tag{2}$$

Satellite attitude kinematics can be obtained as follows [2]:

$$\begin{bmatrix} \dot{\phi} \\ \dot{\theta} \\ \dot{\psi} \end{bmatrix} = \frac{1}{\cos\phi} \begin{bmatrix} \left(\omega_{box}\cos\theta + \omega_{boy}\sin\theta\right)\phi \\ \omega_{box}\cos\phi + \left(\omega_{box}\sin\theta - \omega_{boz}\sin\theta\right) \\ \omega_{boz}\cos\theta - \omega_{box}\sin\theta \end{bmatrix} \tag{3}$$

where ϕ, θ and ψ were roll angle, pitch angle, yaw angle along principal axes, respectively. $\omega_{bo} = \begin{bmatrix}\omega_{box} & \omega_{boy} & \omega_{boz}\end{bmatrix}^T$ the angular velocity of a body-fixed reference frame of a satellite with respect to an orbit reference frame expressed in the body-fixed reference frame. $\boldsymbol{\omega} = \omega_{bo} + T_{bo}\omega_{oi}$, where T_{bo} is the attitude transformation matrix. ω_0 is the constant orbital rate, and $\omega_{oi} = [0 \quad -\omega_0 \quad 0]$ is the orbit angular velocity.

Consider the satellite is moving in a small angle maneuver. According to Eq. (1), Eq. (2) and Eq. (3), satellite attitude control system (ACS) model with actuator and sensor faults are established into the following equations:

$$\begin{cases} \dot{x}(t) = Ax(t) + B_w w(t) + B_u u(t) + g(x,t) \\ \qquad + \psi(x,u,t) + F_a f_a(t) \\ y(t) = Cx(t) + F_s f_s(t) \end{cases} \tag{4}$$

$$\dot{z}(t) = -A_z z(t) + A_z y(t) \tag{5}$$

where $x = [\phi \quad \theta \quad \psi \quad \omega_x \quad \omega_y \quad \omega_z]^T$ is state variable vector, $w = [w_x \quad w_y \quad w_z]^T$ denotes the disturbance torque which is consist of gravitational perturbation solar radiation pressure, electromagnetic force, $u = [u_x \quad u_y \quad u_z]^T$ represents the input vector, $f_a(t)$ denotes actuator fault, and $f_s(t)$ is the a sensor fault. The system matrix can be expressed as

$$A = \begin{bmatrix} 0 & 0 & \omega_0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ -\omega_0 & 0 & 0 & 0 & 0 & 1 \\ 3\omega_0^2(J_z - J_y)J_x^{-1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 3\omega_0^2(J_z - J_x)J_y^{-1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$, \quad g(t,x) = \begin{bmatrix} 0 \\ \omega_0 \\ 0 \\ \omega_y\omega_z(J_z - J_y)J_x^{-1} \\ \omega_x\omega_z(J_x - J_z)J_y^{-1} \\ \omega_x\omega_y(J_y - J_x)J_z^{-1} \end{bmatrix} , \quad C = I_6 \quad ,$$

$$B_u = B_w = F_a = \begin{bmatrix} 0_{3\times3} \\ \boldsymbol{J}^{-1} \end{bmatrix}, F_s = I_6 .$$

Assume the microsatellite is moving in a small angle maneuver, hence, the nonlinear function $g(x,t)$ is locally Lipschitz nonlinear with a Lipschitz constant l_c, which means,

$$\left\| g(t,x_1) - g(t,x_2) \right\| \le l_c \left\| x_1 - x_2 \right\| \tag{6}$$

We can have the augmented system by inserting Eq. (5) into

$$\begin{cases} \dot{x}(t) = Ax(t) + B_w w(t) + B_u u(t) + g(x,t) \\ \qquad + \psi(x,u,t) + F_a f_a(t) \\ \dot{z}(t) = A_z Cx(t) - A_z z(t) + A_z F_s f_s(t) \end{cases} \tag{7}$$

$$\begin{bmatrix} \dot{x}(t) \\ \dot{z}(t) \end{bmatrix} = \begin{bmatrix} A & 0 \\ A_z C & -A_z \end{bmatrix} \begin{bmatrix} x(t) \\ z(t) \end{bmatrix} + \begin{bmatrix} B_w \\ 0 \end{bmatrix} w(t) + \begin{bmatrix} B_u \\ 0 \end{bmatrix} u(t) + \begin{bmatrix} F_a & 0 \\ 0 & A_z F_s \end{bmatrix} \begin{bmatrix} f_a(t) \\ f_s(t) \end{bmatrix} + \begin{bmatrix} g(x,t) + \psi(x,u,t) \\ 0 \end{bmatrix} \tag{8a}$$

$$y(t) = [C \quad 0] \begin{bmatrix} x(t) \\ z(t) \end{bmatrix} + [0 \quad F_s] \begin{bmatrix} f_a(t) \\ f_s(t) \end{bmatrix} \tag{8b}$$

To the convenience purpose, we re-express Eq. (8) as

$$\dot{\bar{x}} = \bar{A}\bar{x} + \bar{B}_w w + \bar{B}_u u + \bar{F}\bar{f} + \bar{g} \tag{9a}$$

$$y = \bar{C}\bar{x}(t) + \bar{E}\bar{f} \tag{9b}$$

where $\bar{x} = [x^T \quad z^T]^T$ represents the new state, $\bar{f} = [f_a^T \quad f_s^T]^T$, $\bar{A} = \begin{bmatrix} A & 0 \\ CA_z & -A_z \end{bmatrix}$, $\bar{B}_w = \begin{bmatrix} B_w \\ 0 \end{bmatrix}$, $\bar{B}_u = \begin{bmatrix} B_u \\ 0 \end{bmatrix}$, $\bar{F} = \begin{bmatrix} F_a & 0 \\ 0 & A_z F_s \end{bmatrix}$, $\bar{g}(x,u,t) = \begin{bmatrix} g(x,t) + \psi(x,u,t) \\ 0 \end{bmatrix}$

$[0 \quad F_s] = [0 \quad A_z^{-1}] \begin{bmatrix} F_a & 0 \\ 0 & A_z F_s \end{bmatrix} = \bar{E}$.

3. Nonlinear Fault Estimation Observer

The nonlinear fault estimation observer is described as

$$\begin{aligned} \dot{\hat{\bar{x}}}(t) &= \bar{A}\hat{\bar{x}}(t) + \bar{g}(t, \hat{\bar{x}}(t)) + \bar{B}_u u(t) \\ &\quad + \bar{F}\hat{\bar{f}}(t) - L(\hat{y}(t) - y(t)) \end{aligned} \tag{10}$$

$$\hat{y}(t) = \bar{C}\hat{\bar{x}}(t) + \bar{E}\hat{\bar{f}}(t) \tag{11}$$

$$\dot{\hat{\bar{f}}}(t) = \hat{\bar{f}}(t) - F(\hat{y}(t) - y(t)) \tag{12}$$

where $\hat{\bar{x}}(t) \in \mathbb{R}^{n+p}$ is the state vector, $\hat{y}(k) \in \mathbb{R}^{p}$ is the output vector, $\hat{\bar{f}}(t) \in \mathbb{R}^{q+r}$ is the estimation vector of $\bar{f}(t)$, $L \in \mathbb{R}^{(n+p)\times p}$ and $F \in \mathbb{R}^{(q+r)\times p}$ are matrix parameters waited to be determined.

Therefore, from the dynamical error function (12) we have

$$\dot{e}_x(t) = (\bar{A} - L\bar{C})e_x(t) + (\bar{F} - L\bar{E})e_f + G(t,x,\hat{x}) - \bar{B}_w w - (\bar{A}\bar{x} + \bar{B}_w w + \bar{B}_u u + \bar{F}\bar{f} + \bar{g}(t,x(t))) \tag{13}$$

$$G(t,x,\hat{x}) = \bar{g}(t,x(t)) - \bar{g}(t,\hat{x}(t))$$

$$e_y(t) = \hat{y}(t) - y(t) = \bar{C}e_x(t) + \bar{E}e_f(t) \tag{14}$$

4. H_∞ Robust Fault Estimation Observer Design

Combining Eq. (14) and Eq. (15), we can obtain the augmented error system,

$$\begin{bmatrix} \dot{e}_x(t) \\ \dot{e}_f(t) \end{bmatrix} = \left(\begin{bmatrix} \bar{A} & \bar{F} \\ 0 & I \end{bmatrix} - \begin{bmatrix} L \\ H \end{bmatrix} \begin{bmatrix} \bar{C} & \bar{E} \end{bmatrix} \right) \begin{bmatrix} e_x(t) \\ e_f(t) \end{bmatrix} + \begin{bmatrix} \bar{B}_w & 0_{\times n} \\ 0 & -I_{q+r} \end{bmatrix} \begin{bmatrix} w(t) \\ \Delta f(t) \end{bmatrix} + \begin{bmatrix} I_n \\ 0_{(q+r)\times n} \end{bmatrix} G(t,x(t),\hat{x}(t)) \tag{16}$$

and

$$e_y(t) = \begin{bmatrix} \bar{C} & \bar{E} \end{bmatrix} \begin{bmatrix} e_x(t) \\ e_f(t) \end{bmatrix}.$$

5. Conclusion

This paper proposes a structure combined with robust fault estimation and fault-tolerant accommodation for the microsatellite attitude control system. It is designed on the basis of the observer and the Lipschitz condition which constrains the nonlinear terms of the system. We study a new multi-objective fault estimation method to improve the fault estimation performance. This new method has the property of low conservative compared to other traditional methods. We then propose an integrated system which is composed of fault estimation observer and fault-tolerant controller. The uncoupling characteristic of this integrated system allows us to design the two parts separately.

References

1. C. Gao, Q. Zhao, G. Duan, Robust actuator fault diagnosis scheme for satellite attitude control systems, Journal of the Franklin Institute350(9) (2013)2560-2580.
2. S. M. Azizi, K. Khorasani, A hierarchical architecture for cooperative actuator fault estimation and accommodation of formation flying satellites in deep space, IEEE Transactions on Aerospace and Electronic Systems, 48(2) (2012) 1428–1450.
3. Y. Shen, Y. Zhang, Z. Wang, Satellite fault diagnosis method based on predictive filter and empirical mode decomposition, Journal of Systems Engineering and Electronics22(1) (2011)83–87.
4. T. Jiang, K. Khorasani, S. Tafazoli, Parameter estimation-based fault detection, isolation and recovery for nonlinear satellite models, IEEE Transactions on Control Systems Technology16(4) (2008)799–808.
5. D. Henry, Fault diagnosis of microscope satellite thrusters using H∞/H_ filters, Journal of Guidance, Control, and Dynamics 31(3) (2008)699-711.
6. W. Chen, M. Saif, Observer-based fault diagnosis of satellite systems subject to time-varying thruster faults, Journal of Dynamic Systems, Measurement, and Control129(2007)352–356.
7. J. Chen, P.R. Patton, Robust Model-based Fault Diagnosis for Dynamic Systems, Kluwer Academic Publishers, Boston, 1999.
8. P. M. Frank, X. Ding, Survey of robust residual generation and evaluation methods in observer-based fault detection systems, J. Proc. Control 7 (1997) 403–424.

Authors Introduction

Mr. Hsin-Yu Lai

He was born in Chiayi, Taiwan in 1996. He received his B.S. degree from the Department of Electrical Engineering, National Chung Cheng University, Taiwan in 2019. He is acquiring the master's degree in Department of Electrical Engineering / Institute of Computer and Communication Engineering, National Cheng Kung University in Taiwan.

Dr. Ho-Nien Shou

He received the Ph.D. degree in Electrical Engineering from National Cheng Kung University, 2002. From 1999 to 2001, he was with the National Space Organization (NSPO) as an assistant researcher, working with satellite attitude control system. Currently, he works as an associate professor at Department of Aviation & Communication Electronics Air Force Institute of Technology.

Key Indicators for Successful E-oriented Operation and Management of the Nutrition Consulting Service System

Ling-Mei Hsu
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: lianyingtw@yahoo.com.tw

Abstract

According to the application theory of information systems, it is a tendency that the e-management of the nutrition consulting service industry by the Internet generation, and service quality is the key indicator of operational performance, which is of great importance. This research purpose was to analyze the key dimensions of online service quality when customers were learning online. This research used the E-SERVQUAL to design five major dimensions and 15 criteria, and through the analytic hierarchy process (AHP) constructed a set of successful operation management modules. This research verified the order of the key indicators were: Security/Privacy, Reliability, Reactivity, Efficiency, Tangibility.

Keywords: E-SERVQUAL, Nutrition Consulting, Analytic hierarchy process (AHP)

1. Introduction

With economic prosperity, human life styles around the world are facing unprecedented changes. The prevalence of obesity in countries around the world is becoming more and more serious. Data from the World Health Organization show that more than 1.9 billion adults worldwide are overweight, and about 13% of the population is obese [1]. The World Obesity Federation predicts that the world's overweight population will rise from 2 billion in 2014 to 2.7 billion in 2025 [2]. At present, obesity seriously affects people's health in Taiwan. According to the National Nutrition and Health Status Change Survey, the weight of nearly 40% of adults in Taiwan is in the high-risk range [3]. Therefore, obesity prevention and control must be vigorously promoted and everyone should face the health problems caused by obesity and actively prevent and treat the chronic complications caused by obesity, so as to consolidate public health. The nutrition consultation agency is the most ideal professional medical unit to help control weight, keep health, and improve chronic diseases [4]. It makes a tangible contribution to safeguarding public.

Enterprise operation requires considerable professional knowledge. The nutritional consulting and medical industry is classified as the preventive medicine health industry in accordance with the regulations of the Ministry of Health and Welfare of the Executive Yuan of Taiwan [4]. However, nutrition consulting and medical work mainly requires service quality in addition to nutrition medicine specialty, and thus, it is extremely necessary to have the ability of business management and the quality of customer service to operate it continuously. In the 21st century, according to the application theory of information systems, it is indispensable to use advanced information technology to manage their business activities and promote the growth of corporate organizations due to the fierce competition and service quality requirements in the business environment [5,6]. In response to the situation, the improvement of e-service quality of nutrition consulting agencies is an important element of business success in the nutrition consulting industry in the modern society. According to this, the key aspects and criteria of e-service quality are to conduct sound management and meet customer service quality for nutrition consulting agency operators and related demands, competitiveness enhancement and profitability creation are indispensable and important indicators [6,7].

The main purpose of this research is to: (1) conduct questionnaires and expert interviews through relevant service quality literature on the website, and then develop a measurement framework suitable for the nutrition consulting and medical industry to evaluate the quality of online services; (2) explore the key aspects and criteria for the quality of online services, and discuss its service performance to provide a reference for sustainable operation in the nutrition consulting industry.

2. Literature Review

2.1. *Nutrition consultation*

According to the definition of nutrition counseling proposed by Katharine R. Curry and Amy Jaffe [8]: "Nutrition counseling is to give dietary guidance to healthy people or diseased patients, plan to make the diet meet the nutritional needs on each stage of growth, or modify the treatment to adapt to the disease. Solving the difficulties encountered in diet changes is to achieve the principle of a balanced diet and obtain a proper establishment of a healthy lifestyle".

2.2. *E-Service Quality*

Zeithaml et al. [9] believe that the quality of website services can be regarded as the degree to which the website promotes efficient and effective shopping and product or service delivery, this study defines website service quality as "the website service provider delivers services to customers through the special environment of "website," and the customer's overall evaluation of the service."

In the past researches on measuring service quality, most of them were based on the service quality measurement model SERVQUAL proposed by Parasuraman et al. [10]. Zeithaml et al. [9] propose a conceptual framework for online service quality and 11 aspects to measure website service quality. Later, in 2002, the online service quality dimension and conceptual framework proposed in 2000 were revised in 2002, and seven dimensions were proposed, which would indeed affect online service quality. Among them, efficiency, reliability, fulfillment, and privacy form the E-SERVQUAL scale, which mainly measures customer satisfaction with the online functions provided by the website. In addition, the recovery E-SERVQUAL scale consists of responsiveness, compensation and contact, which is mainly used for the services that customers hope to get when they have questions or encounter trouble [11]. Based on the views of VIP customers of nutrition consulting organizations, the study cited the five aspects of the online service quality E-SERVQUAL scale proposed by scholar Zeithaml et al. [11], 15 suitable evaluation criteria are selected, as shown in Table 1.

Table 1. Key Indicators of E-SQ Success of Nutrition Organizations

Dimensions	No.	Indicators
Tangibility	A1	The teaching material is clear.
	A2	The course content is updated instantly.
	A3	The information is complete.
Reliability	B1	Professional consultation process.
	B2	The nutritionist has sufficient professional training.
	B3	Effective online consultation.
Security/Privacy	C1	Safe nutritional guidance.
	C2	Helpful performance improvement.
	C3	Safe transaction system provision.
Reactivity	D1	Kindly online service attitude.
	D2	Quick customer question reply.
	D3	Immediate services.
Efficiency	E1	Quick benefits on online consultation.
	E2	Simple online consultation.
	E3	Online consultation saves traffic time.

3. Research Methods

In this study, AHP is adopted as a research method to explore the key factors for the success of E-service quality management in a nutrition consulting organization. Through a systematic level, the complex website service quality evaluation system is transformed into a clear framework, and the scale of one to nine is compared. For the weights between the evaluation criteria, we establish a paired comparison matrix, calculate its eigenvalues and eigenvectors, and finally use the largest feature to perform a consistency test to obtain the relative weights between the evaluation criteria. In this study, such method, one of the most effective methods for cohesive judgment, is applied to obtain the E-service quality of the nutrition consulting agency, the relative weight of each aspect and each factor, and pairwise comparison [12]. Through the hierarchical process analysis method, an expert questionnaire is made to interview VIP customers of nutrition consulting agencies, and the order of weight for the quality of online services is ranked.

4. Empirical Analysis

4.1. *Questionnaire design and data collection*

This research is based on the perspective of "VIP consumer customers" to conduct a study on the "key success factors of E-service quality of nutrition consulting organizations." A total of 20 questionnaires were issued, with 17 valid questionnaires, including six males and eleven females.
The questionnaires are all VIP customers of the nutrition consulting agency, and three copies of the questionnaire data collected are deleted. Because the values of CI and CR values are not less than or equal to the standard 0.1, the remaining 17 copies are consistent, so they are included.

4.2. *Research results*

In this study, the collected VIP customer questionnaires are calculated with Microsoft Excel to obtain the weights of each evaluation dimension and each evaluation element and rank them separately, and the weights of each level are shown in Table 2.

Table 2. Overall evaluation results of key elements of E-SQ of nutrition consulting agencies

Dimension	Factor weight (A)	Indicator factor weight (B)		Integration weight (C) = (A)∗(B)
Tangibility	0.1584 (5)	A1	0.3203 (2)	0.0507 (11)
		A2	0.2754 (3)	0.0870 (13)
		A3	0.4043 (1)	0.0461 (8)
Reliability	0.2155 (2)	B1	0.3734 (2)	0.0805 (5)
		B2	0.3984 (1)	0.1017 (3)
		B3	0.2282 (3)	0.0937 (12)
Security/ Privacy	0.2566 (1)	C1	0.4231 (1)	0.1086 (1)
		C2	0.3215 (2)	0.0825 (4)
		C3	0.2554 (3)	0.0655 (6)
Reactivity	0.2034 (3)	D1	0.3106 (2)	0.0632 (9)
		D2	0.4787 (1)	0.0974 (2)
		D3	0.2107 (3)	0.0429 (14)
Efficiency	0.1661 (4)	E1	0.3833 (1)	0.0637 (7)
		E2	0.3629 (2)	0.0603 (10)
		E3	0.2538 (3)	0.0422 (15)

Note: The brackets after the weight numbers are sorting.

This research refers to the number of key success factors proposed by Daniel [13]. He believes that many industries have two to six key factors that determine whether they can succeed. Therefore, the results of this study integrate the weight ranking, the e-service quality of nutrition consulting agencies, and the six most important success factors, shown as follows.

(1) C1 Safe nutritional guidance.
(2) D2 Quick customer question reply.
(3) B2 The nutritionist has sufficient professional training.
(4) C2 Helpful performance improvement.
(5) B1 Professional consultation process.
(6) C3 Safe transaction system provision.

4.3. *Discussion for six key success factors*

4.3.1 Nutritional institutions can provide safe guidance

Nutrition consulting agencies belong to the service of the medical system. Therefore, the customers seeking for assistance mostly want to improve obesity problems, chronic diseases, and other requirements. Therefore, the quality of the nutritionist online service is available, which is the primary key factor for the success of the nutrition consulting agency.

4.3.2 Institutional nutritionists can quickly answer customer questions

Customers can get answers immediately when they encounter problems about online services.
The adaptability of nutritionists and their ability to provide customer trust can assist customers as soon as possible when they encounter difficulties, provide quick service with immediate answers, and make customers feel safe and valued, which is the key to the successful development of E-service quality of nutrition consulting agencies.

4.3.3 Institutional nutritionists have sufficient professional knowledge

Nutritionists have professional knowledge and abilities. Professional knowledge and ability can provide customers with specific and effective online professional plans to achieve the best results of professional plans. In this way, the ability to build customer trust becomes the third important factor for the success of nutrition consulting E-service quality management.

4.3.4 Online service consultation values customer privacy

During the online service of the dietitian, the privacy of the individual case will not be leaked to the third party, and the personal information and privacy of the individual will be protected. The promise of confidentiality can be fulfilled. This is the fourth key success factor in the development of E-service quality for nutrition consulting organizations.

4.3.5 Institutional nutritionist can provide a professional consultation process

The process of consulting services provided by nutrition institutions is crucial to the success of professional programs. Therefore, correct and professional information transmission is very important. This is the fifth key factor for the success of nutrition consulting organizations in developing E-service quality.

4.3.6 The institution provides an online secure transaction system

The organization has a complete online credit card or payment system to allow consumers to have a safe transaction method and guarantee during the payment process. This is the sixth focus of the nutrition consulting organization's success in developing E-service quality.

5. Conclusions and Recommendations

5.1. *Conclusions*

In this study, it is believed that exploring the e-service quality viewpoints provided by VIP customers of the nutrition consulting agencies will help the management and sustainable development of the nutrition consulting and medical industry.

The main findings of this research are as below.

E-service quality of the nutrition consulting agencies in this study is important for the evaluation of the key success factors. In terms of evaluation aspects, "security/privacy" is the most important aspect considered by customers; the second is "reliability"; the third is "reactivity"; the fourth is "efficiency"; and "tangibility" is considered to be less influential.

The adaptability of nutritionists and their ability to provide customer trust can assist customers as soon as possible when they encounter difficulties, provide quick service with immediate answers, and make customers feel safe and valued, which is the key to the successful development of E-service quality of nutrition consulting agencies.

5.2. *Management Implications and Recommendations*

5.2.1. Recommendations to the government

Nutrition consultation is the front-end of preventive medicine. It can be used as the "promoter of preventive medicine and the patron saint of human health." It is recommended that the government pay attention to the function of nutrition consultation and diet care, establish an e-policy and training plan for nutrition consultation, and support the network of the nutrition consultation industry. At the same time, it is recommended to inject the energy of academic research to promote the enhancement of industrial talents' E-level ability, strengthen E-level service quality, online operation management, marketing strategy, and other professional skills, so that the nutrition consulting industry can contribute the most outstanding force in Taiwan's health promotion industry.

5.2.2. Suggestions for the nutrition consulting industry

Business operators should focus on the professional training of nutrition professionals, strengthen professional training on a regular basis, improve the professional knowledge and service quality of nutritionists, and provide customers with online trust and a sense of security, so that professional competence can have the best competition advantage. The growth and advancement of nutritionists in knowledge or intelligence is the best driving force for nutrition consulting companies in their operations. Only with excellent nutritionists and sufficient professionals can nutrition consulting companies create opportunities for sustainable operation.

References

1. “World Health Organization”, https://health.udn.com/health/story/6008/5559421
2. “World Obesity Federation”, https://www.hpa.gov.tw/Pages/Detail.aspx?nodeid=1136&pid=3162
3. Obesity Prevention and Control Network of the National Health Administration of the Ministry of Health and Welfare (2020) http://obesity.bhp.gov.tw
4. Ministry of Health and Welfare (Dietitians Act, 2021)
5. R. Nath, M. Akmanligil, K. Hjelm, T. Sakaguchi, M. Schultz, "Electronic commerce and the internet: Issues, problems, and perspectives”, International Journal of Information Management, Vol. 18, No. 2, pp. 91-101, 1998.
6. Hu Yizhong, Cai Jiawen, Wang Renhong, “Discuss the key service indicators of online banking”, Management Information Calculation, Vol. 6, No. 1, pp. 39-61, 2017.
7. P. R. Varadarajan, M. S. Yadav, “Marketing strategy and the internet: Anorganizing framework”, Journal of the Academy of Marketing Science, Vol. 30, No. 4, pp. 296-312, 2002.
8. Katharine R. Curry, PhD, RD and Amy Jaffe, MS, RD, LD Nutrition Counseling &Communication Skills. pp.48, 1998.
9. V. A. Zeithaml, A. Parasuraman, A. Malhotra, “A Conceptual Framework for Understanding E-Service Quality: Implications for Future Research and Managerial Practice, Cambridge”, Marketing Science Institute, 2000.
10. A. Parasuraman, V. A. Zeithaml, L. L. Berry, “SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality”, Journal of Retailing, Vol. 64, No. 1, pp. 12-40, 1988.
11. V. A. Zeithaml, A. Parasuraman, A. Malhotra, “Service quality delivery through web sites: A critical review of extant knowledge”, Journal of the Academy of Marketing Science, Vol. 30, No. 4, pp. 362-375, 2002.
12. J. Ding, M. Zhou, C. Zhou, “Research on the key service quality factors of Kaohsiung Port International Tourist Center”, Shipping Quarterly, Vol. 27, No. 4, pp. 49-63, 2018.
13. D.R. Daniel, “Management Information Crisis”, Harvard Business Review, Vol. 39, No. 5, pp.111-121, 1961.

Authors Introduction

Ms. Ling-Mei Hsu

She received her Master's degree from the IEMBA, National Kaohsiung University, Taiwan in 2010. She is currently a Doctoral course student in Chang Jung Christian University, Taiwan. She is a person in charge of Xianghe Nutrition Consulting Agency.

The key factors for the application of blockchain into ocean Freight Forwarders.: An Industry Perspective

Chu-Ting Hsu
The Ph.D. Program in Business and Operations Management, College of Management,
Chang Jung Christian University
711301 No.1, Changda Rd., Gueiren District, Tainan City 71101, Taiwan (R.O.C.)

Ming-Tao Chou
Department of Aviation and Maritime Transportation Management, Chang Jung Christian University
711301 No.1, Changda Rd., Gueiren District, Tainan City 71101, Taiwan (R.O.C.)

Ji-Feng Ding
Department of Aviation and Maritime Transportation Management, Chang Jung Christian University
711301 No.1, Changda Rd., Gueiren District, Tainan City 71101, Taiwan (R.O.C.)
E-mail: 109d00027@mail.cjcu.edu.tw, mtchou@mail.cjcu.edu.tw, jfding@mail.cjcu.edu.tw

Abstract

Blockchain is an emergent technology concept that enables the decentralized and immutable storage of verified data and is often considered to be applied to help the maritime industry to manage innovation. This study subdivided the four dimensions into 20 appropriate evaluation indicators. The indicators of key factors based on the results are as follows: 1. 「The application of blockchain will reduce intermediary costs and increase revenue when switching transportation vehicles.」 2. 「Using blockchain will save a lot of manpower and correspondence in the traditional model, and build trust, and fully secure information.」 Ocean freight forwarders can explore the standard service model as a reference to build competitive advantages and ensure sustainable business decisions.

Keywords: Blockchain, Ocean Freight Forwarders, AHP

1. Introduction

The economy of Taiwan is oriented mainly towards import and export trade. Taiwan has achieved good results in the global shipping industry due to the utilization of government resources in cooperation with private enterprises. According to Taiwan's Ministry of Finance and WTO, Taiwan's total trade value in 2020 ranked 17th in the world, and in terms of export and import value, it ranks 15th and 18th, respectively. Whether these import and export trades can effectively achieve international distribution tasks is mainly achieved through Ocean Freight Forwarder－the main shipping assistants of container carriers. In the 21st century, the transportation industry has gradually set off a wave of electronized, mobilized and ubiquitous industrial revolutions to promote smart transportation systems. The non-temperable, traceable, and decentralized features of the blockchain have also begun to receive extensive attention from the international shipping industry. However, the development of blockchain in shipping is still in its infancy. There is no news about the development of the ocean freight forwarders to blockchain[1]. Currently, in the face of high competition and uncertainty in the post-epidemic era, if the use of blockchain can be introduced to enhance the convenience, loyalty and satisfaction of cargo owners in the orientation of profit and high technology can be applied to enhance competitiveness and move towards the era of knowledge economy, continuous support from the shipper can be obtained, which will improve the operating performance, investment profits, and sustainable development of the enterprise in the ocean freight forwarder industry[2]. Since the blockchain is a key factor for the future trend of the ocean freight forwarder industry to save costs and increase convenience, the SAVE marketing structure is used as the main aspect of this study. According to the basic characteristics of the ocean freight forwarder industry and the relevant elements of the blockchain, 20 suitable evaluation indicators are constructed to prepare expert questionnaires. The results of this study will be used as a reference for the ocean freight forwarder industry to improve the direction of future blockchain applications and to find out the key factors for success to plan the future introduction of management strategies.

2. Literature Review

2.1 *Blockchain*

Even though the development of blockchain technology has existed for a long time, there is very little research on the use of blockchain in transportation and supply chain fields. Blockchain technology is likely to change the operation of the logistics and supply chain industries and improve business models to provide more reliable and convenient services to the transportation industry[3]. Blockchain has the following characteristics namely: 1. Decentralization, 2. Openness, 3. Independence, 4. Security, and 5. anonymity. These experiences enable shipping companies to introduce blockchain in transportation services, which will effectively improve operational efficiency and transaction security and promote the competitiveness of shipping companies in the digital age[4].

2.2 *SAVE Model of Marketing Architecture*

In the modern transportation market, there are many types of business to business (B2B) transactions. Since McCarthy founded the 4Ps (product, price, promotion, and place) marketing theory[5], scholars at home and abroad have since been based on this.

Table 1. The evaluation aspects and references of the key factors affecting the use of blockchain in the ocean freight forwarder industry.

Dimensions	No.	Indicators
SOLUTIONS	A1	It can prevent counterfeiting of bills of lading and solve the problem of counterfeiting during the transfer bills of lading.
	A2	Blockchain technology can be used as the guarantee of credit for both sides of international trade transactions.
	A3	The transparency and unforgeability of blockchain records will lead to trust in the context of decentralization.
	A4	The encryption and decryption mechanism by using blockchain technology can protect information security and privacy, which can effectively reduce theft or mistaking of goods.
	A5	Integrating the exchange of information between supply chain participants can save time and simplify paperwork related to cross-border international trade processes.
ACCESS	B1	Participants in the supply chain can quickly track and understand the current status of goods conveniently in real time.
	B2	By using blockchain traceability technology, participants can realize cross-border union into shared trust under the influence of smart contracts.
	B3	Customers can instantly and conveniently inquire logistics-related costs, freight prices and integrated quotation services.
	B4	Customers can immediately and conveniently understand the full-stage logistics arrangement, which is known as intermodal service, and let the participants in the supply chain know the current status at any time.
	B5	Customers can keep abreast of the current shipping space status and container information, which is convenient for participants to dispatch and respond to the dispatching needs of transportation vehicles at any time.
VALUE	C1	The application of blockchain will reduce intermediary costs and increase revenue when switching transportation vehicles.
	C2	It can be more efficient and reliable to make cargo transportation safer and more professional.
	C3	C3 Blockchain technology tracking can improve the efficiency of customer logistics, thereby attracting customers to use the company's service and creating service differentiation.
	C4	The success of invoking blockchain technology will attract customer support and greater confidence in the company's goodwill and financial position.
	C4	Using blockchain will save a lot of manpower and correspondence in the traditional model, build trust, and fully secure information.
EDUCATIONS	D1	Blockchain can provide customers with relevant market business information on time.
	D2	Blockchain can automatically update the special cases and solutions in the transportation process for customers' reference.
	D3	The introduction of blockchain enables employees to respond quickly to customer needs for improvement.
	D4	Blockchain makes the continuing professional education, training of operators, and sales easier.
	D5	Blockchain makes it easier to share international trade customs regulations and maritime knowledge online.

With the rise of the Internet, the use of Internet information and multimedia functions to provide a large number of opportunities for exposure has also begun to change the way of marketing. In response to market changes, Ettenson et al. proposed a SAVE (solution, access, value, and education) marketing architecture based on past marketing theories. From SAVE's solutions to education, everything is customer-centric, which is lacking in McCarthy's 4Ps. The SAVE marketing model is not only customer-centric, but all customers can work together to achieve their goals[6]. This concept just coincides with all participants in the blockchain. Through decentralized accounting and storage, each node achieves self-verification, transmission and management of information in a concerted manner.

The ocean freight forwarder industry has five important functions for cargo owners, including consultation, cost reduction, time saving, convenient transportation and improvement of product competitiveness, which indicates that the ocean freight forwarder industry is in the position of guiding, providing consulting services and arranging and coordinating processing during the transportation process[7]. Finally, in this article, the successful elements of the blockchain introduced into the ocean freight forwarder industry are summarized as solution, access, value, and education. The SAVE model almost includes all the success factors required in the current business model. The better the performance of these four dimensions, the better the relationship between the company and its customers will be established. Combining the SAVE architecture of solutions, market access, professional value and education proposed by Ettenson et al., in accordance with the evaluation objectives, dimensions and criteria, an evaluation framework for the influencing factors of the ocean freight forwarder industry to choose the import of the blockchain has been constructed, as shown in Table 1.

3. Research Methods

As a decision-making method systematizing complex problems, AHP decomposes complex problems into dimensions and factors and concludes the weights of all factors after integrating experts' opinions. AHP is mostly used in uncertain situations and multi-attribute decision-making problems.[8]: In evaluating the decision maker's judgment or measuring the overall hierarchical structure, Saaty suggested that the consistency ratio shall not be greater than 0.1 to ensure consistency. [9]

In this study, the hierarchical structure is constructed by the analytic hierarchy process, the questionnaire of "The key factors for the application of blockchain into ocean Freight Forwarders.: An Industry Perspective" is designed, a pairwise comparison matrix is adopted to analyze hierarchical factors such as effects and evaluation indicators, and a quantitative evaluation scale is used (1 to 9 scores). Furthermore, the weights of evaluation criteria are pairwisely compared on a scale from 1 to 9. As pairwise comparison is one of the most effective methods in making judgments, this study establishes a pairwise comparison matrix and calculates the eigenvalue and eigenvector. Finally, the consistency test is carried out with the maximum eigenvalue to obtain the relative weights of the evaluation criteria.

After questionnaires are collected, "Expert Choice11" is applied to calculate the weights and the relevance of layers and indicators. In addition, a consistency test is carried out to select valid samples with CR $\leqq$ 0.1, and then the weights of valid samples are sorted and analyzed. A consistency test must be carried out for the weights of indicators to ensure that the respondents think the same way before and after the pairwise comparison. Then, the weights of indicators in layers are calculated according to the results to identify the key factors for the application of blockchain into ocean Freight Forwarders.

4. Empirical Analysis

The AHP questionnaire consists of 3 parts: the first part includes the basic data of respondents, including title, job seniority and company and department; the second part includes answer examples and descriptions of dimensions and criteria; while the third part evaluates the relative importance of the key factors for the application of blockchain into ocean Freight Forwarders.

In this study, questionnaires are issued to managers who have decision-making rights and management from top companies in the ocean freight forwarding industry, based on Taiwan's top 500 service companies recorded by Common Wealth Magazine. Particularly, the key factors for the application of blockchain into ocean Freight Forwarders. are explored from the perspective of large or well-known ocean freight forwarders in Taiwan. In this study, a total of 9 questionnaires are issued, 8 are collected, 8 are valid, and respondents of valid questionnaires meet the standard consistency ratio, that is, not greater than 0.1. They have at least 15 years of experience in the shipping industry, and most of them are managers with more than 20 years of experience. According to the results calculated by "Expert Choice

11", the weights of dimensions of service quality are as follows: (1) the weight of "SOLUTIONS" is 0.346, (2) the weight of "ACCESS" is 0.176, (3) the weight of "VALUE" is 0.397, (4) the weight of "EDUCATIONS" is 0.081. Among them, "VALUE" is the most important, while "EDUCATIONS" is the least important. A C.I. of 0.02<0.1 indicates that it passes the consistency test, indicates that the matrix consistency is satisfactory.
It will be evaluated according to four system quality dimensions. From the experimental results, we can see the importance of the major dimensions. Overall, experts attach great importance is the application of blockchain will reduce intermediary costs and increase revenue when switching transportation vehicles and using blockchain will save a lot of manpower and correspondence in the traditional model, and build trust, and fully secure information.

5. Conclusion

In this study, indicators are mainly obtained from the judgment on the weights of applicable indicators by "AHP", the four dimensions of SAVE marketing are developed through research to construct the key factors for the application of blockchain into ocean Freight Forwarders, while the weights of all measurement factors are ranked. The process of this study is as follows:
(1) The last 20 appropriate measurement factors are developed by literature review and expert interviews.
(2) In this paper, AHP is used to demonstrate that the top 2 factors with the largest weights are 1. 「The application of blockchain will reduce intermediary costs and increase revenue when switching transportation vehicles.」 2. 「Using blockchain will save a lot of manpower and correspondence in the traditional model, and build trust, and fully secure information.」 The suggestions can build and enhance their own value and resources, actively strengthen service quality and core competitiveness, so as to improve customer support and satisfaction, and achieve sustainable corporate business.

References

1. T.-C. Ho, C.-L. Hsu. An Analysis of Analysis of Key Factors Influencing Integration of Blockchain into Shipping Companies in Taiwan. *Journal of Marine Science and Technology,* **28**(4), 2020, pp. 229-236.
2. H.-Lin, C.-C Chung, C.-H Huang. Benfits of Applying Blockchain Technology to Maritime Container Transportation Operations. *Journal of Marine Technology*, **4**, 2020, pp. 27-40.
3. J.-L. Zhao, S.-K. Fan, J. Yan. Overview of business innovations and research opportunities in blockchain and introduction to the special issue. *Financ Innov*, **28**(2) 2016, pp. 1-7.
4. H.-L. Tsai, T-R, Lin, C.-C., Chang. The Application of Innovative Technologies in Shipping and Port Operations. *Maritime Quarterly,* **28**(2) ,2019 pp. 75–96.
5. McCarthy, E. J., Basic Marketing: A Managerial Approach, Homewood, Illinois, 1960.
6. Ettenson, Richard, E. Conrado, J. Knowles, Rethinking the 4Ps, Harvard Business Review, January-February 2013.
7. J.-P. Zeng, L.-Z. Liao, Theory and Practice of Ocean Freight Forwarding, Wunan Culture Enterprise, Taipei. 2010.
8. G.-S. Liang, J.-F. Ding, S.-M. Wang, Quantitative Research Methods: Theory and Applications. 1th(ed), Kee-lung. 2021.
9. T. L. Saaty, The Analytic Hierarchy Process. McGraw-Hill, New York, 1980.

Authors Introduction

Ms. Chu-Ting Hsu

She received her Master's degree from Shipping and Transportation Management, National Kaohsiung University of Science and Technology, Taiwan in 2009. She is currently studying for a doctoral program at Chang Jung Christian University in Taiwan.

Dr. Ming-Tao Chou

He is a Professor of Department of Aviation and Maritime Transportation Management at Chang Jung Christian University in Taiwan.

Dr. Ji-Feng Ding

He is a Professor of Department of Aviation and Maritime Transportation Management at Chang Jung Christian University in Taiwan.

Key Success Factors for Implementation Quality Assurance of Information Technology in Tourism Industry

Shuen-Huei Yao
Department of Accounting and Information Systems, Chang Jung Christian University,
No. 1, Changda Rd., Gueiren District, Tainan City, 711301, Taiwan

Cheng-Chung Yeh
Center for General Education, National University of Tainan,
No.33, Sec. 2, Shu-Lin St., Tainan City, 700301, Taiwan

Wen Jung Tsai
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: shyao@mail.cjcu.edu.tw, nutn06099@gn.edu.tw, m2.nuttsai@ctbc.edu.tw

Abstract

With the constant change nowadays, various enterprises are facing such a fierce competition that they have to seek for competitive advantages and core competitiveness. With respect to consumers, the implementation quality assurance based on the technology system platform in tourism industry. In order to figure out for implementation quality assurance of information technology in tourism, the Delphi method is applied in this research. For the weights and priorities of various measurement indicators are compared through AHP. According to the results, "training professional tour guides", "foreign language and translation skills" are the most critical success factors for the implementation quality in tourism industry. Besides, it hoped results can provide suggestions for the actual implementation and management of the platform.

Keywords: Key Success Factors (KSF); Delphi Method; Analytic Hierarchy Process (AHP); Quality Assurance

1. Introduction

Generally speaking, the larger the number of tourists, the longer the travel time, and the higher the consumption amount, thereby creating tourism output value, bringing about local job opportunities and promoting national economy. As for the information technology in tourism industry, a proper business operation is both significant to the tourism industry and critical for exploring the key factors for implementation quality assurance. Therefore, it is a great help to tourism industry.

1.1 Research Background and Motivations

Tourism is one of the important industries to a country. At present, information technology is developing vigorously in tourism industry around the world, which means that all countries are actively promoting the high-tech of tourism industry. From an economic perspective, the technological development in tourism contributes to gaining foreign exchange income, accelerating economic prosperity, and increasing employment opportunities. In recent years, the improvement of people's living quality has highlighted the importance of tourism quality in tourism industry.

However, due to the lack of business philosophy and direction, the tourism market in Taiwan only focuses on the business amount rather than look into the factors determinant to the operating performance of travel agencies. Thus, for the sake of finding out the key factors for the implementation quality assurance of travel technology, the domestic travel agencies in tourism industry are taken as the research object. It is also expected that the research results can help the tourism industry to create a targeted enterprise operation information technology strategy.

Based on the above-mentioned motivations and background, this research aims to not only study the key success factors of implementation quality assurance of information technology in the tourism industry, but also provide reference for future business strategies of relevant government agencies, tourism industry and researchers through travel agencies. Specifically, the research objectives include: 1. analyzing various information business strategies and technology implementation status in tourism industry; 2. learning about the key success factors of tourism practitioners' implementation quality of information technology in tourism industry; 3. improving tourism practitioners' mastery of key success factors and business performance of tourism industry.

2. Literature Review

By means of discussing the previous literatures, this chapter sorts out the relevant methods and results, which are regarded as a reference to this research.

2.1 Introduction to Key Success Factors in Tourism

As the significant factor influencing and leading the long-term competitiveness of an enterprise in its engaged fields, Key Success Factor (KSF) or Critical Success Factor (CSF) was initially proposed by (Daniel, 1961), and it current has many different interpretations and definitions. Commonly used methods for assessing key success factors include Delphi, Factor Analysis, Regression analysis, and Analytic Hierarchy. The research on the key success factors for business strategy of travel agency indicates that "development of e-commerce" in terms of costing is considered as an important factor. The research results can provide reference for travel industry players in learning about the industrial environment as well as enterprises in formulating business strategies.

2.2 Section 2 Quality Assurance

Quality is to meet the needs of consumers and maximize the consumption satisfaction. In regard to the service industry, quality refers to consumers' positive evaluation on their perceived products or services after comparing with the previous expectations (Murphy et al., 2000). That is to say, if the performance of a product (or service) is equal to or higher than consumers' expectations, the perception of "quality" will be generated (Fick and Ritchie, 1991). In other words, quality is consumers' evaluation on the product or the overall service provided by enterprises.

3. Research Methods

In this research, Delphi method combined with AHP is adopted to explore the key success factors for the implementation quality assurance of information technology in tourism industry, and the expert questionnaire survey is also conducted. Based on the selected variables, this research is divided into three hierarchies. The first hierarchy is "Key Success Factors for Implementation Quality Assurance of Information Technology in Tourism". Through consulting the relevant literatures, the second hierarchy classifies the critical success factor into dimensions including personnel, travel agency, support service, training and counseling, which are used as the theoretical basis for this research and the structure of questionnaire design. The third hierarchy continues the second level and carries out integrated analysis of the expert questionnaires in first and second hierarchies (Saaty, 1980).

With the help of Delphi method combined with AHP, the key success factors for the implementation quality assurance of information technology in tourism industry are studied. By taking the literatures influencing quality assurance as reference, the questionnaire is designed for data collection after consulting industry experts. The key factors are divided into four dimensions and 26 elements. A total of four questionnaire surveys are conducted in this research so as to obtain the weight of the key factor structure, and then to establish an evaluation scale of key success factors for the implementation quality assurance in tourism industry.

4. Research Results and Statistical Analysis

In the first round of statistical analysis based on Delphi method, a semi-open questionnaire survey is applied to comprise the personal views of all participants and allow fully expressing opinions in the questionnaire so as to supplement the deficiencies. With respect to the questionnaire response design, the importance degree is considered and the Likert five-point scale is taken to check the importance of each indicator, and combine with the ranking of key factors of behavior for analysis through AHP. Among the 12 questionnaires collected, 2 are found to fail the verification. Therefore, there are actually 10 valid questionnaires, accounting for 83% of the questionnaires collected. Based on the questionnaire survey and AHP analysis of weights at all levels, an AHP framework for the key success factors of implementation quality assurance of tourism information technology is established. Under the support of the Delphi expert questionnaire and AHP questionnaire survey results, the structure of the key success factors for the implementation quality assurance of tourism information technology is finally developed, and the weight as well as priority of the key factors at different dimensions and between the measurement indicators are analyzed.

Table 1 Overall Weight Value of Each Key Success Factor for Tourism Implementation Quality Assurance

	Dimension	Key Success Factor	Weight	Ranking of Importance Degree
Key Success Factor for Tourism Implementation Quality Assurance	Cultivation	Training Professional Tour Guides	0.28781	1
	Personnel	Privacy of Tourists' Information	0.12259	2
	Supporting Service	Time for Visa	0.071	3
	Travel Agency	Tourism Healthcare and Safety	0.06105	4
	Personnel	Providing Accurate Consumer Data	0.0524	5
	Supporting Service	Market Survey and Research	0.04717	6
	Travel Agency	Professional Tourism Information	0.04389	7
	Cultivation	Foreign Language and Translation Skills	0.04257	8
	Travel Agency	Tourism Risk Assessment	0.03736	9
	Supporting Service	Promotion of Tourism Products	0.03364	10
	Travel Agency	Realizing Promises to Consumers and Establishing Good Reputation	0.03271	11
	Personnel	Attaching Importance to Consumers' Complaints	0.02728	12
	Personnel	Respecting Consumers	0.02559	13
	Cultivation	Offering Orientation Training and Setting up Staffs' Confidence in Job	0.02394	14
	Supporting Service	Treatment and Avoidance of Travel Disputes	0.01969	15
	Supporting Service	Controlling Tourism Product Quality	0.01925	16
	Cultivation	Enhancing Staffs' Sense of Belonging to Companies	0.0121	17
	Travel Agency	Awards to Staffs	0.00313	18
	Personnel	Friendly Service Attitude	0.00087	19

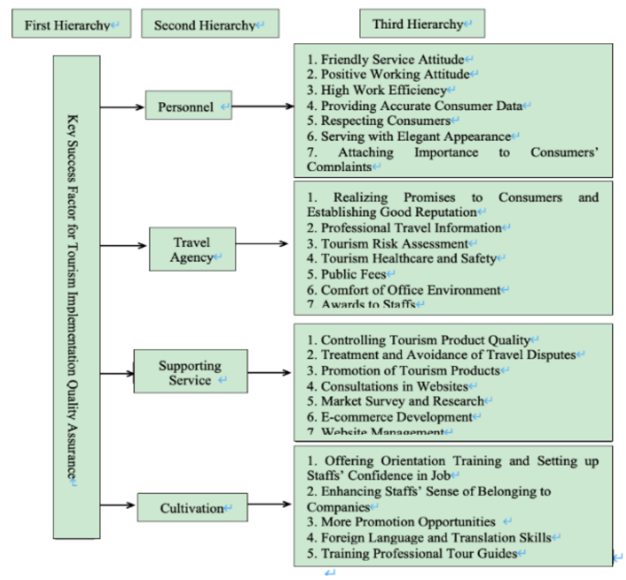


Figure 1 Structure of Research on Key Success Factor for Tourism Implementation Quality Assurance

In this study, the key success factors for the implementation quality assurance of information technology in tourism industry are structured through different approaches such as literature discussion and expert interviews. Based on three times of expert questionnaire surveys by Delphi method, the initial 26 measurement indicators in 4 dimensions are finally summarized into 19 measurement indicators in 4 dimensions. Next, the weight ratio and priority are obtained through AHP questionnaire. In view of too many indicators and for the sake of highlighting the important key factors, the indicators with weight ratio above 0.035 are listed as the key success factors.

5. Conclusions and Suggestions

With the purpose to find out the key success factors for the implementation quality assurance in tourism industry, this research invites experts to assess the importance of each dimension and the implied key factors through Delphi method, so as to establish an AHP hierarchy. Under this hierarchical structure, each key success factor is discussed to obtain the key success factors for the implementation quality assurance of information technology in tourism industry.

This chapter mainly describes the research results and suggestions, and finally proposes suggestions for follow-up research.

5.1 Conclusions

As concluded in this research, the key success factors for the implementation quality assurance of information technology in tourism industry are ranked based on the importance degree as follows, namely, “Training Professional Tour Guides”, “Foreign Language and Translation Skills”, “Offering Orientation Training and Setting up Staffs’ Confidence in Job”, “Time for Visa”, “Market Survey and Research”, “Enhancing Staffs’ Sense of Belonging to Companies”, “Privacy of Tourists’ Information”, “Promotion of Tourism Products”, and “Tourism Healthcare and Safety”. It is hoped that through the results can help the tourism industry avoid mismanagement during information technology operations. In this way, not only consumers will have greater trust in choosing a travel agency, but also the tourism practitioners can make a difference in management, thereby achieving continuous business.

5.2 Suggestions

Considering that this research mainly focuses on information technology in tourism industry, the samples are limited to a specific range. Thus, subsequent researchers may perform key success factor assessments for quality assurance in other industries if more samples can be obtained from different industries, so that the implementation quality in different industries to key success factors can be studied. Meanwhile, with reference to the more weighted consideration levels or evaluation criteria listed in this research, follow-up researchers can carry out in-depth study so as to not only facilitate the assessment on key success factors for the implementation quality of information technology but also enrich related theories.

References

1. Daniel DR. 1961. Management information crisis. Harvard Business Review 39:111-121.
2. Fick GR, Ritchie JB. 1991. Measuring service quality in the travel and tourism industry. Journal of Travel Research 30:2-9.
3. Murphy P, Pritchard MP, Smith B. 2000. The destination product and its impact on traveller perceptions. Tourism management 21:43-52.
4. Saaty TL, Kainen PC. 1977. The four-color problem. McGraw-Hill New York.
5. Saaty TL. 1980. The analytic hierarchy process: Planning, priority setting, resources allocation. McGraw-Hill.

Authors Introduction

Mr. Wen-Jung Tsai

He received his Master's degree from EMBA, Chang Jung Christian University, Taiwan in 2014. He is currently studying for a doctoral program at Chang Jung Christian University in Taiwan.

Dr. Shuen-Huei Yao

He is a Professor of Department of Accounting and Information Studies at Chang Jung Christian University in Taiwan.

Dr. Cheng-Chung Yeh

He is an Assistant Professor at the General Education Center of Tainan University.

AI Big data analysis and application: Patient Safety Culture of Nursing Staff in an Operation Room

Su-Chiu Hsiao

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: n790115@mail.chimei.org.tw

Abstract

The objectives of this study were to investigate patient safety attitudes among operation room (OR) nursing staff by using the Safety Attitudes Questionnaire (SAQ) and to compare their safety attitudes with the entire hospital nurses'. This study investigated the factors affecting patient safety attitudes by the multivariate regression models. There was a significant association between resilience and safety climate after adjusting for age, gender, and career experience using the multiple-regression model. Perception of management was also significantly associated with job satisfaction. The survey data provide baseline information on OR nurses. The results can be used to follow-up on the effectiveness of quality improvement campaigns, caring Strategy, and patient safety education in the future. Establish a caring notification system. Achieve effective management.

Keywords: operating room, nursing staff, patient safety culture, safety attitudes scale

1. Introduction

"To Err Is Human," which is IOM's report, has aroused international attention to patient safety issues. The risk of medical errors in the operating room is higher than that of other medical units. However, the occurrence of medical errors is related to organizational culture, personal beliefs, attitudes, and values. Patient safety culture is considered to be the key to patient safety. Besides, safety culture surveys can assist in understanding patient safety issues, assessing personnel safety attitudes, and creating a safe medical environment that supports patients. Many patient safety studies have shown that improvements in safety attitudes and safety culture are significantly related to the reduction of medical errors. The purpose of this study is to understand the safety culture of nursing staff in the operating room and the factors that may affect the safety culture, to further compare with the data of the whole hospital, and to explore the difference between the safety of patients in the operating room and the whole hospital[1,2,3,6,8].

2. Materials and Methods

2.1. *Research tools: SAQ (Safety Attitudes Questionnaire)*

The Chinese version of the SAQ questionnaire has 40 questions in five dimensions-"teamwork climate," "safety climate," "job satisfaction," "perception of management, and "working conditions," distinguished according to Likert scale into strongly disagree, somewhat disagree, neutral, somewhat agree, and strongly agree. The content also includes basic demographic changes.

2.2. *Research objects*

The research objects are the nurses at a medical center and it takes about 5 minutes to fill out the online questionnaire.

2.3. *Data analysis*

The scores of each dimension are calculated as the average scores of the questions in the dimension, showing the scores of the safety attitudes of nurses of different genders, ages, and working years in the operating room.
By nonparametric statistics, differences in scores of various dimensions of safety attitudes of nurses in operating rooms of different genders, ages, and years of

work are verified. By multiple regression, the influence of the SAQ dimension scores of the nurses in the operating room is explored. In order to compare the SAQ score difference between the operating room and the whole hospital nurses, the chi-square test is adopted to check whether the distribution of demographic characteristics is consistent to ensure the comparability of the nurses in the operating room with those of the whole hospital, so as to further explore whether different genders, ages, working years, and supervisor backgrounds of the nurses in the operating room and the whole hospital nurses affect the difference in scores.

3. Results

A total of 1613 nurses in the hospital answered, and the recovery rate was 69.6%; 102 nurses in the operating room answered, and the questionnaire response rate was 63.5%, with the internal consistency of the questionnaire 0.65 to 0.89. There is no significant difference between the operating room and the whole hospital in terms of supervisory positions, age, and working years of nurses ($p> 0.05$) (Table 3).

The scores of nurses in the operating room on each dimension are "teamwork climate" (71.6 ± 11.3), "safety climate" (69.3 ± 9.5), "job satisfaction" (67.5 ± 11.0), "working conditions" (67.8 ± 10.3), and "perception of management" (65.5 ± 10.2). There is no difference in the scores of nurses in operating rooms of different age groups. Operating room nurses with a working experience of 2 two to ten years have scores lower than those of more than ten years and less than two years, reaching significant difference on working conditions (61.2 vs. 75.1 vs. 73.3, $p <0.01$), shown in Table 1.

Table 1. SAQ dimension scores of the nursing staff in the operating room

	Teamwork Climate	*p* value	Safety Climate	*p* value	Job Satisfaction	*p* value	Perception of Management	*p* value	Working Conditions	*p* value
Nurses	71.6 (11.3)		69.3 (9.5)		67.5 (11.0)		65.5 (10.2)		67.8(10.4)	
Gender		1.00		0.25		0.06		0.13		0.03
Male	81.6 (7.5)		77.0(15.1)		78.5 (2.2)		75.0 (8.7)		81.5 (7.2)	
Female	66.0 (9.2)		68.1 (9.0)		66.7 (11.2)		65.0 (10.3)		66.3 (9.5)	
Supervisor Position		0.23		0.91		0.96		0.12		0.22
Yes	80.1 (-)		68.5 (-)		68.1 (-)		85.0 (-)		80.1 (-)	
No	67.2 (10.0)		69.4 (9.1)		67.8 (11.4)		65.3 (10.0)		67.1 (10.3)	
Age (years old)		0.18		0.81		0.91		0.64		0.49
< 31	67.6 (9.4)		68.5 (10.3)		68.1 (11.5)		65.0 (8.9)		67.5 (9.6)	
31~40	65.8 (14.9)		70.9 (10.65)		65.5(13.6)		68.0 (16.0)		65.9 (14.4)	
> 40	75.1 (7.1)		67.1 (2.1)		68.5(0)		72.5 (17.7)		75.5 (7.3)	
Working Experience (years)		0.19		0.33		0.21		0.11		<0.01
< 2	73.3 (10.1)		70.9 (10.3)		72.6 (9.2)		69.1 (10.5)		73.6 (10.2)	
2~10	61.2 (7.2)		66.1 (9.7)		65.1 (11.0)		61.5 (9.6)		61.5 (7.1)	
> 10	75.1 (9.0)		72.8 (7.5)		68.1 (9.8)		73.2 (11.0)		75.2 (9.3)	

Among the factors that affect the scores of nurses in the operating room, gender, supervisor position, age and working experience have no significant differences in the scores of nurses in the operating room. The scores of "teamwork climate" and "safety climate" of nurses in the operating room are affected by each other's dimensions. The scores of "job satisfaction" and "perception of management" also influence each other due to different job satisfaction or management perception degrees, as shown in Table 2.

The scores of nurses in the operating room in all dimensions are significantly lower than those of in the whole hospital. According to different gender, whether to hold a supervisory position, age, and working experience, the difference in scores between operating room nurses and whole hospital nurses is compared. Females, non-supervisors, and operating room nurses with 2-10-year working experience have significantly lower scores in each dimension (p <0.05), as shown in Table 3.

Multiple regression is used to analyze the influence of variables on the dimension score. Work unit is one of the factors that affect the dimension score. The scores of the nurses in the operating room in “teamwork climate" (= 75.3, SE = 1.1, R2 = 0.05), “safety climate" (= 74.2, SE = 1.2, R2 = 0.05), “perception of management" (= 70.2, SE = 1.5, R2 = 0.03), and "working conditions" (= 72.5, SE = 1.5, R2 = 0.02) are significantly lower than those in the whole hospital. As the age group decreases, the scores of the nurses in the operating room also decrease.

Table 2. Regression analysis of the influence on patient safety culture in the operating room

	Teamwork Climate			Safety Climate			Job Satisfaction			Perception of Management			Working Conditions		
Intercept	*β*(SE)	*p* value	R^2	*β*(SE)	*p value*	R^2	*β*(SE)	*p* value	R^2	*β*(SE)	*p* value	R^2	*β*(SE)	*p* value	R^2
	1.8 (12.8)	0.88	0.49	14.9 (12.7)	0.27	0.50	-3.7 (12.5)	0.78	0.63	7.7 (12.2)	0.54	0.64	26.3 (10.6)	0.02	0.62
SAQ Dimensions															
Teamwork Climate	–			0.5 (0.2)	0.01		0.2 (0.2)	0.23		0.1 (0.2)	0.55		< 0.1 (0.2)	0.95	
Safety Climate	0.8 (0.2)	< 0.01		–			0.2 (0.2)	0.36		-0.1 (0.2)	0.72		0.1 (0.2)	0.53	
Job Satisfaction	-0.1 (0.2)	0.68		0.2 (0.2)	0.36		–			0.5 (0.2)	0.03		0.2 (0.2)	0.43	
Perception of Management	0.5 (0.3)	0.09		-0.1 (0.2)	0.71		0.5 (0.2)	0.03		–			0.3 (0.2)	0.16	
Working Conditions	< 0.1 (0.3)	0.98		0.2 (0.2)	0.53		0.2 (0.2)	0.42		0.3 (0.2)	0.16		–		

Table 3. Comparison of SAQ scores of the staff between the operating room and the hospital

	Teamwork Climate		Safety Climate		Job Satisfaction		Perception of Management		Working Conditions	
	OR/All	*p* value	OR/All	*p* value	OR/All	*p* value	OR/All	*p* value	OR/All	*p* value
Gender										
Male	81.8(7.6) / 70.0 (11.5)	0.18	77.0 (15.0) / 78.2 (3.2)	0.89	78.7 (2.3) / 81.1 (5.0)	0.50	75.0 (8.7) / 75.1 (12.9)	1.00	81.7 (7.6) / 70.1 (11.5)	0.18
Female	66.3 (9.4) / 73.3 (13.2)	< 0.01	68.2 (9.0) / 76.2 (11.7)	<0.01	66.5 (11.3) / 73.2 (17.3)	< 0.01	65.0 (10.3) / 71.4 (14.8)	<0.01	66.1 (9.3) / 73.2 (13.3)	<0.01
Supervisor Position										
Yes	80.1 (-) / 80.5 (10.6)	0.96	68.8 (-) / 85.7 (10.2)	0.14	68.0 (-) / 86.4 (15.6)	0.27	85.1 (-) / 82.3 (11.9)	0.83	80.0 (-) / 80.7 (10.6)	0.96
No	67.6 (10.2) / 72.9 (13.5)	0.02	69.1 (9.8) / 75.9 (11.5)	<0.01	67.6 (11.5) / 72.7 (17.1)	0.03	65.3 (10.0) / 71.5 (14.7)	<0.01	67.3 (10.0) / 72.8 (13.5)	0.02
Age (years old)										
< 31	67.5 (9.5) / 72.3 (13.1)	0.08	68.5 (10.2) / 74.7 (11.7)	0.02	68.4 (11.6) / 70.7 (17.4)	0.28	65.1 (8.9) / 70.2 (14.1)	0.02	67.4 (9.4) / 72.2 (13.0)	0.08
31~40	65.7 (14.8) / 73.5 (13.1)	0.21	70.9 (10.6) / 77.6 (11.3)	0.22	65.4 (13.3) / 74.4 (16.5)	0.24	68.1 (16.0) / 71.5 (15.9)	0.63	65.6 (14.9) / 73.8 (13.3)	0.21
> 40	75.1 (7.1) / 77.0 (14.8)	0.83	67.1 (2.0) / 80.7 (12.6)	0.15	68.1 (0) / 81.2 (15.0)	< 0.01	72.75 (17.7) / 77.4 (13.9)	0.63	75.1 (7.2) / 77.3 (14.7)	0.83
Working Experience (years)										
< 2	73.4 (10.3) / 72.3 (13.6)	0.77	70.9 (10.3) / 74.6 (10.7)	0.29	72.8 (9.2) / 70.0 (17.6)	0.41	69.1 (10.5) / 70.3 (13.6)	0.78	73.5 (10.5) / 72.2 (13.7)	0.77
2~10	61.6 (7.1) / 73.4 (12.6)	< 0.01	66.3 (9.7) / 76.2 (12.5)	< 0.01	65.1 (11.0) / 73.5 (16.6)	0.04	61.8 (9.7) / 71.7 (15.01)	< 0.01	61.5 (7.1) / 73.4 (12.5)	<0.01
> 10	75.1 (9.2) / 74.5 (14.0)	0.94	72.9 (7.6) / 78.7 (11.9)	0.33	68.0 (9.8) / 77.6 (16.9)	0.27	73.6 (11.2) / 73.3 (15.0)	0.93	75.0 (9.2) / 74.4 (14.2)	0.94

4. Discussion

Due to the special working environment, the risk of medical errors in the operating room is higher than that in other medical units. The American College of Emergency Physicians (ACEP) pointed out that the risks of medical errors in the operating room include crowded medical environment, communication barriers for patients or their families, shortage of medical manpower, failure in prediction of workload factors, poor patient health, failure in the establishment of a long-term medical relationship with patients as the outpatient routine, and the condition that physicians need to make rapid medical treatment decisions in emergencies[5].

Research results show that operating room nurses have the highest scores in the "teamwork climate" dimension. The operating room is the epitome of the hospital and summarizes all medical services. Due to the uncertain and unpredictable characteristics of operating room medical care, a good teamwork is needed. In this research, "teamwork climate" and "safety climate" are the dimensions of mutual influence. "Job satisfaction" and "perception of management" are also significantly related in this study, showing that the satisfaction of nurses has a considerable relationship with the management mode of the unit supervisor. Junior and senior operating room nurses scored higher in all dimensions and it is speculated that it may be related to the high turnover rate of nursing manpower in the operating room. The operating room is a high-pressure, high-risk, and fast-paced medical environment. Different medical service features of the operating room will automatically select the appropriate nurses. Experienced nurses themselves have more awareness of patient safety, and it is not easy for junior nurses to relax in dealing with urgent and serious situations[3,6].

In this study, the scores of patient safety attitudes of operating room nurses are lower than those of the whole hospital. It is necessary to cooperate with other measurement tools to analyze the correlation with the medical event to clarify whether the safety of the patient is truly affected. The research results are only suitable for individual hospitals to understand the perceptions of unit nurses' patient safety attitudes, which cannot be fully extended to all hospitals. Medical institutions can refer to this research method and analysis tools for further in-depth discussion. It is recommended that the scores of the patient safety culture scale can be used in the future to compare and discuss patients' prognostic results, the occurrence of medical error, and changes in behavior of medical staff for more comprehensive information[7].

In this study, SAQ is used to investigate patient safety culture, which can not only understand patient safety attitudes of nurses in the operating room but assist in tracking the connection between the safety attitude of operating room personnel and patient safety with quality management methods and medical information system interventions in the future so as to improve the safety culture of the operating room[4].

References

1. Ahmed, N.G., Adam, S.M., & Abd Al-Moniem, I.I. (2011). Patient safety: assessing nurses' perception and developing an improvement plan. life science journal, 8(2), 53-64.
2. Colla JB, Bracken AC, Kinney LM, Weeks WB. (2005). Measuring Patient Safety Climate: A Review of Surveys. Qual Saf Health Care, 14(5): 364-366.
3. Charles R., Denham J. (2005). Patient Safety-Patient safety practices: Leaders Can Turn Barriers Into Accelerators. Journal of Patient Safety, 1(1):41-55.
4. Kim, M., & Windsor, C. (2015). Resilience and work-life balance in first-line nurse manager. Asian nursing research, 9(1), 21-27.
5. Lee, Y.C., Huang, C.H., Weng, S.J., Hsieh, W.L., & Wu, H.H. (2016). Analyzing emotional exhaustion from viewpoints of physicians and nurses–A case of a regional teaching hospital. TEMJ, 5(2), 231-235.
6. Linda T. Kohn, Janet M, Corrigan, Molla S. Donaldson. To Err Is Human-Building a safer health System, National Academies Press, pp15-58, 2000.
7. Roohi, G., Asayesh, H., Rahmani, H., & Abbasi, A. (2011). The relationship job satisfaction and organizational commitment of nurses in hospitals of Golestan University of Medical Sciences. Journal of Payesh, 10(2), 285-292.
8. The Joint Commission. Sentine l Event Alert 57: The Essential Role of Leadership in Developing a Safety Culture. (2017). Cited April 14, 2018. Available at: https://www.jointcommission.org/assets/1/18/SEA_57_Safety_Culture_Leadership_0317.pdf

Authors Introduction

Ms. Su-Chiu Hsiao

She received her Master's degree from the Department of Nursing, Chang Jung Christian University, Taiwan in 2008. She working as a nursing supervisor at Chi Mei Hospital. She is currently a Doctoral course student in Chang Jung Christian University, Taiwan.

Experience in Efficient Real
Office Environment Modelling in Gazebo: a Tutorial

Bulat Abbyasov
*Intelligent Robotics Department, Kazan Federal University, 35 Kremlyovskaya street,
Kazan, 420111, Russian Federation*

Kirill Kononov
*Intelligent Robotics Department, Kazan Federal University, 35 Kremlyovskaya street,
Kazan, 420111, Russian Federation*

Tatyana Tsoy
*Intelligent Robotics Department, Kazan Federal University, 35 Kremlyovskaya street,
Kazan, 420111, Russian Federation*

Edgar A. Martínez-García
*Department of Industrial Engineering and Manufacturing, Universidad Autónoma de Ciudad Juárez,
Manuel Díaz H. No. 518-B Zona Pronaf Condominio,
Juarez, 32315, Mexico*

Evgeni Magid
*Intelligent Robotics Department, Kazan Federal University, 35 Kremlyovskaya street,
Kazan, 420111, Russian Federation*
E-mail: abbyasov@it.kfu.ru, k.g.kononov@gmail.com, tt@it.kfu.ru, edmartin@uacj.mx, magid@it.kfu.ru
http://robot.kpfu.ru/eng

Abstract

New robotic solutions should be carefully verified before executing with real robots in real environments. Simulation provides a significant support in testing, but requires test sites with a high level of realism. In this case, 3D modeling can be used to produce the necessary 3D digital representation of real objects with varying difficulty. This article presents a step-by-step tutorial on modeling a realistic office environment. The environment contains a building skeleton frame, windows, building tiles, and furniture. Blender modeling toolset was used to create high-quality 3D models in the Gazebo simulator. The constructed virtual environment was validated with a lidar-based SLAM task for a UGV.

Keywords: 3D modeling, UGV, lidar-based SLAM, Gazebo.

1. Introduction

Real-world objects 3D modeling is the most commonly used technique for producing a 3D digital representation real 3D objects or surfaces[1]. In robotics field the modeling approach is often used for prototyping 3D robot models[2,3], designing test sites[4,5]. The 3D model is a replica of a real object and has actual physical and visual characteristics: size, scale, shape, texture. 3D modeling toolsets allow creating 3D objects, e.g., Blender[6].

However, to interact with a 3D model, it should first be loaded into a virtual environment.

Simulation is an essential part in robotics research field. The purpose of a simulation is to imitate real environment conditions (e.g., gravity). The benefits of using simulation: no special hardware is required; virtual experiments take less time than real experiments; quick algorithm errors detection in early implementation stages[7].

Developed novel concepts, approaches, and algorithms should be verified before executing with robots in real environments[8]. In robotics, 3D modeling provides a significant support in proving the correctness of robotics tasks performed. One of these tasks is simultaneous localization and mapping (SLAM)[9]. During a mapping procedure, a robot explores an environment and creates a map. 3D modeling allows creating necessary virtual environments with a high realism level: landscapes, terrains, buildings, building interiors and other objects with varying difficulty.

This paper presents a step-by-step tutorial on creating a realistic office virtual environment. The 14th floor of the second study building of Kazan Federal University (KFU) was used as a real office environment. Blender was chosen as a modeling software tool due to its intuitive interface, good documentation, and free license. For simulating and testing the modeled virtual environment we used Gazebo simulator. Gazebo is a popular well-designed free 3D robotics simulator that makes it possible to design and emulate complex virtual environments and operate robots within them. The constructed virtual environment is validated through a lidar-based SLAM task for a UGV performed using Robot Operating System (ROS). ROS provides necessary interfaces to simulate robots in Gazebo simulator.

2. Environment Modeling

3D modeling requires special 2D CAD drawings of a real model. A 2D drawing represents the top view or side view of an object. An artist uses 2D drawings to get a basic understanding of the model object and its technical characteristics (size, shape, etc.). We used a 2D plan of the 14th floor of KFU 2nd study building that is shown in Fig. 1.

We used the following characteristics of the floor plan:

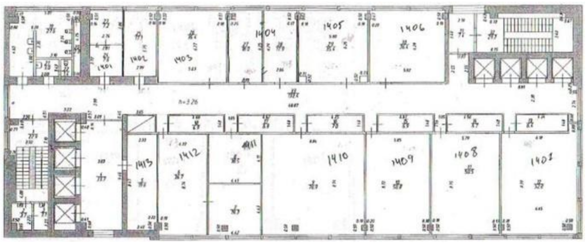


Fig. 1. Top view 2D drawing of the 14th floor of KFU 2nd study building.

- Size of the rooms: length, width, and height
- Corridor width and length
- Types of windows

Other details of the 14th-floor interior were omitted for simplifying the 3D model. We created only main model parts: a building skeleton frame, windows, building tiles, and furniture.

2.1. *Building skeleton frame*

Constructing a skeleton frame is done by creating a 14th-floor skeletal frame of pieces of parallelepipeds and cuboids of different sizes and thicknesses. A 3D model of the 14th-floor building frame is shown in Fig. 2. Dimensions of the model are 45 × 15 × 3 (m). The model is a set of building blocks: walls and room dividers.

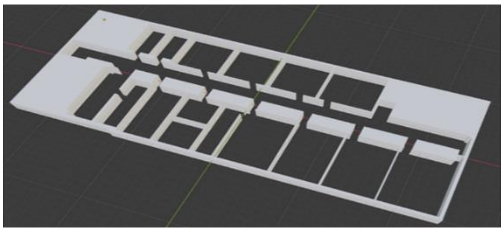


Fig. 2. Top view 2D drawing of the 14th floor of KFU 2nd study building.

2.2. *Windows*

We modeled two types of windows: transparent glass and tinted glass (Fig. 3). Tinted glass is used between the corridor and classrooms, transparent glass – for exterior windows.

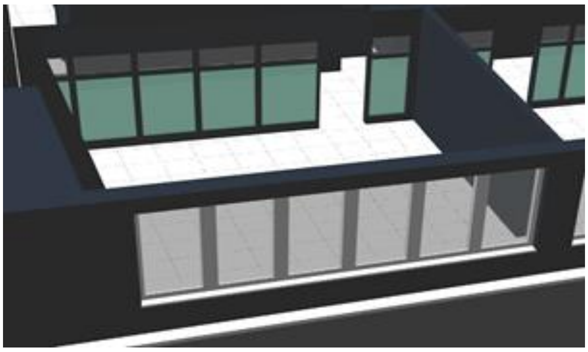


Fig. 3. Window types: transparent glass and tinted glass windows.

2.3. *Building tile*

A building tile is modeled as a parallelepiped (Fig. 4). Dimensions of the shape are 0.3 × 60 × 50 (cm). In order to optimize simulation, we joined several pieces of building tile into one continuous solid chunk. A tile has a material property that defines color or texture. The material can be easily changed to the necessary material.

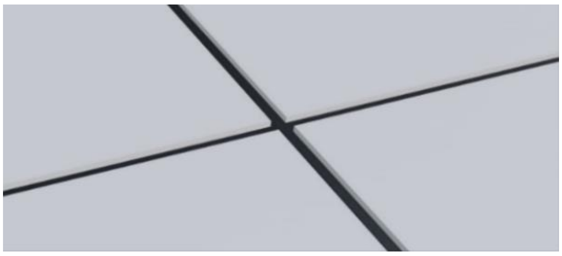


Fig. 4. 3D model of building tiles.

2.4. *Classroom furniture*

In order to design classroom interior, a chair, a wardrobe and tables were modeled. We modeled two types of tables: a rectangular activity table and a trapezoid activity table. They both have the same size: 80 × 55 × 175 (cm).

The standard chair has an overall height of 90 (cm), width of 40 (cm), and depth of 41 (cm). The average standard depth of a wardrobe is 80 (cm), height and length is 160 and 80 cm respectively. An example of virtual classrooms is shown in Fig. 5.

3. Use of the Real Office Environment

One of the most interesting application of virtual environments is testing and evaluation of Lidar-based 2D SLAM methods. We used *hector_slam* 2D SLAM ROS

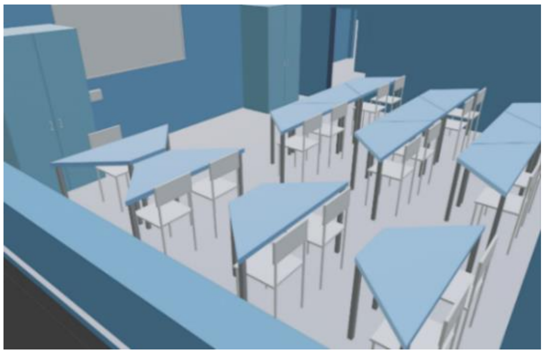


Fig. 5. Example of a virtual classroom.

algorithm[10], which employs LRF sensory data and odometry to build a 2D occupancy grid map.

The created environment usability was evaluated by running *hector_slam* with Servosila Engineer mobile crawler-type robot. In our virtual environment testing, *hector_slam* SLAM method produced the 2D occupancy grid map (Fig. 6, right) of the 14th floor of the KFU second building (Fig. 6, left).

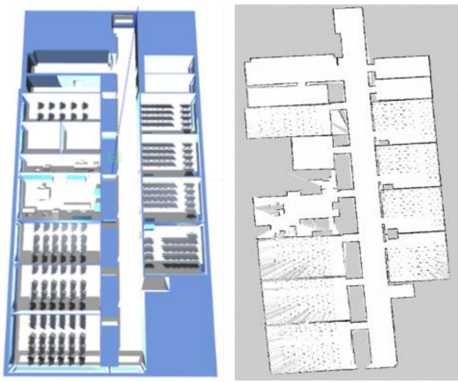


Fig. 6. The original 3D environment (left) and sensory.

4. Conclusions

This paper presented a tutorial on modeling a realistic office environment in Gazebo. The environment consists of a building skeleton frame, transparent glass and tinted glass windows, building tiles, and furniture. The constructed virtual environment was validated with *hector_slam* SLAM for Servosila Engineer UGV.

Acknowledgements

This paper has been supported by the Kazan Federal University Strategic Academic Leadership Program ("PRIORITY-2030").

References

1. T. Linner, A. Shrikathiresan, M. Vetrenko, B. Ellmann, and T. Bock, "Modeling and Operating Robotic Environment Using Gazebo/ROS", Chair for Building Realization and Robotics, 2011.
2. K. Shabalina, A. Sagitov, K.-L. Su, K.-H. Hsia, and E. Magid, "Avrora Unior Car-like Robot in Gazebo Environment", International Conference on Artificial Life and Robotics, pp. 116-119, 2019.
3. A. Dobrokvashina, R. Lavrenov, E. A. Martinez-Garcia, Y. Bai, "Improving Model of Crawler Robot Servosila "Engineer" for Simulation in ROS/Gazebo", Proceedings of the 13th International Conference on Developments in eSystems Engineering (DeSE), pp. 212-217, 2020.
4. B. Abbyasov, R. Lavrenov, A. Zakiev, K. Yakovlev, M. Svinin, E. Magid, "Automatic Tool for Gazebo World Construction: from a Grayscale Image to a 3D Solid Model", IEEE International Conference on Robotics and Automation (ICRA), pp. 7226-7232, 2020.
5. T. Wright, A. West, M. Licata, N. Hawes, B. Lennox, "Simulating Ionising Radiation in Gazebo for Robotic Nuclear Inspection Challenges", Robotics, Vol. 10, No. 3, pp. 1-27, 2021.
6. M. Zając, S. Paszkiel, "Using Brain-Computer Interface Technology for Modeling 3D Objects in Blender Software", Journal of Automation, Mobile Robotics and Intelligent Systems, pp. 18-24, 2021.
7. K. Takaya, T. Asai, V. Kroumov, and F. Smarandache, "Simulation environment for mobile robots testing using ROS and Gazebo", International Conference on System Theory, Control and Computing, pp. 96-101, 2016.
8. C. S. Timperley, A. Afzal, D. S. Katz, J. M. Hernandez, C. Le Goues, "Crashing Simulated Planes is Cheap: Can Simulation Detect Robotics Bugs Early?", 11th International Conference on Software Testing, Verification and Validation (ICST), pp. 331-342, 2018.
9. Y. Liu et al., "The Design of a Fully Autonomous Robot System for Urban Search and Rescue", International Conference on Information and Automation, pp. 1206-1211, 2016.
10. S. Kohlbrecher, J. Meyer, T. Graber, K. Petersen, U. Klingauf, O. Stryk, "Hector Open Source Modules for Autonomous Mapping and Navigation with Rescue Robots", Robot Soccer World Cup, pp. 624-631, 2013.

Authors Introduction

Mr. Bulat Abbyasov

He received a BSc degree in Computer Science at the Institute of Information Technology and Intelligent Systems (ITIS) of Kazan Federal University (KFU) in 2020. Currently, he is a second-year student of Master degree program in Intelligent Robotics at the ITIS, KFU

Mr. Kirill Kononov

He is a forth year undergraduate student at the Institute of Information Technology and Intelligent Systems (ITIS), Kazan Federal University

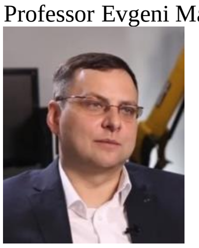
Ms. Tatyana Tso

In 2012 she graduated from International Area Studies Master's Degree Program at the University of Tsukuba. Since 2018 she has been a PhD student in Robotics at the Institute of Information Technology and Intelligent Systems (ITIS) of Kazan Federal University.

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



He is a Head of Intelligent Robotics Department, a founder, and a Head of Laboratory of Intelligent Robotic Systems (LIRS) at Kazan Federal University, Russia. Senior member of IEEE. He obtained Ph.D. degree in Robotics Engineering from the University of Tsukuba, Japan. His research interests include urban search and rescue robotics, mobile robotics, path planning, robotic teams, and human-robot interaction. He authors over 200 publications in English, Russian and Japanese languages.

Graphical User Interface Design for a UAV Teleoperation

Roman Lavrenov
Intelligent Robotics Department, Kazan Federal University, 35 Kremlyovskaya street, Kazan, 420111, Russian Federation

Ramil Safin
Intelligent Robotics Department, Kazan Federal University, 35 Kremlyovskaya street, Kazan, 420111, Russian Federation

Yang Bai
Information Science and Engineering Department, Ritsumeikan University, 1-1-1 Noji-higashi, Kusatsu, Shiga, 525-8577, Japan

Edgar A. Martínez-García
Department of Industrial Engineering and Manufacturing, Universidad Autónoma de Ciudad Juárez, Manuel Díaz H. No. 518-B Zona Pronaf Condominio, Juarez, 32315, Mexico

Roman Meshcheryakov
Cyber-physical Lab, V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, 65 Profsoyuznaya street, Moscow, 117997, Russian Federation
E-mail: lavrenov@it.kfu.ru, safin.ramil@it.kfu.ru, yangbai@fc.ritsumei.ac.jp, edmartin@uacj.mx, mrv@ipu.ru
http://robot.kpfu.ru/eng

Abstract

The number of drones being used around the world grows at a high speed. New drones' manufacturers are emerging and new drone designs are being developed. Most drones are controlled by remote control, while UAV joysticks and communication protocols are different. However, the Robotic Operating System (ROS) unifies the control process for drones. In this article, we present a universal graphical interface for controlling drones using ROS. The program is written in C++ and Qt Framework. It enables to control UAVs, receive and visualize the data from drones. Due to the use of ROS topics, this program can be applied to any drone integrated with ROS.

Keywords: Unmanned Aerial Vehicle, UAV, GUI, Teleoperation, Control.

1. Introduction

Unmanned aerial vehicles (UAVs) are used in a large number of applications. Drones are utilized in the entertainment industry to create 3D shapes and luminous objects[1]. They are also taking part in more serious military and dual-use missions. They find their use in in the tasks of terrain survey, mapping of the environment[2], static and dynamic objects detection in the environment[3].

Fig. 1. Graphical user interface design from Zul and Husry[6].

Drones are also employed for specific indoor applications. For example, UAVS are able to read RFID tags as a part of the autonomous inventory of warehouses and operate together with ground robots[4].

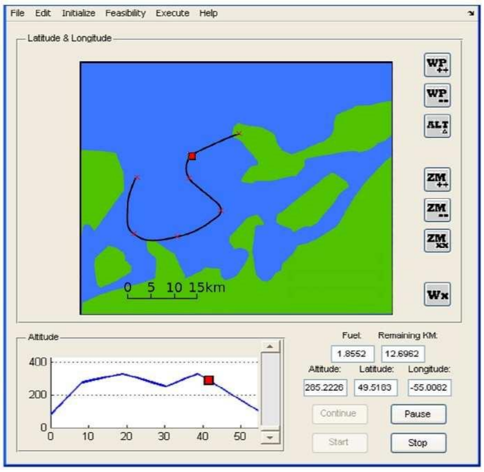


Fig. 2. Graphical user interface design from Perry and Taylor[7].

UAVs are able to operate alone or act in a group. A group of drones can have a robotic leader or not. The user can control robots one by one or the whole group, like a swarm[5].

Traditionally, drones are controlled using joysticks. A group of drones is controlled using specially developed programs. However, there is a unified framework for controlling robots – Robot operating system (ROS). In this paper, we present a new design for controlling drones using ROS.

2. Existing GUIs for Drones

Drones are usually controlled by an operator using a joystick or special software. Currently, several different interfaces exist for controlling drones. In some of them, one UAV is controlled, while in other cases there is a user that controls a group or swarm of drones. In this work, we surveyed the existing GUIs to find the most necessary widgets for UAV control.

Fig. 3. Graphical user interface design presented in the paper[8].

The GUI from Zul and Husry[6] contains altitude, position, and compass widgets (Fig. 1). In addition, the GUI contains data on the fuel level and the current battery charge. However, the graphical interface does not provide video stream from the cameras.

Another GUI for unmanned air vehicle mission planning and execution was presented by Sean R. Perry and James H. Taylor[7]. It allows tracking the drone over long distances and provides route planning (Fig. 2). The graph of change in height is displayed in the GUI. However, the interface does not provide real-time control of drones and does not track the status and speed.

Yet another GUI was presented in 2020 year in the Iraqi journal of computers, communications, control and systems engineering[8]. It uses a map from images from space, displays the speed and altitude of the drone. But there is no way to control the drone and there is no way to enter exact values in the GUI.

More options for controlling the drone were presented in the article "Natural user interfaces for multimodal human-drone interaction"[9]. According to the article, GUI is an auxiliary tool that allows one to perform high-level operations (takeoff / landing / flight in place). The interface enables to track the position of the drone relative to the coordinate axes. This GUI is programmatically associated with Aerostack[*]. The GUI cannot be used on every drone.

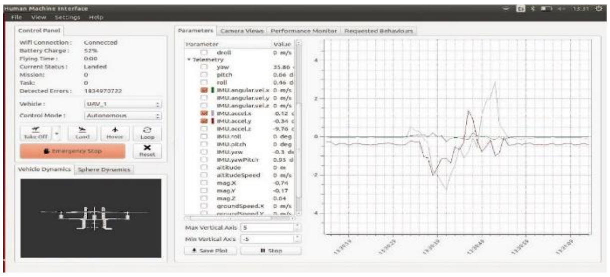


Fig. 4. Graphical user interface design presented in the paper[9].

3. Drones in ROS

The ROS framework has multiple drones that can be modeled and programmed. A lot of researchers use the hector_quadrotor UAV model to test drone control algorithms[10]. However, there are other drone simulation packages based on the PIXHawk and Mavlink controllers[4]. The new GUI should be able to control all drones that contain the ROS framework.

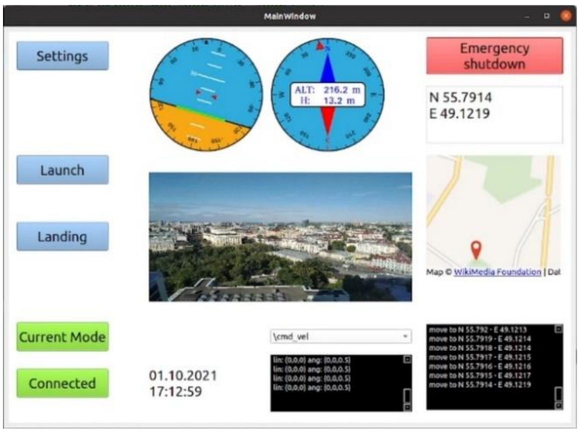


Fig. 5. New graphical user interface design.

4. Graphical User Interface Design

The graphical interface is written in C++ \ Qt Framework. It is designed to control a UAV using ROS. It includes the following widgets (Fig. 5):

- button “Settings”;
- button “Launch”;
- button “Landing”;
- button “Emergency shutdown”;
- info form “Current Mode”;
- info form “Status of connect”;
- info form “Date and time”;
- info form “Coordinates”;
- aerial widget (compass, height, roll, pitch indicators);
- image viewer (for video stream);
- map widget;
- combobox “topic select”;
- topic logs;
- other logs.

The map widget is created with QML, it renews UAV’s position in real-time. The coordinates of the UAV are displayed in a designated graphical interface’s form. Users are able to change the scale of the map using a mouse wheel.

The aerial widget includes several data. It shows compass, altitude, height, roll, pitch angles data. The source code of the widget can be found by link[†].

Users can select a ROS-topic in the combobox. Messages that are published in this ROS-topic will be displayed in the “topic logs” form as text.

Clicking the “Settings” button opens the window with UAV connection parameters. The drone control buttons are also configured in the "Settings" form.

The developed GUI is able to control all drones that are a part of the ROS framework.

Acknowledgements

This paper has been supported by the Kazan Federal University Strategic Academic Leadership Program ("PRIORITY-2030").

References

1. M. A. Azam, H. D. Mittelmann, S. Ragi, “UAV Formation Shape Control via Decentralized Markov Decision Processes”, Algorithms, Vol. 14, No 3, pp. 1-12, 2021.
2. J. Ma, D. Guo, Y. Bai, M. Svinin, E. Magid, “A Vision-Based Robust Adaptive Control for Caging a Flood Area via Multiple UAVs”, 18th International Conference on Ubiquitous Robots (UR), pp. 386-391, July, 2021.
3. B. Abbyasov, R. Lavrenov, A. Zakiev, T. Tsoy, E. Magid, M. Svinin, E.A. Martinez-Garcia, “Comparative Analysis of ROS-based Centralized Methods for Conducting

Collaborative Monocular Visual SLAM Using a Pair of UAVs", International Conference on Climbing and Walking Robots and Support Technologies for Mobile Machines (CLAWAR), pp. 113-120, August, 2020.
4. A. Khazetdinov, A. Aleksandrov, A. Zakiev, E. Magid, K.-H. Hsia, "RFID-based Warehouse Management System Prototyping Using a Heterogeneous Team of Robots", International Conference on Climbing and Walking Robots and Support Technologies for Mobile Machines (CLAWAR), pp. 263-270, August, 2020.
5. A. Zakiev, T. Tsoy, E. Magid, "Swarm Robotics: Remarks on Terminology and Classification", International conference on interactive collaborative robotics, pp. 291-300, 2018.
6. A. Zul Azfar, and D. Hazry, "Simple GUI Design for Monitoring of a Remotely Operated Quadrotor Unmanned Aerial Vehicle (UAV)", 7th International colloquium on signal processing and its applications, pp. 23-27, 2011.
7. S.R. Perry, J.H. Taylor, "A Prototype GUI for Unmanned Air Vehicle Mission Planning and Execution", IFAC Proceedings Volumes, Vol. 47, No. 3, pp. 12214-12219, 2014.
8. H.M. Qays, B.A. Jumaa, A.D. Salman, "Design and Implementation of Autonomous Quadcopter using SITL Simulator", Iraqi Journal of computers, communications, control and systems engineering, Vol. 20, pp. 1-15, 2020.
9. R.A.S. Fernandez et al., "Natural User Interfaces for Human-Drone Multi-Modal Interaction", 2016 International Conference on Unmanned Aircraft Systems (ICUAS), pp. 1013-1022, 2016.
10. J. Meyer et al., "Comprehensive Simulation of Quadrotor UAVs using ROS and Gazebo", International conference on simulation, modeling and programming for autonomous robots, pp. 400-411, 2012.

Authors Introduction

Assistant Professor Roman Lavrenov

Currently he works as an Assistant Professor at Kazan Federal University. He obtained his Ph.D. at Kazan Federal University in 2020. He teaches the following master's courses: «Robotic Operating System», «Industrial robotics» and others. Author of an online course and methodological guide on the Robotic Operating System (ROS). His main interests are path planning, robot operating system, and autonomous mobile robots.

Mr. Ramil Safin

Ramil Safin is a Ph.D. student and a research assistant at the Laboratory of Intelligent Robotic Systems, Institute of Information Technology and Intelligent Systems, Kazan Federal University, Russia. He teaches the following courses: «Fundamentals of technical vision», «Robotic systems sensors», «Algorithms and data structures» His main interest is computer vision for robotic applications.

Assistant Professor Yang Bai

Currently he works as an Assistant Professor at the College of Information Science and Engineering of the Ritsumeikan University. He teaches Information Systems Science and Engineering Course. His interests are robotics, non-holonomic systems and motion planning.

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 and leader of the academic body of Mechatronics 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, sensing and actuating models in all robots locomotion modalities.

Professor Roman Meshcheryakov

In 2014 he received his degree of Professor in the specialty "System analysis, management and information processing".
Since July 2018, he has been the Chief Researcher at the V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences and Head of the Laboratory of Cyberphysical Systems.

Numerical Solution Approach for the ROBOTIS OP2 Humanoid Hand Inverse Kinematics

Zagidullin Linar
Intelligent Robotics Department, Kazan Federal University, 35 Kremlyovskaya street, Kazan, 420111, Russian Federation

Tatyana Tsoy
Intelligent Robotics Department, Kazan Federal University, 35 Kremlyovskaya street, Kazan, 420111, Russian Federation

Roman Meshcheryakov
Cyber-physical Lab, V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, 65 Profsoyuznaya street, Moscow, 117997, Russian Federation

Kuo-Hsien Hsia
Department of Electrical Engineering, National Yunlin University of Science and Technology, 123 University Road, Section 3, Douliou, Yunlin, 64002, Taiwan, R.O.C

Evgeni Magid
Intelligent Robotics Department, Kazan Federal University, 35 Kremlyovskaya street, Kazan, 420111, Russian Federation
E-mail: lfzagidullin@stud.kpfu.ru, tt@it.kfu.ru, mrv@ipu.ru, khhsia@yuntech.edu.tw, magid@it.kfu.ru
http://robot.kpfu.ru/eng

Abstract

Small-size humanoids are widely used in human-robot interaction (HRI) projects and activities. To operate robot limbs in HRI and pick-and-place tasks it is required to solve an inverse kinematics problem. Classical approaches are closed-form solutions with algebraic or geometric approaches or a numerical solution. While a typical numerical solution is supposed to search for joint variables using an iterative optimization, in this paper we suggest an off-line solution for a ROBOTIS OP2 humanoid upper limb via a forward kinematics approach that allows to calculate in advance all possible solutions for an end effector pose within a robot workspace with several levels of the workspace discretization. The solution was obtained in a simulation and successfully validated with a real ROBOTIS OP2 humanoid.

Keywords: Humanoid, Inverse Kinematics, Control, Manipulator.

1. Introduction

Robots target to replace humans in dangerous locations or monotonous activities[1]. All activities that require object manipulation, e.g., manufacturing tasks[2], object grasping by a humanoid hand[3], door opening[4] etc., involve inverse kinematics calculations[5]. To solve inverse kinematics, algebraic, geometric or numerical methods are employed[6]. While industrial manipulators are initially designed to allow closed form solutions, design of many humanoid robots might make it difficult to find a consistent solution with algebraic or geometric approaches[7]. This is one of the reasons that makes numerical methods popular in humanoid robots[8].

This work demonstrates a simple brute-force approach for obtaining required values of joint angles for a given end effector (EE) pose of the ROBOTIS OP2 humanoid upper limb with 3 degrees of freedom (DoF). The approach employs kinematics analysis of the robot[9] and forward kinematics to generate data tables and matrices of angle values of a predefined precision in the off-line mode. The computational feasibility of the method was evaluated on a PC and the proper performance was confirmed with real ROBOTIS OP2 humanoid.

2. System Setup

Robotis OP2 (Fig. 1) is a humanoid robot developed by Korean company Robotis[10]. For today it is one of the most popular platform in research, education and competitions that involve applications in human-robot interaction, mobility, and manipulation[11].

It has 20 DoFs (6 per each leg, 3 per each arm and 2 in the neck), which are operated with 20 MX-28T servos. The robot is equipped with onboard Intel Atom Processor N2600 (dual core, 1.6 GHz) CPU, 4GB RAM and 32 GB SSD mSATA storage.

This work studies the operation of the 3DoF upper limb of the robot. Table 1 demonstrates the DH-table for the upper limb[2]. The DH-table generates the following forward kinematics equations:

$$\begin{cases} x = 60s_1s_2 + 60s_1 - 16c_1 - 129(c_1s_3 - c_3s_1s_2) \\ y = 60c_2 + 129c_2c_3 \\ z = -16s_1 - 16c_1 - 60c_1s_2 - 129(s_1s_3 + c_1c_3s_2) \end{cases} \quad (1)$$

where $c_i = \cos(\theta_i)$, $s_i = \sin(\theta_i)$ and $\theta_1, \theta_2, \theta_3$ are angles of the corresponding joints (Fig. 1). These lead to the following inverse kinematics equations:

$$\begin{cases} xc_1s_2 + ys_1s_2 + zc_2 - 16s_2 - 60 = 16s_3 + 129c_3 \\ xc_1c_2 + ys_1c_2 - zs_2 - 16c_2 = 0 \\ -xs_1 + yc_1 - 76 = -16c_3 + 129s_3 \end{cases} \quad (2)$$

Table 1. DH parameters for Robotis Darwin OP2 robot hand.

i	α_{i-1}	a_{i-1}	d_i	θ_i
1	0	0	0	θ_1
2	-90	16.0	16.0	$\theta_2 - 90$
3	90	60.0	0	θ_3

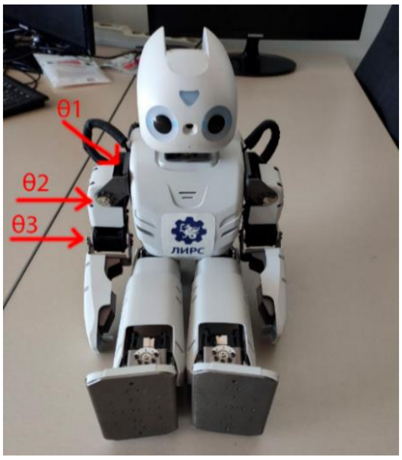


Fig. 1. Robotis Darwin OP2.

3. Brute Force Approach

The basic solution uses a prepared in advance text file, which is populated with data as follows:

(i) A value of an angle step discretization α is selected in advance by a user. We selected several values empirically, between 2 and 6 degrees. This value determines the step of each joint angle increment.

(ii) Based on the robot construction, for each joint angle minimal θ_i (*Min*) and maximal θ_i (*Max*) possible values are set, where *i=1,2,3*.

(iii) The three variables of joint angles $\theta_1, \theta_2, \theta_3$ are iteratively incremented with step α in three nested loops in order to provide all possible combinations within θ_i(*Min*) and maximal θ_i(*Max*), *i=1,2,3* joint variables. The resulting θ_i angles are used with forward kinematics equations (1) to provide *x,y,z* are coordinates of the EE, which are rounded to the closest real number with two digits after the decimal point. At this point, each line of the file contains $x, y, z, \theta_1, \theta_2, \theta_3$ values that correlate joint angles in the Joint space and an EE position in the Cartesian space. Note that the lines of the file are not ordered with regard to *x,y,z* coordinates and therefore a brute force search of the corresponding joint angles requires O(n) time, where n is the number of lines in the file.

4. Structured Brute Force Approach

This method uses the same brute force idea but structures the resulting data, which allows to greatly reduce the number of join angles calculations (actually, search in the file) on the robot at execution time. The data file is constructed in the off-line mode as follows:

(i) A matrix step value h is empirically selected in advance; it reflects the discretization level of the workspace and corresponds to the required precision of manipulation. An empty matrix M of reachable 3D Cartesian workspace is constructed with its dimensions corresponding to [X_{max}/h, Y_{max}/h, Z_{max}/h], where K_{max} are maximal possible coordinates of reachable workspace with regard to the appropriate axis K={X,Y,Z} of the global coordinate frame F_G, whose origin is placed in the center of the first joint between the robot body and the shoulder. All measurements are performed in this coordinate frame F_G.

(ii) Robot reachable workspace is divided into cubes with edges of length h. For each cube the brute force approach (Section 3) is used to calculate the joint angles' values that bring the EE into the cube's center.

(iii) Finally, the triples of $\theta_1, \theta_2, \theta_3$ angles are stored within a 3D matrix in the appropriate cell. The matrix stored as a file is uploaded to the robot.

Note that unreachable poses of the EE (that were not discovered at matrix construction time) did not produce joint variables and thus their cells within the matrix remain empty. The [0,0,0] cell corresponds to the origin of F_G that is used to descried the reachable workspace.

Next, at the execution time, to obtain the necessary joint angles, desired coordinates (X_{EE},Y_{EE},Z_{EE}) of the EE are transformed into matrix coordinates of (X_{EE}/h, Y_{EE}/h, Z_{EE}/h) with each value being rounded to the nearest integer index of the matrix.

5. Testing

A PC with 6C/12T Ryzen 5600H CPU, 16 GB DDR4 RAM and SSD NVME m.2 storage was employed for testing. To determine optimal tradeoff between the accuracy (the joint step value s) and the computation of pseudo inverse kinematics on the PC using brute force approach (Section 3) experiments were run with s=2 to s=6 degrees with a step of 0.5 degrees. Computation time was calculated as an average time to provide $\theta_1, \theta_2, \theta_3$ angles given (X_{EE},Y_{EE},Z_{EE}) coordinates.

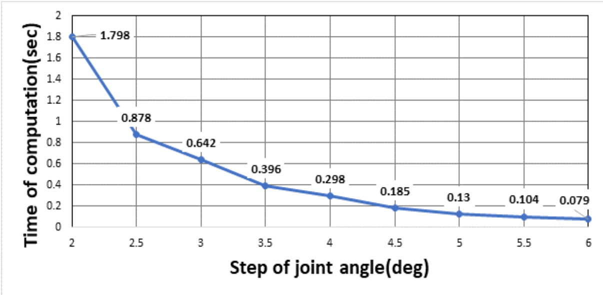


Fig. 2. Time consumption of pseudo inverse kinematics execution on the PC using brute force approach.

6. Conlcusions

The paper presents a simple off-line solution for a ROBOTIS OP2 humanoid upper limb via a forward kinematics approach that allows to calculate in advance all possible solutions for an end effector pose within a robot workspace with several levels of the workspace discretization. The solution was obtained in a simulation and successfully validated with a real ROBOTIS OP2 humanoid. The results allow determining an optimal tradeoff between the required accuracy (the joint angles' step) and the computation of joint angles using brute force approach.

Acknowledgements

This paper has been supported by the Kazan Federal University Strategic Academic Leadership Program ("PRIORITY-2030").

References

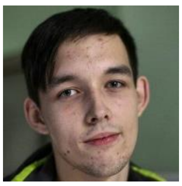
1. I. Ha, Y. Tamura, H. Asama, "Development of Open Platform Humanoid Robot DARwIn-OP", Advanced Robotics, Vol. 27, No. 3, pp. 223-232, 2013.
2. A. Kheddar, S. Caron, P. Gergondet, A. Comport, A. Tanguy, C. Ott, B. Henze, G. Mesesan, J. Englsberger, M. A. Roa, P.B. Wieber, "Humanoid Robots in Aircraft Manufacturing: The Airbus Use Cases", IEEE Robotics & Automation Magazine, Vol. 26, No. 4, pp. 30-45, 2019.
3. K. Khusnutdinov, A. Sagitov, A. Yakupov, R. Meshcheryakov, K.-H. Hsia, E. A. Martinez-Garcia, E.

Magid, “Development and Implementation of Grasp Algorithm for Humanoid Robot AR-601M”, 16th International Conference on Informatics in Control, Automation and Robotics, Vol. 2, pp. 379-386, 2019.

4. K. Otani, K. Bouyarmane, “Adaptive Whole-Body Manipulation in Human-to-Humanoid Multi-Contact Motion Retargeting”, IEEE-RAS 17th International Conference on Humanoid Robotics (Humanoids), pp. 446-453, 2017.
5. A. Dobrokvashina, R. Safin, Y. Bai, R. Lavrenov, “Improved Graphical User Interface for Crawler Robot Servosila Engineer”, International Siberian Conference on Control and Communications (SIBCON), pp. 1-5, 2021.
6. J.J. Craig, “Introduction to Robotics: Mechanisms and Controls”, Third Edition, Addison-Wesley, 2005.
7. M. A. Ali, H. A. Park, and C. S. G. Lee, “Closed-form Inverse Kinematic Joint Solution for Humanoid Robots”, 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 704-709, 2010.
8. M. Mistry, J. Nakanishi, G. Cheng, S. Schaal, “Inverse kinematics with floating base and constraints for full body humanoid robot control”, Proceedings of the International Conference on Humanoid Robots, pp. 22-27, 2008.
9. R.L. Williams II, “DARWIN-OP Humanoid Robot Kinematics”, International Design Engineering Technical Conferences & Computers and Information in Engineering Conference, Vol. 4: 36th Mechanisms and Robotics Conference, pp. 1187-1196, 2012.
10. “Robotis Company”, http://en.robotis.com/
11. T. Tsoy, L. Sabirova, R. Lavrenov, E. Magid, “Master Program Students Experiences in Robot Operating System Course”, 11th International Conference on Developments in eSystems Engineering (DeSE), pp. 186-191, 2018.

Authors Introduction

Mr. Linar Zagidullin



Undergraduate student of the fourth-year of the Laboratory of Intelligent Robotic Systems (LIRS), Institute of Information Technology and Intelligent Systems (ITIS), Kazan Federal University.

Ms. Tatyana Tsoy

She graduated from International Area Studies Master's Degree Program at Tsukuba University. Since 2018 she has been a Ph.D. student in Robotics at the Institute of Information Technology and Intelligent Systems (ITIS) of Kazan Federal University. She works as a research assistant at the Laboratory of Intelligent Robotic Systems (LIRS) at KFU.

Professor Roman Meshcheryakov

In 2014 he received his degree of Professor in the specialty "System analysis, management and information processing". Since July 2018, he has been the Chief Researcher at the V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences and Head of the Laboratory of Cyberphysical Systems.

Associate Professor Kuo-Hsien Hsia

Received a Ph.D. degree in Electrical Engineering at the National Sun Yat-Sen University, Kaohsiung, Taiwan. He works at the National Yunlin University of Science and Technology.

Professor Evgeni Magid

A Head of Intelligent Robotics Department and a Head of Laboratory of Intelligent Robotic Systems (LIRS) at Kazan Federal University, Russia. Senior IEEE member. He earned his Ph.D. degree from the University of Tsukuba, Japan. He authors over 200 publications in English, Russian and Japanese languages.

Alvus Modeling in Gazebo

Liaisan Safarova
Intelligent Robotics Department, Kazan Federal University, 35 Kremlyovskaya street, Kazan, 420111, Russian Federation

Bulat Abbyasov
Intelligent Robotics Department, Kazan Federal University, 35 Kremlyovskaya street, Kazan, 420111, Russian Federation

Tatyana Tsoy
Intelligent Robotics Department, Kazan Federal University, 35 Kremlyovskaya street, Kazan, 420111, Russian Federation

Hongbing Li
Department of Instrument Science and Engineering, Shanghai Jiao Tong University, 800 Dongchuan Rd., Shanghai, Minhang, 200240, China

Evgeni Magid
Intelligent Robotics Department, Kazan Federal University, 35 Kremlyovskaya street, Kazan, 420111, Russian Federation
E-mail: liaisansafarova@it.kfu.ru, abbyasov@it.kfu.ru, tt@it.kfu.ru, lihongbing@sjtu.edu.cn, magid@it.kfu.ru
http://robot.kpfu.ru/eng

Abstract

Insufficient testing of medical robots can lead to accidents during a surgery and damage an expensive equipment. A simulated 3D patient permits a preliminary checking of robotics-based medical scenarios without threatening a real patient's health. This article presents a 3D model of a human abdomen, which contains vital organs: intestine, liver, stomach, and kidneys. There are 3 layers of an abdominal wall: skin, adipose, and muscle. Blender modeling software was used to create realistic 3D models of organs with their distinctive features for Gazebo simulator. The model is presented as a ROS package with necessary configuration files and can be used by other researchers to simulate medical operations in Gazebo environment.

Keywords: 3D Modeling, Blender, Medical Robotics, Gazebo.

1. Introduction

Over the last two decades, robot-assisted procedures have become popular in many surgical scenarios[1]. Robots are used in various aspects of medicine – from surgical intervention and palpation to therapy and rehabilitation. This avoids the risk of wound infection, postoperative pain and reduces the need for blood transfusions[2]. But despite the many positive aspects, there are also disadvantages: the high cost of operations, limitation of the surgeon's movement and the absence of a three-dimensional image that interferes with coordination and reduces maneuverability[3]. Robots are used only for auxiliary tasks, e.g., Da Vinci, which is widely used for laparoscopic surgery, and ARTAS, which is used for hair transplant operations.

An increasing number of robots are being used in a wide variety of medical fields. These operations are not ubiquitous due to their high cost[4]. Surgical operations require special technical equipment and qualified personnel. An example of such an operation is laparoscopic surgery. Laparoscopic surgery is an operation on the internal organs that is performed through small holes, while traditional surgery requires large incisions.

Simulation avoids these difficulties because the medical environment, which includes the patient and special medical equipment, is emulated on a computer. This emulation approach is most effective when testing medical algorithms, since the simulation is close to real medical tests, which allows get reliable results of testing the developed medical algorithms. It is possible to work out all the outcomes and scenarios without compromising expensive equipment.

The ROS/Gazebo environment was chosen as the environment for simulating the medical environment, which includes the patient and medical equipment[5]. This choice is due to the fact that most of the real medical robotic systems are developed on the basis of the ROS framework[6]. This allows effortlessly transfer the simulation code to a real robot, thereby reducing the time required to adapt the design for use in real conditions.

Our analysis of scientific sources and publications showed that at the moment there is no full-fledged medical robotic complex in the ROS/Gazebo environment, and there are also no developments related to modeling the human body: head or body. There is no high-quality assembly of a complex model of the abdominal cavity in the Gazebo simulator, so it is impossible to find a world or sdf file, which will allow run ready-made modules or models in a complex.

This article presents a medical software package for the ROS/Gazebo environment, which allows to simulate a complex model of the abdominal. The package includes 7 3D models of organs and tissues. Each model of organs and tissues has unique structure: shape, color, location and quantity.

2. Complex Abdomen Model

The 3D model of abdomen is divided into two main parts: tissues (skin, muscles and adipose tissue) and organs (stomach, kidneys, liver and intestines).

2.1. *Tissues*

For the complex model tissues were modeled, differing in their shape, structure and color. Fig 1 shows the skin. The main features of our model are the beige color and the shape of the prelum. Human skin color is determined by the brown pigment melanin[7]. The prelum perform functions: twisting the lumbar spine, rotating the torso, stabilizing the movement of the chest and participating in the breathing process.

Fig. 2 shows adipose tissue models. For the reliability of the model, the relief was modeled, imitation of adipocytes. Adipocytes are cells of adipose tissue.

Muscle tissue has pronounced muscle fibers, colored red. The color is due to the pigment[8]. These features have been demonstrated in our model. In Fig. 3 depicts muscle tissue models.

Fig. 1. Human skin: real human abdomen (left) and simplified 3D model skin of a human with a waist outline and prelum (right). The left image is borrowed from ocalaplasticsurgery.com.

Fig. 2. Adipose tissue: exemplary adipose with adipocyte samples (left) and 3D adipose model with relief and pronounced adipocytes simulated in Gazebo (right). The left image source: science.org.

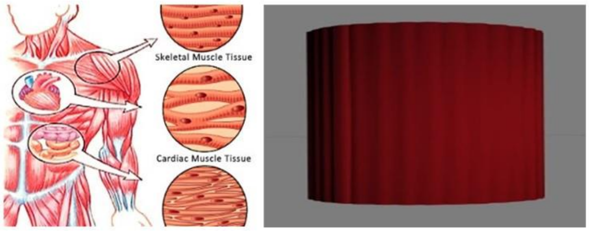


Fig. 3. Muscle tissue: standard muscle tissue 3D model with specified fiber direction (left) and created muscle tissue 3D model with pronounced fibers (right). The left image source: lifelinecelltech.com.

2.2. *Organs*

Fig. 4. Human intestine: original intestines 3D model location taking into account the location of other organs (left) and simplified intestine 3D model of small and large intestines (right). The left image is borrowed from webmd.com.

For the abdominal complex, 4 models of organs were modeled in the Blender modeling toolset, with their different peculiarities. Fig. 4 shows models of intestine. The intestine consists of the small intestine, the large intestine and rectum. The intestines are pink in color due to the very dense network of blood vessels[9].

Fig. 5 shows models of liver. The liver has 2 brown lobes. To create a more believable model, each lobe was modeled separately and fastened with connective tissue.

Fig. 6 shows models of stomach. A hollow muscular organ, part of the digestive tract. The model was created in red due to the network of blood vessels. The reliefs of the gastric mucosa with characteristic longitudinal folds, fields, dimples were modeled[10].

Fig. 7 shows models of kidneys. The kidneys are paired bean organs. This model demonstrates not only the kidneys organs, but also the adrenal glands, simulated with imitation of adipocytes, made in yellow. Arteries and ureter are also modeled.

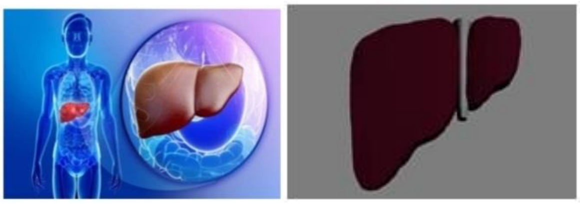


Fig. 5. Liver organ: sample liver 3D model with structure and location (left) and designed liver 3D model with two lobes (right). The left image is appropriate from webmd.com.

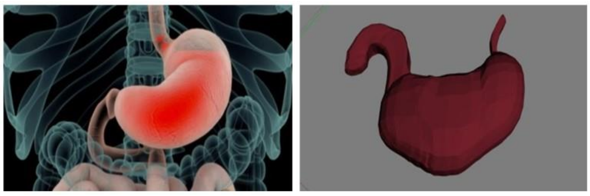


Fig. 6. Human stomach: exemplary stomach 3D model with location features (left) and created stomach 3D model taking into account the structure and shape (right). The left image is borrowed from webmd.com.

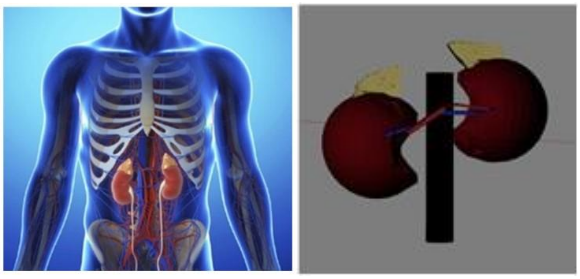


Fig. 7. Models of kidneys: a paradigmatic kidneys 3D model taking into account the structure and location (left) and simplified kidneys 3D model with adrenal glands (right). The image is imported from intermountainhealthcare.org.

3. Conclusions

This article presents a 3D model of a human abdomen, which contains the following organs: intestine, liver, stomach, and kidneys. There are 3 panniculus of an abdominal wall: skin, adipose, and muscle. The model is implemented as a ROS/Gazebo package with necessary configuration files and can be used to simulate medical operations in Gazebo environment.

Acknowledgements

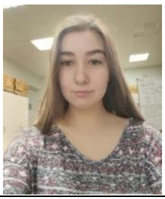
This paper has been supported by the Kazan Federal University Strategic Academic Leadership Program ("PRIORITY-2030").

References

1. R. Bogue, "The Rise of Surgical Robots", Industrial Robot: the international journal of robotics research and application, pp. 335-340, 2021.
2. D. Ilic, "Laparoscopic and Robotic-Assisted versus Open Radical Prostatectomy for the Treatment of Localised Prostate Cancer", Cochrane Database of Systematic Reviews, Vol. 9, No. 9, pp. 1-34, 2017.
3. M. Tonutti, "The Role of Technology in Minimally Invasive Surgery: State of the Art, Recent Developments and Future Directions", Postgraduate Medical Journal, Vol. 93, No. 1097, pp. 159-167, 2017.
4. E. Magid, A. Zakiev, T. Tsoy, R. Lavrenov, A. Rizvanov, "Automating pandemic mitigation", Advanced Robotics, pp. 1-18, 2021.
5. B. Abbyasov, A. Dobrokvashina, R. Lavrenov, E. Kharisova, T. Tsoy, L. Gavrilova, E. Magid, "Ultrasound Sensor Modeling in Gazebo Simulator for Diagnostics of Abdomen Pathologies", International Siberian Conference on Control and Communications, pp. 1-6, 2021.
6. C. Sachin, "ros_control: A Generic and Simple Control Framework for ROS", The Journal of Open Source Software, Vol. 2, No. 20, pp. 456-456, 2017.
7. M. Brenner, V. Hearing, "The Protective Role of Melanin against UV Damage in Human Skin", Photochemistry and photobiology, Vol. 84, No. 3, pp. 539-549, 2008.
8. A. Murray, "The Evaluation of Muscle Quality", Quality and grading of carcasses of meat animals, pp. 83-107, 2020.
9. M. Dewhirst, T. Secomb, "Transport of Drugs from Blood Vessels to Tumour Tissue", Nature Reviews Cancer, Vol. 17, No. 12, pp. 738-750, 2017.
10. W. Rickesha, C. Stevenson, "Anatomy and Physiology of the Stomach", Shackelford's Surgery of the Alimentary Tract, Vol. 2, pp. 634-646, 2019.

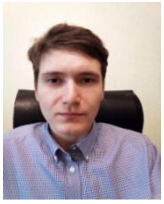
Authors Introduction

Ms. Liaisan Safarova



She is a third-year student of the Institute of Information Technology and Intelligent Systems at the Kazan Federal University (KFU). Currently she works as a research assistant at the Laboratory of Intelligent Robotic Systems (LIRS) at KFU.

Mr. Bulat Abbyasov



He received a BSc degree in Computer Science from the Institute of Information Technology and Intelligent Systems (ITIS) at the Kazan Federal University (KFU) in 2020. Currently, he is a second-year student of Master degree program in Intelligent Robotics at ITIS, KFU.

Ms. Tatyana Tsoy

She graduated from International Area Studies Master's Degree Program at Tsukuba University. Since 2018 she has been a Ph.D. student in Robotics at the Institute of Information Technology and Intelligent Systems (ITIS) of Kazan Federal University. She works as a research assistant at the Laboratory of Intelligent Robotic Systems (LIRS) at KFU.

Dr. 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 Evgeni Magid

A Head of Intelligent Robotics Department and a Head of Laboratory of Intelligent Robotic Systems (LIRS) at Kazan Federal University, Russia. Senior IEEE member. He earned his Ph.D. degree from the University of Tsukuba, Japan. He authors over 200 publications in English, Russian and Japanese languages.

Testing Procedures Architecture for Establishing a Fiducial Marker Recognition Quality in UAV-based Visual Marker Tracking Task in Gazebo Simulator

Mikhail Kilin
Intelligent Robotics Department, Kazan Federal University, 35 Kremlyovskaya street, Kazan, 420111, Russian Federation

Roman Lavrenov
Intelligent Robotics Department, Kazan Federal University, 35 Kremlyovskaya street, Kazan, 420111, Russian Federation

Yang Bai
Information Science and Engineering Department, Ritsumeikan University, 1-1-1 Noji-higashi, Kusatsu, Shiga, 525-8577, Japan

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, 35 Kremlyovskaya street, Kazan, 420111, Russian Federation
E-mail: MVKilin@stud.kpfu.ru, lavrenov@it.kpfu.ru, yangbai@fc.ritsumei.ac.jp, svinin@fc.ritsumei.ac.jp, magid@it.kpfu.ru
http://robot.kpfu.ru/eng

Abstract

Fiducial markers could be used in different tasks, including UAV and UGV marker-based localization. In most cases developers do not consider features of fiducial markers' systems (FMS) while selecting a particular FMS for a project. However, this selection might significantly influence results of experiments and thus the quality of a resulting product, an algorithm or a software. In this work, we define an architecture of experimental framework that allows finding an optimal marker for a UAV in a mobile ground object following task. The proposed framework estimates an average deviation of a detected Aruco marker position and an accuracy of the UAV landing on the marker. The framework uses Robot Operating System and employs UAV PX4 LIRS model in the Gazebo simulator.

Keywords: UAV, ROS, Gazebo, PX4, Fiducial Markers, ArUco.

1. Introduction

Fiducial markers are used for a wide range of tasks, including mapping, localization, and navigation in mobile robotics. For example, a robot can determine its current location using markers in a camera stream. Markers have different parameters, and thus it is important to use a marker that serves a particular function in a better way.

Several examples of the most popular markers are listed in Fig. 1. ArUco marker[1] (Fig. 1 (a)) is a square image containing a square binary matrix. A robot detects the marker using its black boundary and uses the internal matrix to identify the marker and obtain additional transformations' data. AprilTag marker[2] (Fig. 1 (b)) successfully overcomes issues with rotations and false positive detections, but its computational efficiency is lower relatively to the ArUco. ArTag marker[3] (Fig. 1 (c)) attempts to maximize the minimum Hamming distance between each instance of the family and has an improved search algorithm. Stag marker[4] (Fig. 1 (d)) employs a black square frame that helps to detect it, yet uses a circular internal pattern, which increases recognition stability of the algorithm.

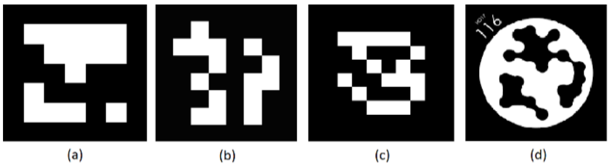


Fig. 1. Fiducial marker types: ArUco (a), AprilTag (b), ArTag (c), Stag (d).

Thoughtless choice of a marker can negatively affect performance results, therefore it is important to choose criteria for markers' comparison. For example, in[5] authors proposed to use marker resistance to rotations (with regard to different axes) as a metric and performed virtual experiments in Gazebo simulator in order to compare different families of markers.

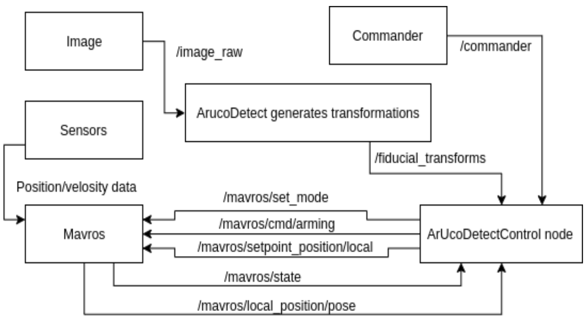


Fig. 2. UML diagram of the solution architecture.

In[6] the authors employed an overlap quality, a frequency of false detections, and a probability of inter-marker confusion as criteria for comparison. In some cases, it might be more reasonable to compare technologies within a specific application, e.g., authors in[7] compared fiducial markers in medical field using accuracy and noise as criteria since in such applications immunity is more important than computational costs.

This work presents an architecture of experimental framework that allows finding an optimal marker for a UAV in a mobile ground object following task. The framework estimates an accuracy of the UAV landing on the Aruco marker.

2. System Setup

We used Ubuntu 20.04 operating system with Noetic version of Robot Operating System (ROS)[8]. ROS is a set of tools, conventions, and libraries for robot software. ROS consists of two main parts: a system for interfacing software code and a set of packages with some standard robotics functions, e.g., planning, localization, getting data from sensors, etc. For virtual experiments Gazebo 11 simulator is employed. The Gazebo is an open-source 3D robotics simulator[9] that allows constructing robot models in SDF format, simulating various real-world conditions and sensors for tracking data about these conditions.

We use the PX4-LIRS model as UAV simulation[10]. This model contains four motors, where the opposing motors rotate in opposite directions. The model uses a camera with 800x800 resolution and 30 FPS, and a GPS sensor. The UAV employs PX4 open-source firmware[11] for UAV control, which provides a set of libraries to create custom solutions and scale them. Robotics APIs allow the PX4 to be controlled from outside the flight stack computing environment using a companion computer or other computing environment. The APIs interact with the PX4 using Micro Air Vehicle Link (MAVLink[12]), which is a protocol for communication

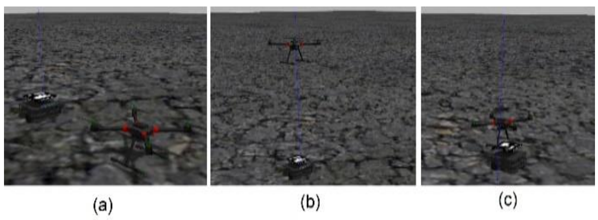


Fig. 3. UAV modes: initial pose (a), following (b), landing (c).

with UAVs. The MAVROS package provides MAVLink communication between computers running ROS.

3. Solution architecture

A scheme of the experimental framework architecture for selecting an optimal marker for a UAV in a mobile ground object following task is illustrated in Fig. 2. A heart of the scheme is an algorithm for UAV landing and following. We used *aruco_detect* node[13] from fiducials package for marker recognition. It contains a set of tools for searching and recognizing fiducials markers. After image processing, it sends data to */fiducial_transforms* topic. Next, the information about the recognized markers could be obtained from this topic.

Node *aruco_detect_control* subscribes to topics */commander*, */mavros/state*, */fiducial_transforms*, and */mavros/local_position/pose*. From */mavros/state* topic the node obtains data about the state of the UAV; from */mavros/local_position/pose* - a current position of the UAV using GPS; from */fiducial_transforms* - data about recognized markers from *aruco_detect* node.

Node *commander* receives data from the keyboard and sends it to *aruco_detect_control* node allowing to control the UAV flight. The initial position of the UAV at the simulation start is shown in Fig. 3 (a). The node starts sending data to *mavros/setpoint_position/local* topic with the starting position when *commander* topic receives data with "start" command. If *commander* topic receives "follow" command, the UAV starts following the marker (Fig. 3 (b)) and uses transformation data from *fiducial_transforms* topic. Fig. 4 presents a scheme of the following method of *aruco_detect_control* node. After "land" command, the UAV performs landing (Fig. 3 (c)). An altitude gradually decreases as long as the marker is in the field of view during the landing. The vertical landing is activated when the marker disappears from the field of view.

To compare markers, several metrics were selected[14]:

1) Ratio of time when the marker was recognized to time when it was not recognized
2) Average distance between the marker and UAV
3) Variance
4) Average speed of the UAV. Knowing the average speed of the target, we can estimate quality of communication.
5) Landing accuracy

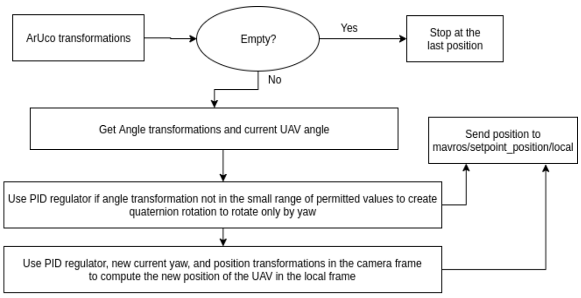


Fig. 4. UAV following UML.

To store these metrics, we created a special statistics node. It obtains data from all topics from the *aruco_detect_control* node, prepares them, and saves into a file.

To automate the marker movement, in our experiments we attached it to a top of the TurtleBot3[15] robot and recorded topics of its movement, including velocity topic and the commander topic.

Acknowledgements

This paper has been supported by the Kazan Federal University Strategic Academic Leadership Program ("PRIORITY-2030"). The third and the for the authors acknowledge the support of the Japan Science and Technology Agency, the JST Strategic International Collaborative Research Program, Project No. 18065977.

References

1. S. Garrido-Jurado, R. Mu˜noz-Salinas, F. J. Madrid-Cuevas, M. J.Mar´ın-Jim´enez, "Automatic Generation and Detection of Highly Reliable Fiducial Markers Under Occlusion", Pattern Recognition, Vol. 47, No. 6, pp. 2280-2292, 2014.
2. E. Olson, "AprilTag: A Robust and Flexible Visual Fiducial System", IEEE International Conference on Robotics and Automation, pp. 3400-3407, 2011.
3. M. Fiala, "ARTag, a Fiducial Marker System Using Digital Techniques", IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), Vol. 2, pp. 590-596, 2005.
4. B. Benligiray, C. Topal, C. Akinlar, "Stag: A Stable Fiducial Markersystem", Image and Vision Computing, Vol. 89, pp. 158-169, 2019.
5. T. Tsoy, A. Zakiev, K. Shabalina, R. Safin, E. Magid, & S. K. Saha, "Validation of Fiducial Marker Systems Performance with Rescue Robot Servosila Engineer Onboard Camera in Laboratory Environment", 12th

International Conference on Developments in eSystems Engineering (DeSE), pp. 495-499, 2019.
6. M. Fiala, "Comparing ARTag and ARToolkit Plus Fiducial Marker Systems", IEEE International Workshop on Haptic Audio-Visual Environments and their Applications, pp. 148-153, 2005.
7. C. Kunz, V. Genten, P. Meißner, B. Hein, "Metric-based Evaluation of Fiducial Markers for Medical Procedures", Medical Imaging Image-Gui|ded Procedures, Robotic Interventions, and Modeling, pp. 690-703, 2019.
8. M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, A. Y. Nget al., "ROS: an Open-Source Robot Operating System", ICRA workshop on open-source software, Vol. 3, No. 3, pp. 1-6, 2009.
9. N. Koenig, A. Howard, "Design and Use Paradigms for Gazebo, an Open-Source Multi-Robot Simulator", IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vol. 3, pp. 2149-2154, 2004.
10. A. Khazetdinov, A. Aleksandrov, A. Zakiev, E. Magid, K. H. Hsia, "RFID-Based Warehouse Management System Prototyping Using A Heterogeneous Team Of Robots", Robots in human life, pp. 263-270, 2020.
11. K. D. Nguyen, C. Ha, "Development of Hardware-in-the-Loop Simulation Based on Gazebo and Pixhawk for Unmanned Aerial Vehicles", International Journal of Aeronautical and Space Sciences, Vol. 19, No. 1, pp. 238-249, 2018.
12. A. Koubˆaa, A. Allouch, M. Alajlan, Y. Javed, A. Belghith, M. Khal-gui, "Micro Air Vehicle Link (MAVLink) in a Nutshell: A Survey", IEEE Access, Vol. 7, pp. 87658-87680, 2019.
13. A. Khazetdinov, A. Zakiev, T. Tsoy, M. Svinin and E. Magid, "Embedded ArUco: a Novel Approach for High Precision UAV Landing", 2021 International Siberian Conference on Control and Communications (SIBCON), pp. 1-6, 2021.
14. J. A. Ward, P. Lukowicz, H. W. Gellersen, "Performance Metrics for Activity Recognition", ACM Transactions on Intelligent Systems and Technology (TIST), Vol. 2(1), pp 1-23, 2011.
15. H. A. Ahmed, J. W. Jang, "Cloud Based Simultaneous Localization and Mapping with Turtlebot3", Proceedings of the Korean Institute of Information and Communication Sciences Conference, pp. 241-243, 2018.

Authors Introduction

Mr. Mikhail Kilin

He is a fourth year undergraduate student of the Kazan Federal University. He is interested in UAVs and fiducial markers.

Assistant Professor Roman Lavrenov

Currently he works as an Assistant Professor at Kazan Federal University. He obtained his Ph.D. at Kazan Federal University in 2020. His main interests are path planning, robot operating system, and autonomous mobile robots.

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. Mikhail is a member of the American Society of Mechanical Engineers (ASME), the International Conference on Intelligent Robots and Systems (RSJ), and the Society of Instrument and Control Engineers (SICE).

Professor Evgeni Magid

He is a Head of Intelligent Robotics Department, a founder, and a Head of Laboratory of Intelligent Robotic Systems (LIRS) at Kazan Federal University, Russia. Senior member of IEEE. He obtained Ph.D. degree from the University of Tsukuba, Japan. His research interests include urban search and rescue robotics, mobile robotics, path planning, robotic teams, and human-robot interaction. He authors over 200 publications in English, Russian and Japanese languages.

Feature Importance Evaluation Method for Multi-Agent Deep Reinforcement Learning in Advanced Robotics Task Allocation

Sergey Ryabtsev
Department of Computer Security, North-Caucasus Federal University, 1, Pushkin Street Stavropol, 355017, Russia

Mikhail Gurchinsky
Department of Computer Security, North-Caucasus Federal University, 1, Pushkin Street Stavropol, 355017, Russia

Igor Struchkov
Department of Organization and Technology of Information Security, North-Caucasus Federal University, 1, Pushkin Street, Stavropol, 355017, Russia

Vyacheslav Petrenko
Department of Organization and Technology of Information Security, North-Caucasus Federal University, 1, Pushkin Street, Stavropol, 355017, Russia

Fariza Tebueva
Department of Computer Security, North-Caucasus Federal University, 1, Pushkin Street Stavropol, 355017, Russia

Sergey Makarenko
Laboratory of Information Technologies in System Analysis and Modeling, Saint Petersburg Federal Research Center of the Russian Academy of Sciences, 39, 14 Line, Saint Petersburg, 199178, Russia
E-mail: nalfartorn@yandex.ru, gurcmikhail@yandex.ru, selentar@bk.ru, vipetrenko@ncfu.ru, ftebueva@ncfu.ru, mak-serg@yandex.ru
www.ncfu.ru

Abstract

The need to tackle intelligent tasks using advanced robotics multi-agent systems (MAS) actualize the use of artificial neural networks (ANNs) and multi-agent deep reinforcement learning technology. The article aims to solve the problem of exponential growth of ANN complexity with an increase in the number of agents in the MAS. To solve this problem, we propose an evaluation method for input data features importance. This method allows to optimize the input data feature set to reduce the computational complexity of the ANN inference while providing the same level of performance.

Keywords: Multi-agent systems, artificial neural networks, multi-agent deep reinforcement learning, feature importance

1. Introduction

The specific problems of learning ANN to control MAS agents have led to the emergence of such an area of machine learning as multi-agent deep reinforcement learning (MDRL)[1]. In contrast to single-agent deep reinforcement learning (SDRL), in the case of a MDRL, the ANN receives both features of the state of the environment and features of the state of other MAS agents:

$$F = \{X, F_1, .., F_N\}, \qquad (1)$$

where F – set of features supplied to the input of the ANN; X – set of features of the state of the environment; F_i – set of features of the state of the i-th agent, $i = \overline{1, N}$; N – number of MAS agents. Thus, in the transition from SDRL to MDRL, a component of variable dimension $\{F_i | i = \overline{1, N}\}$ appears in the feature set F, which depends on the number of MAS agents N. To overcome this limitation, in work[2] in relation to the problem of tasks allocation[3], it was proposed to input the ANN with a set of features F', which includes the features of the state of not all agents, but only some of them:

$$F' = \{X, F_1, .., F_{N'}\}, N' < N, F' \subset F, \qquad (2)$$

where N' – the number of agents whose features of the state are fed to the input of the ANN. This approach allows the use of MDRL for training MAS with a large number of agents, or MAS with a variable number of agents, for example, swarm robotic systems[4]. However, the problem of determining the optimal number of agents remains unsolved. To solve this problem, the in this article proposed a method for evaluation the importance of features in MDRL in problems of task allocation.

2. Related work

In work[5] proposed a combinatorial method for evaluation the importance of features based on combining features into subgroups, rearranging them, and retraining. The paper[6] raises the problem of combinatorial selection of the optimal composition of the used features. The authors proposed a method for evaluation the importance of features based on the use of the architecture of a double ANN – "operator" and "selector". The disadvantages of these methods are their high computational complexity. In work[7] proposed effective from the point of view of computational complexity, methods for evaluation the importance and choice of used features. However, these methods are focused on application in learning with a teacher.

The purpose of evaluation the importance of features in MDRL is to optimize their composition in order to achieve a compromise between computational complexity and the efficiency of MAS learning using some specific learning algorithm.

3. The proposed method for evaluation the importance of features

3.1. *Method description*

The proposed method based on the following hypothesis: in the case of a spatially distributed problem of task allocation features with low importance on average will make a small contribution to the value generated at the output of the ANN. Let us re-designate the grouped set of features F from Eq. (1) into an expanded set of features:

$$F = \{f_i | i = \overline{1, n}\}, \quad (3)$$

where n – the total number of features (environment and all agents).

Let us denote the MDRL process as a function l that returns some trained ANN π that approximates the behavior of MAS agents. The MDRL process is a function of the set of features F used:

$$l(F) \rightarrow \pi. \quad (4)$$

Let us also introduce the function $e(\pi)$ for evaluating the efficiency of the ANN π from the point of view of the purpose of the MAS:

$$e(\pi) \sim e. \quad (5)$$

The purpose of the estimates calculated using the proposed method is to determine such a subset of features $F' \subset F$, which can be excluded from the MDRL process l. In this case, an ANN π' should be obtained, the efficiency of which will be in a certain admissible neighborhood of the ANN efficiency π, trained on the full set of features F. Mathematically, this statement can be described as follows:

$$l(F \setminus F') \rightarrow \pi', \quad (6)$$

$$F \setminus F' = \{f_i \in F | f_i \notin F'\}, \quad (7)$$

$$|e(\pi') - e(\pi)| < \varepsilon, \quad (8)$$

where $F \setminus F'$ – the operation of subtracting the set F' from the set F, ε – given radius of the admissible neighborhood of the efficiency of the original ANN π.

To solve this problem the proposed method includes the following steps:

1. Training the ANN on the full set of features F using the training method l. As a result, ANN π is formed.

2. Calculation the estimate e of the efficiency of the ANN π.

3. Performing a test run of the ANN π at a certain number of steps. In this case, each state s_j supplied to the input of ANN π at the j-th step is recorded. As a result, a sample of data S is formed, supplied to the input of the ANN π:

$$S = \{s_j | j = \overline{1, m}\}, \quad (9)$$

where m – number of test run steps.

4. Further, for each feature $i = \overline{1, n}$, its elimination is emulated and the change in the ANN output π is evaluated based on the data of the test run S:

$$d_i = |\pi(S) - \pi(S \cdot M_i)|, i = \overline{1, n}, \quad (10)$$

$$M_i = \{m_j | j = \overline{1, n}\}, \quad (11)$$

$$m_j = \begin{cases} 0, \text{если } j = i, \\ 1, \text{если } j \neq i, \end{cases} \quad (12)$$

where M_i – the vector of masking of the i-th feature.

The obtained d_i values can then be used to evaluation the importance of the i-th feature.

Based on this method, the optimization of the set of used features F can be performed according to a scheme that includes the following steps.

1. The generated row d is processed by normalizing to the range $[0;\ 1]$ and ordering in descending order:

$$\bar{d} = \text{sort}\left\{\frac{d_i}{\max d_i - \min d_i} \middle| i = \overline{1, n}\right\}. \quad (13)$$

2. Some threshold value d_t is selected. If the value of the feature $\bar{d}_i$ is less than the threshold value d_t, this feature is included in the set of insignificant features F':

$$F' = \{f_i \in F | \bar{d}_i < d_t\}. \quad (14)$$

3. Based on the reduced set of features $(F \setminus F')$ the ANN π' is retrained and the efficiency is evaluation.

3.2. *Experimental results*

To test the efficiency of the proposed method, the simulated MAS was trained to solve the task allocation problem[2] using the modified MADDPG[2] method. The diagram of the series $\bar{d}$ obtained as a result of the application of the proposed method (Fig. 1). As follows from Fig. 1, starting with the features of the fourth agent's state, there is a sharp decrease in their importance.

To test the validity of this assumption, an optimization was carried out by brute force and training was carried out for a different number of agents (Fig. 2).

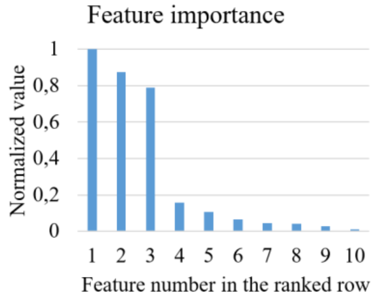


Fig. 1. Feature importance diagram.

As follows from Fig. 2 the effectiveness of functioning with a different number of agents more than two changes insignificantly. This confirms the adequacy of the results shown in Fig. 1. Similar efficiency is likely to be observed in other spatially distributed tasks, such as collective movement, or covering a certain area[8].

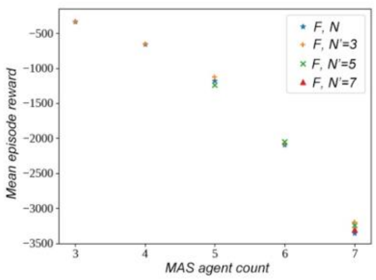


Fig. 2. Efficiency of trained ANN.

4. Conclusions

The article describes a proposed method for evaluation the importance of features. The essence of the proposed method is to evaluation the influence of the considered input feature on the change in the output signal generated at the ANN output. As follows from the results of experimental studies, the method can significantly reduce the number of used features without losing effectiveness. The advantage of the proposed method is its relatively low computational complexity.

References

1. P. Hernandez-Leal, B. Kartal, M.E. Taylor, "A survey and critique of multiagent deep reinforcement learning", Autonomous Agents and Multi-Agent Systems, Vol. 33(6), pp. 750-797, 2019.

2. V.I. Petrenko, F.B. Tebueva, A.S. Pavlov, M.M. Gurchinsky, "Multi-agent reinforcement learning method in systems with a variable number of agents", Information Technologies for Intelligent Decision-Making Support (ITIDS), 2021.
3. H. Hamann, "Swarm Robotics: A Formal Approach", Springer, Cham, 2018.
4. A. Zakiev, T. Tsoy, E. Magid, "Swarm Robotics: Remarks on Terminology and Classification", Lecture Notes in Computer Science, Vol. 11097, pp. 291-300, 2018.
5. C. Molnar, G. König, B, Bischl, G. Casalicchio, "Model-agnostic Feature Importance and Effects with Dependent Features - A Conditional Subgroup Approach", arXiv, 2021.
6. M. Wojtas, K. Chen, "Feature Importance Ranking for Deep Learning", arXiv, 2020.
7. G. Wei, J. Zhao, Y. Feng, A. He, J. Yu, "A novel hybrid feature selection method based on dynamic feature importance", Applied Soft Computing, Vol. 93(6):106337, 2020.
8. Y. Wang, M. Svinin, E. Magid, R. Sun, Y. Bai, "Adaptive Multi-Agent Coverage Control With Obstacle Avoidance", IEEE Control Systems Letters, Vol. 6, pp. 944-949, 2022.

Authors Introduction

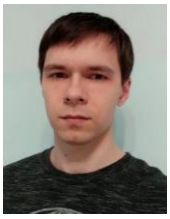
Mr. Sergey Ryabtsev

He received his Specialist's degree from the Department of Computer Security, North-Caucasus Federal University, Russia in 2016. He is currently a Master student in North-Caucasus Federal University, Russia.

Mr. Mikhail Gurchinsky

He received his Master's degree from the Department of Information Systems and Technologies, North-Caucasus Federal University, Russia in 2016. He is currently a postgraduate student in North-Caucasus Federal University, Russia.

Mr. Igor Struchkov



He received his Specialist's degree from the Department of Computer Security, North-Caucasus Federal University, Russia in 2019. He is currently a postgraduate student in North-Caucasus Federal University, Russia.

Mr. Vyacheslav Petrenko

He is Head of the Department of Organization and Technology of Information Security, North-Caucasus Federal University in Russia. He received his Ph.D. degree in Military control, communications and navigation systems in 1994. His research interests are Artificial intelligence.

Ms. Fariza Tebueva

She is Head of the Department of Computer Security, North-Caucasus Federal University in Russia. She received her Doctor of Physics and Mathematical Sciences degree in Mathematical modeling, numerical methods and program complexes in 2014. Her research interest is Swarm Robotics.

Mr. Sergey Makarenko

He is a Professor of Department of Information Security, Saint Petersburg Electrotechnical University "LETI" in Russia. He received his Doctor of Engineering Sciences degree in Systems, networks and telecommunication devices in 2018. His research interest is telecommunication networks.

Iterative Method of Labor Division for Multi-Robotic Systems

Sergey Ryabtsev
Department of Computer Security, North-Caucasus Federal University, 1, Pushkin Street Stavropol, 355017, Russia

Artur Sakolchik
Department of Mathematical Modeling and Data Analysis, Belarusian State University, 4 Nezalezhnosti Ave. Minsk, 220030, Belarus

Vladimir Antonov
Department of Computer Security, North-Caucasus Federal University, 1, Pushkin Street Stavropol, 355017, Russia

Vyacheslav Petrenko
Department of Organization and Technology of Information Security, North-Caucasus Federal University, 1, Pushkin Street, Stavropol, 355017, Russia

Fariza Tebueva
Department of Computer Security, North-Caucasus Federal University, 1, Pushkin Street Stavropol, 355017, Russia

Sergey Makarenko
Laboratory of Information Technologies in System Analysis and Modeling, Saint Petersburg Federal Research Center of the Russian Academy of Sciences, 39, 14 Line, Saint Petersburg, 199178, Russia
E-mail: nalfartorn@yandex.ru, gurcmikhail@yandex.ru, selentar@bk.ru, vipetrenko@ncfu.ru, ftebueva@ncfu.ru, mak-serg@yandex.ru
www.ncfu.ru

Abstract

Labor division in multi-robotic systems allows distributing tasks between agents in order to increase the efficiency of performing the global task. Collective decision-making methods allow agents to form the “agent-task” pairs. In this paper, we consider the case when the number of tasks significantly exceeds the number of agents. We propose an iterative method of labor division in multi-robotic systems. It uses collective decision-making to assign a cluster of subtasks to an agent. The paper examines different ratios between cluster size, number of clusters, and number of agents in order to find ratios that provide minimal average global task execution time and minimal average energy consumption.

Keywords: Multi-robotic systems, swarm robotics, task allocation, division of labor, collective decision making.

1. Introduction

Robotic technologies are being implemented into all spheres of human activity: when performing routine tasks, space exploration, performing work in response to natural and man-made emergencies, in the agricultural industry, in exploration, geology, in the fight against terrorism, and much more. At the same time, the intensive development of microelectronics has led to the miniaturization of robots. Thus, it became possible to use groups of numerous robots - multi-robotic systems (MRS). The advantages of their use are high mobility, low maintenance costs, the ability to perform many tasks, and the ability to scaling[1].

The problems of constructing an autonomous control system for MRS groups caused by the complexity of objectively existing systemic connections between the agents of the group, the patterns of interaction of the elements of the groups in an uncertain environment and possible instable elements. MRS groups are usually stochastic, nonlinear, so building mathematical models to test and optimize control models is difficult. The lack of methods for the transition from the specific behavior of the agent to the universal behavior of the group does not allow building an effective control system for groups of robots[2].

In this regard, a huge class of tasks appears for managing MRS groups. One of these tasks commonly known as the division of labor task.

Analysis of work[3] suggests a large variety of theoretical methods for solving this problem, especially with an equal number of agents and subtasks. In terms of popularity, one can single out heuristic algorithms, analytical, based on market economy models, potential fields, probabilistic approaches; methods based on machine learning and ANN, fuzzy logic, ant algorithms, dynamic and integer programming, genetic algorithms, mixed algorithms.

Many of the presented methods show good productivity with an equal number of agents and tasks, while in the literature the application of these methods is not widely described for cases when the number of tasks exceeds the number of agents by more than 5-20 times. In this case, an important criterion for performing tasks is the energy efficiency of the method of division of labor to perform the largest number of possible tasks.

2. The proposed method

The essence of the iterative method of division of labor in a group of multi-robotic systems when solving multiple problems is to carry out iterative procedures for establishing relationships of the "cluster-agent" type[4].

The algorithm for performing a global task using the iterative method of division of labor shown in Fig.1.

Decomposition and clustering of a global task into local tasks performed on the equipment of the control center. The input data for the decomposition is the data on the global task, the output data is the set of subtasks $\boldsymbol{Q}$ (2).

Clustering involves the formation of a set of clusters $W = \{w_1, w_2, \dots, w_{n_w}\}$ Clustering involves the formation of a set of clusters[5].

It assumed that in a three-dimensional environment, clustering performed by dividing a zone into cubes of the same size. The scale of the partition made experimentally and given in the results section.

To establish relationships of the "cluster-agent" type, the work considers three variations of the method of

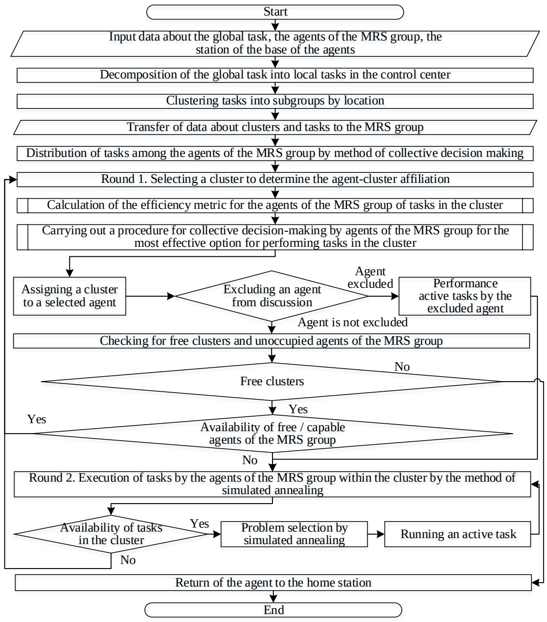


Fig. 1. Algorithm for performing a global task using the iterative method of division of labor.

distributing clusters between agents: the choice of near clusters (1M), the choice of distant clusters (2M) and the uniform choice of clusters (3M). The difference between the methods is in the order in which the clusters defined for the selection of agents. In the case of a variation of the 1M method, the clusters considered from the closest to the most distant from the place of launch of the MRS agents, in the variation of 2M, from the most distant to the nearest clusters. The 3M method selects several evenly distributed clusters in the target field equal to the number of agents. The choice of clusters starts from the farthest from the MRS launch point to the closest one. In the first round, only the clusters included in the list of selected clusters are considered.

Information about the clusters is transmitted to the agents of the MRS group, after which the agents proceed to the distribution of tasks (division of labor), which is carried out in two iterative rounds.

Round 1. Initially, all data about the set of subtasks $\boldsymbol{Q}$ and clusters $\boldsymbol{W}$.

Depending on the chosen method (selection of near clusters (1M), selection of distant clusters (2M) and uniform selection of clusters (3M)), a group of agents receives a certain cluster w_i,

Agents calculate performance metrics and the ability to perform tasks in the cluster. After calculating performance metrics, agents begin collective decision-making. Steps of Round 1 are repeated until there are free clusters and unoccupied agents.

Round 2. When assigning a cluster to an agent, the agent starts the 2nd round - performing tasks within the cluster. The sequence of tasks execution in the cluster determined by the simulated annealing method.

This procedure carried out until there are no free tasks or agents. If there are tasks and agents, the procedure repeated; if there are no tasks, but there are free agents, they sent to the base station. If there are no clusters at all, the agent starts moving to the home station.

3. Result

The software simulation carried out in the CoppeliaSim system. To assess the effectiveness of the proposed solutions for the distribution of tasks in the MPC group, 20'000 computational experiments carried out. The following methods implemented in the simulation:

- iterative task distribution algorithm, variation with distant clusters,
- iterative task distribution algorithm, variation with the nearest clusters,
- iterative task distribution algorithm, variation with distributed clusters,
- analog, greedy task distribution algorithm with collective decision-making.

In the simulation, 100 tasks generated by a uniform distribution, a group of MRSs of 5, 7, 10, and 15 agents performed tasks in clusters. For the purity of the experiment, the division into 18, 32 and 50 clusters checked. The results summarized for each generated map. A total of 250 cards were generated for one set of agents and clusters. The results of the study of the average distance traveled by all agents of the MRS group presented in Figures 3.

Based on the results of the study, the following conclusions drawn: with the average dimension of the clusters, the average distance traveled by the agent's decreases.

Because the distance traveled correlates with the expended energy in direct proportion, we can draw the following conclusions:

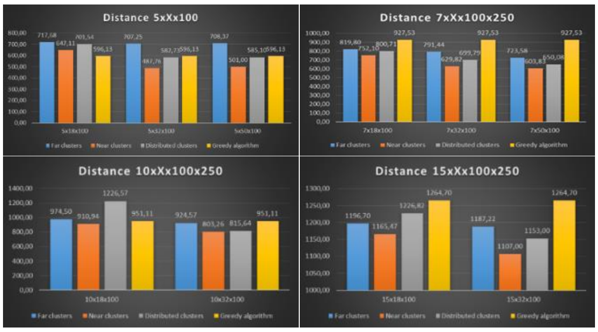


Fig. 2. Average distance traveled by 5, 7, 10, 15 agents in 18, 32, 50 clusters 100 tasks.

- for a group of 5, 7, 10, 15 agents with the ratio of tasks as to 20, to 14, to 10, to 7, respectively, it is optimal to use the iterative distribution method with variation of nearby clusters;
- depending on the size of the cluster, the algorithms show different results. For five agents, the best result shown by the distribution over 32 clusters. For 7 agents - 50 clusters, for 10 agents - 32 clusters, for 15 agents 32 clusters;
- for a different number of agents in comparison with the greedy algorithm, the results improved by 18% for 5 agents, by 35% for 7 agents, by 15% for 10 agents, by 12% for 15 agents.

The gain in energy efficiency in completing tasks offset by the time it takes to complete them. The best time indicator provided by the greedy algorithm.

4. Conclusions

This article proposes an iterative method for the division of labor in an MRS group in the case of an excess of the number of tasks over the number of agents by 5-20 times. The method is based on an iterative procedure for selecting task clusters and collective decision-making by agents of the MRS group. Three variations of the iterative method were proposed, with a difference in the order of selection of clusters for performing tasks by agents. The analogue of the method was a greedy algorithm for the division of labor with collective decision-making. According to the simulation results for a different number of agents, compared to the greedy algorithm, it was possible to achieve an improvement in the results by 18% for 5 agents, by 35% for 7 agents, by 15% for 10 agents, by 12% for 15 agents with a different number of task clusters.

References

1. A. Zakiev, T. Tsoy, E. Magid, “Swarm Robotics: Remarks on Terminology and Classification”, Lecture Notes in Computer Science, Vol. 11097, pp. 291-300, 2018.
2. E. Magid, R. Lavrenov, M. Svinin, A. Khasianov, “Combining Voronoi Graph and Spline-based Approaches for a Mobile Robot Path Planning”, Informatics in Control, Automation and Robotics, 14th International Conference on Informatics in Control, Automation and Robotics (ICINCO), pp. 475-496, 2017.
3. V.H. Pshihopov, V.V. Soloviev, A.E. Titov, V.I. Finaev, I.O. Shapovalov, “Group control of moving objects in uncertain environments”, Fizmatlit, 2015. (In Russian)
4. A.R. Gaiduk, I.A. Kalyaev, S.G. Kapustyan, A.P. Kukharenko “Models and methods of control of large groups of robots: a tutorial”, SFedU Publishing House, 2014. (In Russian)
5. F. Tebueva, V. Antonov, N. Svistunov, “Labour division algorithm for a group of unmanned aerial vehicles in a clustered target field”, E3S Web of Conferences, Vol. 270, No. 01035, pp. 1-9, 2021.

Authors Introduction

Mr. Sergey Ryabtsev

He received his Specialist’s degree from the Department of Computer Security, North-Caucasus Federal University, Russia in 2016. He is currently a Master student in North-Caucasus Federal University, Russia.

Mr. Artur Sakolchik

He is studying for a Bachelor's degree of the Department of Mathematical Modeling and Data Analysis, Belarusian State University in Belarus. His research interests are Swarm Robotics and Optimization methods.

Dr. Vladimir Antonov

He is an Associate Professor of the Department of Computer Security, North-Caucasus Federal University in Russia. He received his Ph.D. degree in Math. modeling, numerical methods and program complexes in 2019. His research interests are Swarm Robotics.

Dr. Vyacheslav Petrenko

He is Head of the Department of Organization and Technology of Information Security, North-Caucasus Federal University in Russia. He received his Ph.D. degree in Military control, communications and navigation systems in 1994. His research interests are Artificial intelligence.

Dr. Ms. Fariza Tebueva

She is Head of the Department of Computer Security, North-Caucasus Federal University in Russia. She received her Doctor of Physics and Math. Sciences degree in Mathematical modeling, numerical methods and program complexes in 2014. Her research interest is Swarm Robotics.

Prof. Sergey Makarenko

He is a Professor of Department of Information Security, St. Petersburg Electrotechnical University "LETI" in Russia. He received his Doctor of Engineering Science degree in Systems, networks and telecommunication devices in 2018. His research interest is telecommunication networks.

Development of Bowling Machine Using VEX IQ

Kuo-Hsien Hsia
Department of Electrical Engineering, National Yunlin University of Science and Technology,
123 University Road, Section 3,
Douliou, Yunlin 64002, Taiwan, R.O.C

Ya-Chun Chen
Graduate School of Engineering Science and Technology, National Yunlin University of Science and Technology,
123 University Road, Section 3,Douliou,
Yunlin 64002, Taiwan, R.O.C

Evgeni Magid
Intelligent Robotics Department, Kazan Federal University,
18 Kremlyovskaya Street,
Kazan 420008, Russian

Xin-Ying Zeng
Department of Electrical Engineering, National Yunlin University of Science and Technology,
123 University Road, Section 3,Douliou,
Yunlin 64002, Taiwan, R.O.C
E-mail: khhsia@yuntech.edu.tw, D10810210@yuntech.edu.tw,
magid@it.kfu.ru, a0954139931@gmail.com
www.yuntech.edu.tw

Abstract

VEX IQ is a kind of educational robotics platform focuses on semi-automatic and semi-remote. In this paper, a fully automatic bowling machine based on VEX IQ educational robotics platform has been developed. Since the brain of VEX IQ cannot communicate to other VEX IQ brains and the components of VEX IQ are made of plastic, it is necessary to overcome these problems to create mechanism similar to steel-made constructions and form an intelligent large system. We use sensors as communication interface for the brains of VEX IQ. Totally 3 brains and 10291 VEX IQ plastic components are used for the construction of the bowling machine. The overall size is about 252x93x90 in centimeters.

Keywords: Educational Robot, Bowling Game, VEX IQ, Robot C.

1. Introduction

Bowling is a well-known sport. The player tries to use the thrown ball to knock down the neatly arranged pins as many as possible. Usually there are three kind of indoor bowling games with 10 pins: tenpins, duckpins, and candlepins[1]. Fig.1 shows the relative sizes of bowling balls and pins for three popular variations of the bowling game[2]. Tenpins is the most popular one of the three variations. It is played everywhere except in the Canadian Maritimes.

It is quite challenging to make a bowling machine by using an educational robot host with plastic building blocks. Someone used LEGO to make a bowling machine with quite complete functions except scoring[3]. However, the pins in Ref. 2 are candlepins-like pins which are symmetrical up and down and have no distinct "top" or "bottom" end, unlike a tenpin. Unlike in tenpin bowling, fallen candlepins are not cleared away between balls within a player's turn [4].

VEX IQ is another kind of educational robotics platform which can be controlled in automatic and remote mode. Mr. Lin tried to use VEX IQ plastic building blocks to complete the implementation of an automatic bowling machine in his master thesis [5], but the functions in this work are independent of each other, and the system stability is insufficient.

Before making the VEX IQ bowling machine, the following issues must be considered:

1) The material of plastic parts is soft and low hardness.
2) VEX IQ host (brain) cannot communicate with each other.
3) The VEX IQ cable is limited in length.

In this paper, we will develop a bowling machine only with VEX IQ plastic building blocks and brains. The paper is organized as follows: Section 1 shows an introduction of this paper. Section 2 shows the hardware structure of the developed bowling machine. Section 3 shows the overall structure and some experiments. Finally, a conclusion and discussion is made in the section 4.

2. Hardware Structure

The VEX IQ bowling machine is composed of a pin-clamp mechanism, a pin-pressed and scoring mechanism, a host-communication mechanism, a pin-swiping mechanism, a pin sliding-track mechanism, a pin loader mechanism, a ball return mechanism, a rotary pin-elevator mechanism, and a pin conveyor-belt mechanism.

The current bowling machine arranges the pins in an equilateral triangle. However, due to the inherent limitations of VEX IQ building blocks, it is impossible to arrange an equilateral triangle. Therefore, we construct the pin-clamp mechanism in an isosceles triangle arrangement, as shown in Fig. 2. The function of the pin clamp mechanism is to clamp the unknocked pins and raise them, and then lower the pins to put the pins on the deck after the fallen pins are swept away.

Since the Vision Sensor of VEX V5 cannot fully match the VEX IQ to use, it cannot be used for scoring. So we used the touch sensor of VEX IQ to design a

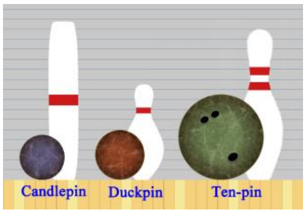


Fig. 1. Relative sizes of bowling balls and pins for three popular variations of the game[2].

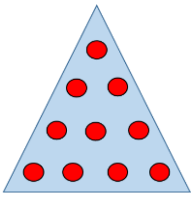



Fig. 2. The designed pin-clamp mechanism.

mechanism that can squeeze the pins. When the touch sensor is pressed, it will return a signal. So we can know the number of pins remaining on the deck, and the returned signal can be used for scoring. There is an ultrasonic sensor on the alley. When the ball is sensed to pass, the pin-pressed and scoring mechanism will be activated to perform the scoring process.

Since VEX IQ hosts (brains) cannot interconnect mutually and ultrasonic sensors will interfere with each other, two VEX motors with touch sensors and L-shaped beams are used to form a communication mechanism between the hosts. It is used for communication between host 1 (pin-pressed and sweeping) and host 2 (scoring).

The pin-swiping mechanism is used to sweep out the pins and the ball on the deck. It is made by combining the VEX IQ cross plates with the cross beams and the motor, and combining the rack and linear slider of the VEX IQ Gear Add-On Kit.

The pin sliding-track mechanism is used to make the pins transported by the pin conveyor belt slide down to the pin loader. Because the center of gravity of the pin is low, it is transported up to the highest point by the conveyor belt with the bottom facing up. The bottle sliding platform composed of the VEX IQ large plate has a slight downward slope to allow the bottle to slide downward due to the heavier bottom. The pin sliding track mechanism, composed of VEX IQ motor and rack, will move left and right to specify the position where the pin falls. A color sensor is installed above the end of the pin slide mechanism to detect whether a pin has slipped.

Whenever a pin slides down, the pin slide mechanism will move left and right to the top of the next designated drop position, waiting for the next pin to slide down.

The function of the pin loader mechanism is to place the pins dropped from the pin-sliding mechanism on the designated 10 positions, and move to the bottom of the pin-clamp mechanism to wait for the pin to be picked up after placing the pins. The pin-loader is composed of VEX IQ's horizontal plate, beam and thrust bearing housing. The isosceles triangle formed by the bearing housing on the front of the loader is just suitable for the arrangement of the pins. There are four gears on the back, two of which are unpowered, and the other two gears are combined with a motor as a power output to drive the loader to move horizontally back and forth with a fixed rack on the left and right sides. Fig. 3 shows the front and rear sides of the pin-loader mechanism.

The ball return mechanism consists of three parts: the guiding track, the ball-lifting mechanism, and the return slope. The thrown ball will be returned to ball rack via the three parts to wait for the next throw. The designed ball-lifting mechanism is similar to a windmill in order to actually lift the ball from low to high. Hence the ball can roll on the return slope to the ball rack naturally.

The most important design of the automatic bowling machine is how to make the 10 pins in the same direction during transportation, and the pins are placed in the same direction on the pin loader. Due to the special shape of the pins and the inconsistent falling directions of the pins, it is not easy to meet the previous requirements. Different from the common pin elevator mechanism, large plates and gears of VEX IQ, and the tank chain (reverse application) were used to create a mechanism for lifting pins in a rotating manner, as shown in Fig. 4. The operation of this mechanism can transport the pins in sequence and guide the pins to the same direction. The pins will slide down to the pin conveyor belt due to the wider and heavier side of the bottom, so the pin directions on the pin conveyor belt will be consistent, as shown in Fig. 5.

3. Overall Structure and Experiments

Considering the structural robustness of the entire structure, the VEX IQ bowling machine is designed with the support like steel structure. For example, the column is composed of a horizontal plate and a corner connector, so that the structural rigidity of the column can be improved and not easily bend, and it can be more stable during the operation of the machine.

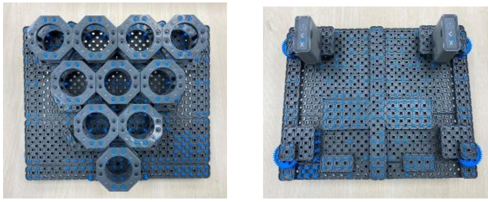


Fig. 3. Front and rear sides of the pin-loader mechanism.

Fig. 4. Pin elevator mechanism.

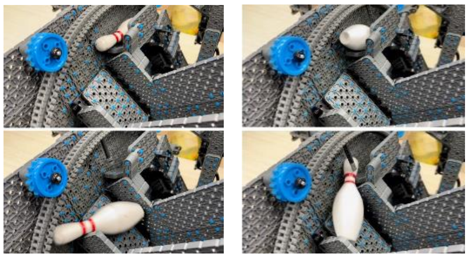


Fig.5. The pin directions will be consistent eventually.

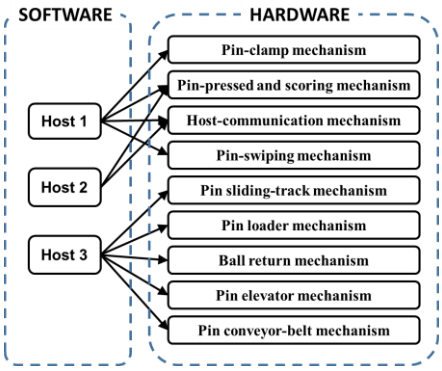


Fig. 6. Relationship of each host and each hardware block.

Three VEX IQ hosts are used in the designed bowling machine. The functions are respectively pressing the pin and sweeping the ball (host 1), scoring (host 2), placing and delivering the pins (host 3). The relationship between each host and the hardware block is shown in Fig. 6.

The completed bowling machine is shown in Fig. 7. Totally 10291 VEX IQ plastic components are used for the construction of the bowling machine. The overall size is about 252x93x90 in centimeters. Fig. 8 shows an actual operation of the bowling machine. In experiments, except for the pin-clamp mechanism, the operation success rate of the other mechanisms can almost reach 100%. The main reason for the failure of the pinning mechanism is that some pins are displaced due to the collision of the ball or other fallen pins, which will cause the mechanism to get stuck on the pins and fail to smoothly clamp. This will also cause subsequent scoring problems.

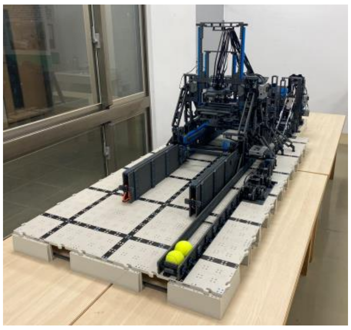


Fig. 7. The completed bowling machine.

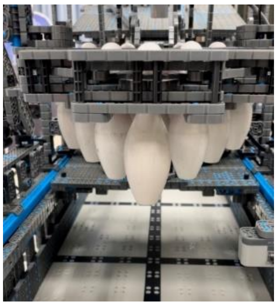


Fig. 8. The completed bowling machine.

4. Conclusions

In this paper, we introduce a bowling machine developed using only VEX IQ plastic building blocks. It is proved by actual operation that our organization can provide the entertainment of bowling game completely automatically. Since VEX's Vision Sensor does not yet support VEX IQ, the visual sensor cannot be used for scoring. The inability of direct communication between VEX IQ hosts also necessitates the use of additional sensing devices as a medium for message transmission. If these two problems are solved in the future, it will be expected that the entire system can be completed with fewer building blocks and parts.

Acknowledgements

The third author acknowledges the support of the Kazan Federal University Strategic Academic Leadership Program ("PRIORITY-2030").

References

1. M. Miller, “Bowling”, Bloomsbury Publishing, 2013.
2. “Bowling”, https:/en.wikipedia.org/wiki/Bowling
3. Lego Bowling Alley 5.0, YouTube, https://www.youtube.com/watch?v=8D8c4CcJobc
4. “Candlepin bowling”, https:/en.wikipedia.org/wiki/Candlepin_bowling
5. G.R. Lin, “Automatic Bowling Machine Implementation with VEX IQ and Teaching Material Development”, Master Thesis, National Yunlin University of Science and Technology, July 2020. (in Chinese)

Authors Introduction

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.

Ms. Ya-Chun Chen

She obtained her Master degree in Electrical Engineering from Kun Shan University, Taiwan in 2018. She is currently a Ph.D. student in the Institute of Engineering Science and Technology of National Yunlin University of Science and Technology, Taiwan. Her research interests are in the area of mobile robotics, electronic circuit and Layout.

Prof. Evgeni Magid

He is an acting Head of Intelligent Robotics Department, a founder, and a Head of Laboratory of Intelligent Robotic Systems (LIRS) at Kazan Federal University, Russia. Senior member of IEEE. Previously he worked at the University of Bristol and Bristol Robotics Laboratory, UK; Robotics Institute at Carnegie Mellon University, USA; the University of Tsukuba, Japan; National Institute of Advanced Industrial Science and Technology (AIST), Japan. His research interests include urban search and rescue robotics, mobile robotics, path planning, robotic teams, and human-robot interaction. He authors over 200 publications in English, Russian and Japanese languages.

Mr. Xin-Ying Zeng

He graduated from the Department of Electrical Engineering, National Yunlin University of Science and Technology, Taiwan, with a master's degree in 2021. He is currently a Mechanical Design Engineer of Gallant Precision Machining Co., Ltd. His work is mainly in the field of mechanical design drawing.

Design of local linear models using Self tuning Control System for PID Tuning According to error

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

PID control is widely used in process systems represented by chemical processes and petroleum refining processes. The reason is that PID control has a simple structure. However, most of the existing systems are non-linear systems, and it is difficult to always obtain good control results with fixed PID control. Therefore, in this study, we propose a method of tuning the PID gain according to the deviation (control error) of the control result, and verify the effectiveness of this method through experiments. For self-tuning PID control using a local linear model, we propose a program that performs PID tuning only when the deviation occurs with a certain magnitude. A simulation is performed on the Hammerstein model, which is a non-linear system. As a result of the experiment, the number of PID gain changes could be significantly reduced.

Keywords: Local Linear Models, PID controller, Control error.

1. Introduction

PID control [1,2] is widely used in industry, especially in process systems such as chemical processes and oil refining processes. PID control is widely used in industry, especially in process systems such as chemical and oil refining processes. However, the characteristics of many real systems vary according to environmental and operating conditions, and they include nonlinear systems. Therefore, it is difficult to obtain good control results with fixed PID control. For this reason, methods using machine learning and data-driven control [3,4] have been proposed as effective control methods for nonlinear systems. However, since these methods perform PID tuning at each step, the computational processing load is large, and only high-precision computers can handle them.

In this paper, we propose a method to tune the PID gain only when the control result does not follow the target value. In this method, a threshold value is set for the deviation, and PID control is performed only when the threshold value is exceeded. If the deviation is smaller than the threshold, the PID gain is not changed. Simulation is performed on a bilinear model, which is a nonlinear system, to verify the effectiveness of this method in reducing the computational load and focusing on the control results. A method of tuning the PID gain according to the deviation of the control results (control error) is proposed, and the effectiveness of this method is verified through numerical examples.

2. A design of a self-tuning control system using a local linear model that performs PID tuning according to deviation

The block diagram of the proposed control system is shown in Fig. 1. The authors have previously proposed a method for calculating control parameters using the idea of the local linear model method [5]. This method can be used to control a nonlinear system by creating a locally linear model. In the proposed control method, the deviation is evaluated for a self-tuning control system using a local linear model.

2.1. *System description*

First, consider the discrete-time nonlinear system represented by

$$y(t) = f(\varphi(t-1)) \tag{1}$$

where, y(t) represents the system output and f(·) represents the nonlinear function. Also, φ(t-1) represents the state of the system before time t-1 (historical data) and is called the information vector. The information vector φ(t-1) is defined by the following equation.

$$\varphi(t-1) := [y(t-1),\ y(t-2), \cdots, y(t-n_y)\quad u(t-1), u(t-2), \cdots, u(t-n_u)\,] \tag{2}$$

Furthermore, u(t) is the control input, and n_y and n_u are the orders of the output and input, respectively. Now, suppose that the nonlinear system represented by equation (1) can be locally represented by a linear model as follows

$$A_i(z^{-1})y(t) = z^{-(k_m+1)}B_i(z^{-1})u(t) \quad (i = 1,2,\cdots,N) \tag{3}$$

where, in Eq. (3), km represents the minimum estimate of the lag time, and when the lag time is known, km is set to that value; when the range of the lag time is unknown, k_m is set to 0. Furthermore, z^{-1} represents a time-delay operator, meaning $z^{-1}y(t) = y(t-1)$. Also, $A(z^{-1})$ and $B(z^{-1})$ are given by

$$A_i(z^{-1}) = 1 + a_{i,1}z^{-1} + \cdots + a_{i,n_y}z^{-n_y} \tag{4}$$

$$B_i(z^{-1}) = b_0 + b_{i,1}z^{-1} + \cdots + b_{i,n_u}z^{-n_u} \tag{5}$$

After the above preparation, the controller is designed for the local linear model.

2.2. *Controller Design*

Design the controller based on the following steps.

[STEP1] Construction of multiple linear models

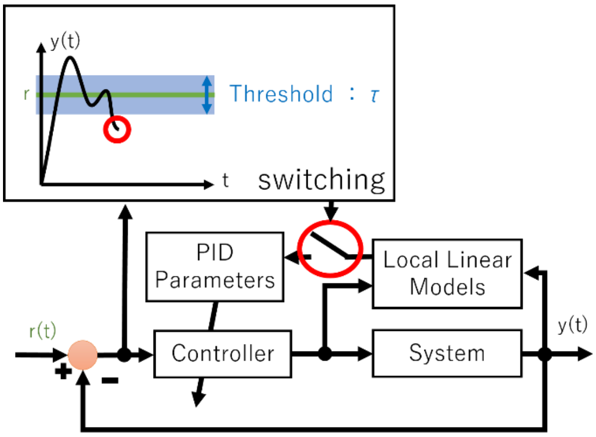


Fig. 1. Block diagram

For the nonlinear model, multiple linear models are constructed, system identification is performed using the lumped least squares method, and the parameters of $A_i(z-1)$ and $B_i(z-1)$ (i = 1, 2, - - - , N; where i takes these values unless otherwise noted) included in the linear model of Eq. (3) are estimated.) parameters in the linear model.

[STEP2] Design of control system

For the linear model represented by Eq. (3), consider the feedback control law given by Eq.

$$R(z^{-1})y(t) + S(z^{-1})\Delta u(t) - R(1)r(t) = 0 \tag{6}$$

where, r(t) represents the target value at step t. $R(z^{-1})$ and $S(z^{-1})$are polynomials designed based on the pole configuration of the closed-loop system, respectively.

$$\left.\begin{aligned} R_i(z^{-1}) &= r_{i,0} + r_{i,1}z^{-1} + \cdots + r_{i,n_1}z^{-n_1} \\ S_i(z^{-1}) &= 1 + s_{i,1}z^{-1} + \cdots + s_{i,n_2}z^{-n_2} \end{aligned}\right\} \tag{7}$$

$R(z^{-1})$ and $S(z^{-1})$ are designed based on the pole placement method. In this case, the input-output relation of the closed-loop system composed of equations (3) and (6) is expressed by the following Eq.

$$y_i(t) = \frac{z^{-(k_m+1)}B_i(z^{-1})R_i(1)}{P(z^{-1})}r(t) \tag{8}$$

The denominator polynomial $P(z^{-1})$ of Eq.(8) is defined by the following Eq.

$$P(z^{-1}) := \Delta A_i(z^{-1})S_i(z^{-1}) + z^{-(k_m+1)}B_i(z^{-1})R_i(z^{-1}) \tag{9}$$

It can be seen that $P(z^{-1})$ is the characteristic polynomial of the closed-loop system. The following equation is used to design this polynomial.

$$P(z^{-1}) = 1 + p_1 z^{-1} + p_2 z^{-2} \tag{10}$$

$$\left.\begin{aligned} P_1 &= -2e^{\frac{\rho}{2\mu}} \cos\left(\tfrac{\sqrt{4\mu-1}}{2\mu}\right) \\ P_2 &= e^{-\frac{\rho}{\mu}} \\ \rho &:= \tfrac{T_s}{\sigma} \\ \mu &:= 0.25(1-\delta) + 0.51\delta \end{aligned}\right\} \tag{11}$$

σ indicates a parameter corresponding to the rise time, and μ is a parameter related to the damping characteristics of the response, which is adjusted by δ. Here, $\mathrm{R}(z^{-1})$ and $\mathrm{S}(z^{-1})$ are calculated based on Eq.(9). In order to obtain $\mathrm{R}(z^{-1})$ and $S(z^{-1})$ uniquely, it is necessary to set their orders to $n_1 = n_y$ and $n_2 = n_u + k_m$, respectively. In this way, the pole placement control system can be designed for each linear model.

[STEP3] Replacement with PID controller

We have described a control method based on the pole placement method. This method can be replaced by a design method based on PID control if it is considered in the same way as in Eq.(6). First, consider the PID control law of the following Eq.

$$\Delta u(t) = K_I e(t) - K_P \Delta y(t) - K_D \Delta^{-2} y(t) \tag{12}$$

where, K_P , K_I , and K_D represent the PID gain, respectively. Furthermore, e(t) is the control deviation, which is given by the following Eq.

$$e(t) := r(t) - y(t) \tag{13}$$

Now, Eq.(6) is rewritten as follows

$$\frac{R_i(z^{-1})}{\nu} y(t) + \Delta u(t) - \frac{R_i(1)}{\nu} r(t) = 0 \tag{14}$$

In this case, from equations (12) and (14), the PID parameter is given by

$$\left.\begin{aligned} K_P &= \frac{-r_{i,1} - 2r_{i,2}}{\nu} \\ K_I &= \frac{r_{i,0} + r_{i,1} + r_{i,2}}{\nu} \\ K_D &= \frac{r_{i,2}}{\nu} \end{aligned}\right\} \tag{15}$$

The above allows us to adjust the PID parameters based on the approximate pole configuration.

$$\nu := 1 + s_{i,1} + s_{i,2} \tag{16}$$

[STEP4] Calculation of weights

Next, for each local linear data calculated in [STEP2], the estimation error $\varepsilon_i(t)$ is calculated for each model, and the weight ω_i is calculated based on this. $\varepsilon_i(t)$ is the error between the system output value y(t) and the estimated output value ŷ(t) of each linear model. Here, ŷ(t) is calculated based on equation (3) by the following formula

$$\hat{y}_i(t) = -A_i(z^{-1}) y(t) + z^{-(k_m+1)} B_i(z^{-1}) u(t) \tag{17}$$

where, $\mathrm{A}_i(z^{-1})$ and $\mathrm{B}_i(z^{-1})$ are the system parameters of each linear model estimated in [STEP1].

$$\varepsilon(t) = \left| y(t) - \hat{y}_i(t) \right| \tag{18}$$

$$\omega_i(t) = \frac{\frac{1}{\varepsilon_i(t)}}{\sum_{i=1}^{N} \frac{1}{\varepsilon_i(t)}} \tag{19}$$

In addition, $\omega_i(t)$ is the weight corresponding to the selected i-th information vector. The smaller the difference between the actual outputs value of the system and each linear model, the larger the value of this weight. Note that the calculation of $\omega_i(t)$ based on equation (18) satisfies the following Eq.

$$\sum_{i=1}^{N} \omega_i(t) = 1 \tag{20}$$

[STEP5] Generation of weighted PID parameters

Using the weights obtained in [STEP4] and the PID parameters in Eq.(15), calculate the weighted PID parameters using the following Eq.

$$\mathbf{K}(N) = \sum_{i=1}^{N} \omega_i(t) \mathbf{K}(i) \tag{21}$$

$$\mathbf{K}(i) := \left[K_P(i), K_I(i), K_D(i) \right] \tag{22}$$

3. PID tuning method in response to deviation

In contrast to the local linear model-based competitive tuning PID control described in Section 2.2, this section describes a method that performs PID tuning only when a deviation of a certain magnitude occurs. First, we define the threshold as τ, and the condition on τ as follows.

$$\tau > |e(t)| \tag{23}$$

The PID control described in Section 2.2 is performed only when this condition is satisfied; otherwise, the control is performed without changing the PID gain. Here, the parameter τ in equation (23) represents a design parameter given by a certain positive constant. The setting of this parameter requires some trial and error.

4. Simulation

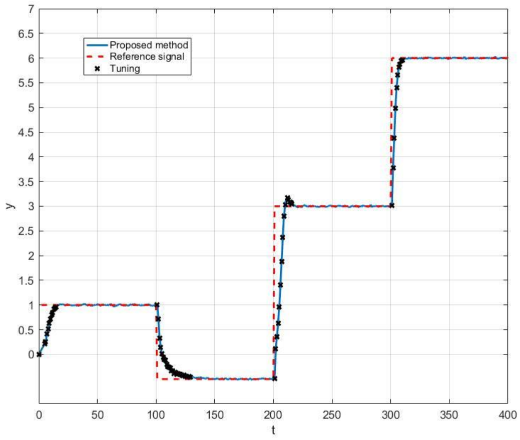


Fig. 1 Control result using the newly proposed PID control scheme

In order to verify the effectiveness of this method, a numerical example for a nonlinear system is presented. The Bilinear model is used as the control target. The Bilinear model, which is expressed by the following equation, is considered as the control target.

$$\begin{aligned} y(t) = 0.4y(t-1) - 0.99y(t-2) + 0.3u(t-1) \\ -0.1u(t-2) + 0.1y(t-1)u(t-1) \\ +0.05y(t-2)u(t-2) + \xi(t) \end{aligned} \tag{24}$$

However, ζ(t) is a Gaussian white noise with mean 0 and variance 0.012. The target value is given as follows.

$$r(t) = \begin{cases} 1.0(0 \le t < 100) \\ -1.0(100 \le t < 200) \\ 3(200 \le t < 300) \\ 6(300 \le t \le 400) \end{cases} \tag{25}$$

Next, based on the static characteristics, a linear model corresponding to the control input range is constructed as follows. However, the number of divisions is set to N=3. The various design parameters included in this method are $n_y = 2$, $n_u = 1$, and $k_m = 0$.

$$\begin{cases} -4.0 \le u_1 < 2.0 \\ 2.0 \le u_2 < 3.4 \\ 3.4 \le u_3 < 4.0 \end{cases} \tag{26}$$

Furthermore, the threshold τ for determining PID tuning is set to 0.1.

Fig. 2 shows the control results of the proposed method. The points shown in Fig. 2 indicate where the PID gain is being tuned. The results show that the number of times the PID gain is changed in the proposed method is only 56 step compared to 400 step in the conventional method because the conventional method tunes the PID gain sequentially, which is a significant reduction of about 1/7 times compared to the conventional method.

5. Conclusions

In this paper, a self-tuning control system using a locally linear model that performs PID tuning according to deviations for a nonlinear system is discussed. The effectiveness of the proposed method was verified by simulating a nonlinear bilinear model. In the future, we plan to evaluate the control performance of the deviation threshold setting using item response theory.

References

1. K.Heong Ang, G. Chong, Y. Li,"PID control system analysis, design, and technology", IEEE Transactions on Control Systems Technology, Vol.13, Issue.4, 2005
2. C. Gao, F. Wang, Z. Fu,"Fuzzy self-tuning PID control algorithm for belt conveyor driven by multi-motor", Journal of Robotics, Networking and Artificial Life, pp.175-178, 2017
3. T.Yamamoto, K.Takao and T.Yamada: Design of a Data-Driven PID Controller, IEEE Trans. on Control Systems Technology, Vol.17, No.1, pp.29-39, 2009
4. T. Kinoshita, T. Yamamoto,"Design of a Data-Driven Controller with Evaluating Controller Performance", Journal of Robotics, Networking and Artificial Life, pp.257-260,2019
5. S. Imai and T. Yamamoto, "Design of a Self-Tuning Controller using Multiple Local Linear Models for Nonlinear Systems", The transactions of the Institute of Electrical Engineers of Japan. C, A publication of Electronics, Vol.132,No.6, pp.860-865, 2012 (In Japanese)

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.

A Systematic Analysis of the Knee Support Exoskeleton Based on Multibody Dynamics Toward Personalization with 3D printed Spring-Damper Components

Shintaro Kasai[1], Pancho Dachkinov[1], Kohei Tanaka[1], Hiroaki Wagatsuma[1,2]
[1]Kyushu Institute of Technology, 4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan
[2]RIKEN CBS, Japan
E-mail: kasai.shintaro660@mail.kyutech.jp, dachkinov.pancho-nikolaev608@mail.kyutech.jp
tanaka.kohei886@mail.kyutech.jp, waga@brain.kyutech.ac.jp
https://www.kyutech.ac.jp/

Abstract

Exoskeleton-type assistive devices have been studied a long time focusing on the universal design and a simplification for mass production, and recently the concept is extended to the personalization according to the advancement of 3D printing, which allows to embed spring-damper systems in the form of compliant mechanisms. Therefore, a sophisticated kinematic and kinetic analysis is highly important for the realization of integrative systems and theories of multibody dynamics enhance the capability to find best parameters that are suitable for target body requirements. We analyzed a knee support exoskeleton in the form of the linkage system as the rigid-body dynamics and estimated necessary spring-damper components in the system to reduce burden on joint motions, especially persons with joint dysfunctions.

Keywords: Knee support, exoskeleton-type assistive device, multibody dynamics (MBD), joint dysfunction.

1. Introduction

The knee joint is a vulnerable and important joint for walking and then knee injuries have a serious impact on normal living ability of patients [1]. Therefore, understanding of the biomechanics of a normal and damaged knee joint is crucial for providing an actual support and it is beneficial for designing knee assistive devices with an appropriate optimization of parameters of the device to fit for successive rehabilitation program. Traditionally, exoskeleton-type assistive devices have been provided for the patients with motor disorders especially after neurological injuries and those devices have made rapid strides in recent years by integrating motorized parts and spring-damper mechanisms [2][3]. For example, assistive strategies were modeled with a rotational actuator, a simple pendulum model, and a damped pendulum model, which enhance abilities for normal and fast gaits [2][4]. Indeed, those require a dynamic adaptation to motion kinetics depending on the walking environment and model-based analyses highly important for evaluating of the joint torque and knee stiffness [4][5], which extend capabilities of exoskeleton-type assistive devices [6][7][8].

In this study, we introduced a systematic mathematical analysis based on the multibody dynamics [9][10][11][12][13] for motion kinetics occurring the knee joint support. In the realization of the systematic analysis, necessary joint torque and knee stiffness can be estimated clearly and it helps to design supportive mechanisms to provide a load reduction of the joint in the rehabilitation stage. The model-based kinetic and dynamic analysis were designed for detail investigations on exact timing and position for an effective support and it allows to know necessary improvements of the support devices.

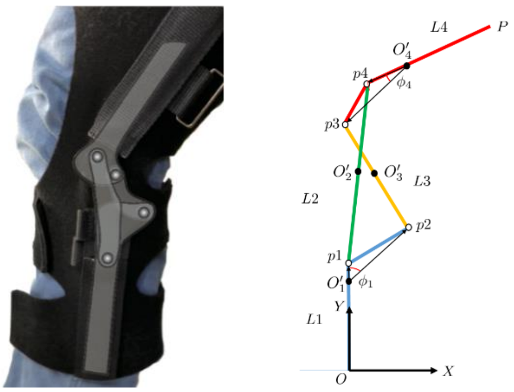
2. Methodology

2.1. Knee link model

The expert orthosis design provided by welfare device companies suggest that the human knee joint cannot be simply replaced by the single rotational joint, but it can be modeled as a combination of rotational joints and translational movements [14] as shown in Fig. 1 (a). In consideration of the knee analysis model, it is possible to analyze the knee movement and trajectory by using a simplified linkage mechanism as illustrated in Fig. 1 (b).

Table 1: Link model specifications

Parameter	**Sides**	**Length [mm]**
l_{10}	$O_1' \, O$	40
l_{11}	$O_1' \, p1$	1
l_{12}	$O_1' \, p2$	47
l_{21}, l_{22}	$O_2' \, p1, O_2' \, p4$	50
l_{31}, l_{32}	$O_3' \, p2, O_3' \, p3$	33
l_{40}	$O_4' \, P$	60
l_{41}	$O_4' \, p3$	100
l_{42}	$O_4' \, p4$	60



(a) (b)

Fig. 1. Rigid link model which reproduces knee mechanism and movement (a) and a simplified linkage mechanism in the rigid orthosis.

2.2. *MBD for knee link model*

In order to obtain the motion of each link and the change in the angle of each joint when a constant angular velocity is given to the joint p3 (Fig. 1 (b)) as the knee joint extension, multibody dynamics (MBD) [9][10][11] [12][13] was introduced. MBD is capable of numerical analysis that handles the motion and force of the multisystem simultaneously by describing the motion of a mechanism or structural system composed of various parts. A set of differential equations in the matrix represents constraints, kinematics and kinetics of the system and the numerical integration in the computer experiment provides the actual solution. Thus, in the MBD analysis, the differential algebraic equation is derived from the generalized coordinates. In forward dynamics analysis by MBD, a differential algebraic equation as Eq. (1) is necessary for the formulation of the target system, which provides individual positions of bodies of the system and velocities, acceleration and other factors for kinematic and kinetic analyses can be obtained.

$$\begin{bmatrix} M & \Phi_q^T \\ \Phi_q & 0 \end{bmatrix} \begin{bmatrix} \ddot{q} \\ \lambda \end{bmatrix} = \begin{bmatrix} Q^A \\ \gamma \end{bmatrix} \quad (1)$$

Table 2. 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

The mass-center coordinates and the angle of each link are set as shown in Fig. 2. The generalized coordinate matrix and the generalized velocity matrix for each mass center are expressed as follows.

$$q_i = [x_i \quad y_i \quad \theta_i]^T$$
$$\dot{q}_i = [\dot{x}_i \quad \dot{y}_i \quad \dot{\theta}_i]^T$$

The whole generalized coordinate matrix and the generalized velocity matrix are expressed as follows.

$$\mathrm{q} = [q_1 \quad q_2 \quad q_3 \quad q_4]^T$$
$$\dot{q} = [\dot{q}_1 \quad \dot{q}_2 \quad \dot{q}_3 \quad \dot{q}_4]^T$$

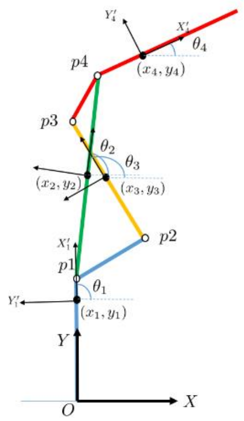


Fig. 2. The definition of generalized coordinates.

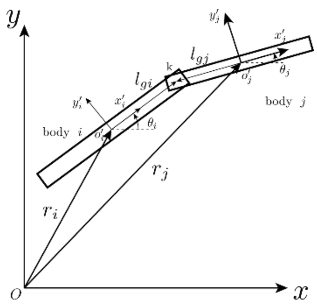


Fig. 3. The definition of kinematic constrains of rotary joint.

The kinematic constrains of rotary joint is defined as shown in Fig. 3 and expressed as follows.

$$\Phi_k^{R(i,j)} = \left(\begin{bmatrix} x_j \\ y_j \end{bmatrix} + A_j \begin{bmatrix} -l_{gj} \\ 0 \end{bmatrix}\right) - \left(\begin{bmatrix} x_i \\ y_i \end{bmatrix} + A_i \begin{bmatrix} l_{gi} \\ 0 \end{bmatrix}\right)$$
$$= \begin{bmatrix} x_j - x_i - l_{gj}\cos\theta_j - l_{gi}\cos\theta_i \\ y_j - y_i - l_{gj}\sin\theta_j - l_{gi}\sin\theta_i \end{bmatrix} = 0$$

The kinematic constraints of rotary joints in rigid knee link are expressed as follows.

$$\Phi^K = \begin{bmatrix} \Phi_{p1}^K \\ \Phi_{p2}^K \\ \Phi_{p3}^K \\ \Phi_{p4}^K \end{bmatrix} = \begin{bmatrix} \left(\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} + A_2 \begin{bmatrix} -l_{21} \\ 0 \end{bmatrix}\right) - \left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + A_1 \begin{bmatrix} l_{12} \\ 0 \end{bmatrix}\right) \\ \left(\begin{bmatrix} x_3 \\ y_3 \end{bmatrix} + A_3 \begin{bmatrix} -l_{31} \\ 0 \end{bmatrix}\right) - \left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + A_1 \begin{bmatrix} l_{41}\cos\phi_1 \\ -l_{41}\sin\phi_1 \end{bmatrix}\right) \\ \left(\begin{bmatrix} x_4 \\ y_4 \end{bmatrix} + A_4 \begin{bmatrix} -l_{41}\cos\phi_4 \\ -l_{41}\sin\phi_4 \end{bmatrix}\right) - \left(\begin{bmatrix} x_3 \\ y_3 \end{bmatrix} + A_3 \begin{bmatrix} l_{32} \\ 0 \end{bmatrix}\right) \\ \left(\begin{bmatrix} x_4 \\ y_4 \end{bmatrix} + A_4 \begin{bmatrix} -l_{42} \\ 0 \end{bmatrix}\right) - \left(\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} + A_2 \begin{bmatrix} l_{22} \\ 0 \end{bmatrix}\right) \end{bmatrix} = 0$$

The link 1 set vertically, and then x_1, y_1, θ_1 is provided as follows according to the absolute constraints.

$$\Phi^A = \begin{bmatrix} x_1 \\ y_1 - L1 \\ \theta_1 - \frac{\pi}{2} \end{bmatrix} = 0.$$

The driving constraints when a constant angular velocity is given to the joint p3 is expressed as follows.

$$\Phi^D = \theta_4 - \omega t = 0$$

The kinematic, and absolute and driving constraints are combined in the matrix as follows.

$$\Phi = \begin{bmatrix} \Phi^K \\ \Phi^A \\ \Phi^D \end{bmatrix} = 0$$

Finally, the Jacobian matrix is follow.

$$\Phi_q = \begin{bmatrix}
-1 & 0 & l_{11}\sin\theta_1 & 1 & 0 & l_{21}\sin\theta_2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & l_{11}\cos\theta_1 & 0 & 1 & l_{21}\cos\theta_2 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & l_{12}\sin\phi_1-\theta_1 & 0 & 0 & 0 & 1 & 0 & l_{31}\sin\theta_3 & 0 & 0 & 0 \\
0 & -1 & l_{12}\cos\phi_1-\theta_1 & 0 & 0 & 0 & 0 & 1 & l_{31}\cos\theta_3 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & l_{32}\sin\theta_3 & 1 & 0 & l_{41}\sin\phi_4+\theta_4 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & l_{32}\cos\theta_3 & 0 & 1 & l_{41}\cos\phi_4+\theta_4 \\
0 & 0 & 0 & -1 & 0 & l_{21}\sin\theta_2 & 0 & 0 & 0 & 1 & 0 & -l_{42}\cos\theta_4 \\
0 & 0 & 0 & 0 & -1 & l_{21}\cos\theta_2 & 0 & 0 & 0 & 0 & 1 & -l_{42}\sin\theta_4 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0
\end{bmatrix}$$

In the MBD analysis, the resultant Jacobian matrix is applied to the differential algebraic equation Eq. (1) and then a generalized acceleration matrix is numerically calculated as a numerical solution.

3. Results

The MBD differential algebraic equation of motion for the knee rigid link model was successfully solved with the generalized acceleration matrix, and the angle and angular velocity were obtained by using the numerical integration of the Runge-Kutta Gill's method [15].
In computer experiments, MATLAB was used.

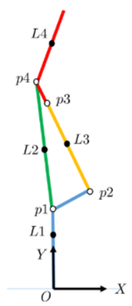


Fig. 4. The singular point posture of the knee model.

3.1. *The movement of the knee link model*

In the process of the knee joint extension, the angle, angular velocity and angular acceleration of each link were analyzed in the condition that a constant angular velocity is given to θ_3. Interestingly, this knee link model has a limitation of the range of the joint extension to prevent the breakage of the knee joint, which was derived in the form of a singular posture (Fig. 4) just before line link L3 and L4 is getting to be on the same straight line. It implies that the rotation of the link is locked in this specific angle. It indicates that the linkage model finely represents the freedom of the knee joint and its limitation. Therefore, the result clarified the importance of the analysis of the singular posture based on MBD [16].

3.2. *The results of the dynamic analysis*

The kinetic analysis was successfully obtained as shown in Fig. 5. The temporal evolution of angle of individual joints as p1, p2, p3 and p4 were denoted by blue, red, yellow and green lines respectively.

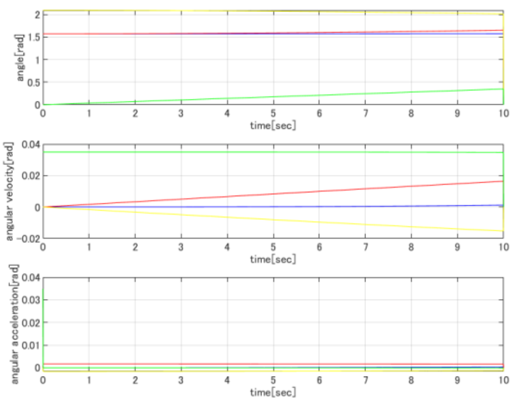


Fig. 5. The result of kinematic analysis ($\theta_4 = 0° \sim 20°$).

4. Conclusion

According to the MBD analysis for the knee joint linkage model, necessary factors in kinematics and kinetics were successfully obtained and the limitation of the movement was clarified in the form of the singular posture. This result demonstrated that MBD-based analysis is beneficial for the reverse engineering to complement the ideal load reduction at a specific posture to avoid risks of the joint movement. In the further analysis, the comparison with/without an additional spring-damper system to prevent the singular posture can be discussed.

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. L. Zhang, G. Liu, B. Han, Z. Wang, Y. Yan, J. Ma and P. Wei, “Knee Joint Biomechanics in Physiological Conditions and How Pathologies Can Affect It: A Systematic Review,” Applied Bionics and Biomechanics, 2020, Article Number 7451683.
2. L. Zhang, Y. Liu, R. Wang, C. Smith and E. Gutierrez-Farewik, “Modeling and Simulation of a Human Knee Exoskeleton’s Assistive Strategies and Interaction,” Frontiers in Neurorobotics, 2021, 15 620928.
3. W. Wu, J. Fong, V. Crocher, P. Lee, D. Oetomo, Y. Tan and D. Ackland, “Modulation of shoulder muscle and joint function using a powered upper-limb exoskeleton,” Journal of Biomechanics, 2018, Apr 27;72:7-16.
4. J. Vantilt, K. Tanghe, M. Afschrift, A. Bruijnes, K. Junius, J. Geeroms, E. Artbelien, F. Groote, D. Lefeber, I. Jonkers and J. Schutter, “Model-based control for exoskeletons with series elastic actuators evaluated on sit-to-stand movements,” Journal of Neuroengineering and Rehabilitation, 2019, Jun 3;16(1):65.
5. S. Schrade, Y. Nager, A. Wu, R. Gassert and A. Ijspeert, “Bio-inspired control of joint torque and knee stiffness in a robotic lower limb exoskeleton using a central pattern generator,” IEEE, 2017, Jul;2017:1387-1394.
6. R. McGrath, M. Ziegler, M. Pires-Faernandes, B. Knarr, J. Higginson and F. Sergi, “The effect of stride length on lower extremity joint kinematics at various gait speeds,” PLoS One, 2019, Feb 22;14(2):e0200862.
7. K. Gui, H. Liu and D. Zhang, “A generalizes framework to achieve coordinated admittance control for multi-joint lower limb robotic exoskeleton,” IEEE, 2017, Jul;2017:228-233.
8. K. Mankala, S. Banala and S. Agrawal, “Novel swing-assist um-motorized exoskeletons for gait training,” Journal of Neuroengineering and Rehabilitation, 2009, Jul 3;6:24.

9. 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.
10. P. E. Nikravesh, *Planar Multibody Dynamics: Formulation, Programming with MATLAB, and Applications*, 2nd edn., CRC Press, Boca Raton, 2018.
11. 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.
12. 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.
13. 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.
14. Arizono orthopedic supplies Co., Ltd.: Products, http://arizono.co.jp/products/products_links/products_2/
15. 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

Mr. Pancho Dachkinov

He received his Master's degree in Engineering in 2018 from the Technical University of Sofia, Bulgaria. He worked in the Institute of Robotics – Bulgarian Academy of Sciences 2016-2019. He is currently a doctoral student in Kyushu Institute of Technology, Japan

Mr. Kohei Tanaka

He is currently a fourth grade undergraduate student in Department of Mechanical and Control Engineering, Faculty of Engineering, Kyushu Institute of Technology, Japan. He is interested in 3D printing and compliant mechanisms.

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 Intelli- gence Systems. His research interests include non- linear dynamics and robotics. He is a member of IEEE.

A Drone-Based Concrete Crack Inspection System by Using Morphological Component Analysis and Sub-Pixel Width Estimation

Ankur Dixit[1], Wataru Oshiumi[1], Hiroaki Wagatsuma[1, 2]
[1]Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-Ku, Kitakyushu, 808-0196, Japan
[2]RIKEN CBS, Japan
E-mail: ankur.dixit003@gmail.com, oshiumi.wataru745@mail.kyutech.jp, waga@brain.kyutech.ac.jp

Abstract

Social infrastructure inspections were relying on human experts, and then a recent topic is a possible implementation to realize an automated inspection based on machines and sophisticated software algorithms. We have studied an advancement of Morphological Component Analysis (MCA) to apply the concrete-crack position estimation especially for submillimeter-width cracks, which are highly difficult for traditional methods to detect finely. We demonstrated a concrete crack detection from images obtained from proximity cameras attached a specialized multi-copter, by using the MCA-based crack position estimation and the linear regression-based sub-pixel width estimation. It will contribute to the actual field work not only for the concrete crack detection but also various social infrastructure inspections.

Keywords Morphological Component Analysis (MCA), multicopter, concrete crack detection, sub-pixel estimation

1. Introduction

In the consideration of the sustainable society, the social infrastructure maintenance is highly important issue, such an early-stage detection of cracks in the concreate buildings[1][2][3][4][5]. Traditionally, human experts have been devoted to careful inspections; however, the inspection work in high places has a certain risk on accidents and the height safety for human workers requires special cares and huge costs. Therefore, a recent trend is to provide an automated and remote detection of flying systems, known as a drone. In the national research project in Japan, called “SIP Infrastructure Regional Implementation Support Teams: Promoting Innovation in Regional Infrastructure Maintenance,” innovative inspection systems were developed in a competitive style of teams organized by academic-industrial partnerships [6]. We have contributed to the project in a team of the project as a part of academic advancement [7], by using Morphological Component Analysis (MCA) [8] for the concrete-crack detection and position estimation, particularly for submillimeter-width cracks, which are highly difficult for traditional methods. We demonstrated not only a concrete crack detection from images obtained from proximity cameras attached a specialized multi-copter, but also the length and width estimation, which is crucial for the completion of the inspection report associated with a database to record a secular change. In this study, the sub-pixel estimation method for lengths and widths of concrete cracks was proposed to adapt to the direction of extension of cracks even they were wandering.

The remainder of this paper is organized as follows. After giving brief introduction of our work in section 1, in section 2, the hardware and software systems were explained as the overview of the inspection procedure from the image acquisition to an inspection report. In section 3, the proposed method was described and in section 4 experimental results were described, which is followed by conclusion in section 5.

2. Inspection System Design Using a Multicopter

2.1. *Hardware design*

In this study, we used a multicopter, or simply drone, with high proximity camera as shown in Figure 1. The video image can be clearly obtained with the same distance *L* between the camera and the concrete surface.

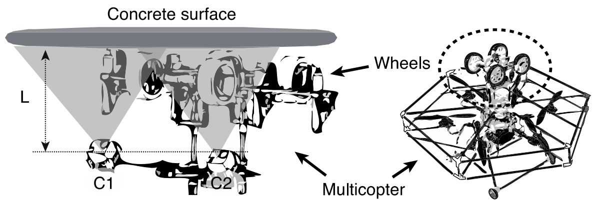


Fig. 1. The multicopter inspection system with wheels to keep the same distance *L* between cameras (*C1* and *C2*) and the monitoring surface (illustrated as a redraw version from [7]).

2.2. *Software design*

As illustrated on Figure 2, acquired images in the form of videos is converted to split images as snapshots of the target concrete surface. Large (more than 1mm width; Fig. 3 (a)), middle (about 1mm width) and minor cracks (less than 1 mm width; Fig. 3 (b)) were detected by using multiple methods with image morphology (IM), anisotropic diffusion filter (AD) and MCA methods as Dixit and Wagatsuma reported [9][10]. After the estimation of crack positions, an accurate gauging is necessary for making a fine inspection report as well as reports provided by human experts.

In this study, we focused on the possible method for the sub-pixel estimation method for lengths and widths of concrete cracks. The problem is that most of cracks are not simply extending straightforward and those extend in a zig-zag manner, which prevent an appropriate measurement.

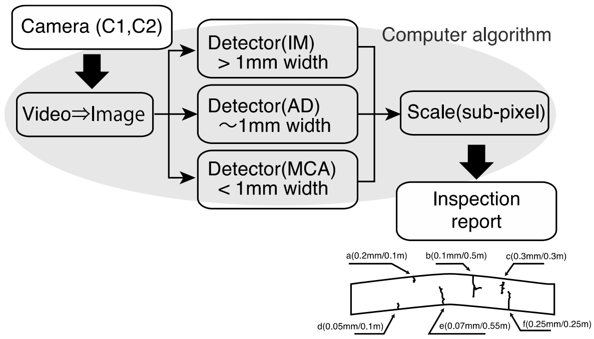
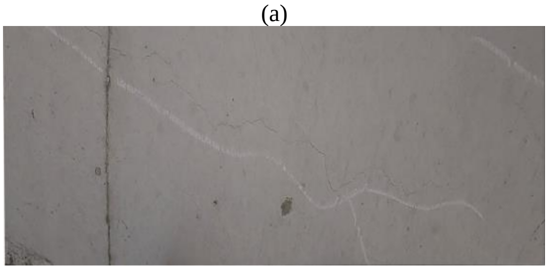


Fig. 2. The whole workflow of the information processing from the image data acquisition to the inspection report.

(a)



(b)

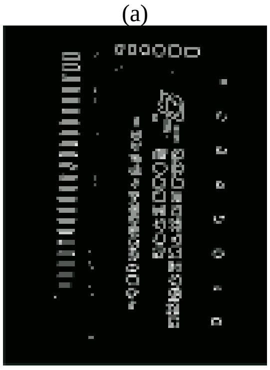
Fig. 3. Example of concrete cracks. (a) >1mm width and (b) <1mm width.

3. Proposed Method

In the consideration of the crack-width estimation, which is traditionally performed by the human expert with the crack gauge (Fig. 4(a)), an edge detection is necessary to estimate the gap of normal regions to represent a crack and then a tangential section is required for an appropriate estimation of the width as illustrated in Figure 5. According to those requirements, the sub-pixel estimation method was proposed as a procedure with steps:

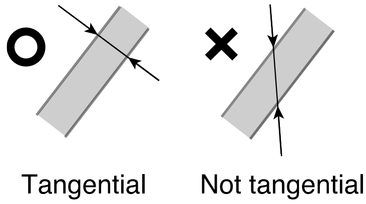
1) Extract a pre-straight line as target
2) Estimate the tangential angle against the target line (Fig. 5 (a))
3) Provide a curve fit, which trace the discrete stair-like color difference due to pixel-based representation in the image (Fig. 5 (b) as the ideal data)
4) Obtain a transition point in the fitted curve, which represents the crack in the form of the negative valley of the brightness of the color (Fig. 5 (b) as the estimated width)
5) Plot the width and length of the crack at the position of the detected crack for making an appropriate inspection report (Fig. 5 (c))

(a)

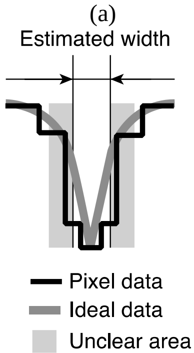


(b)

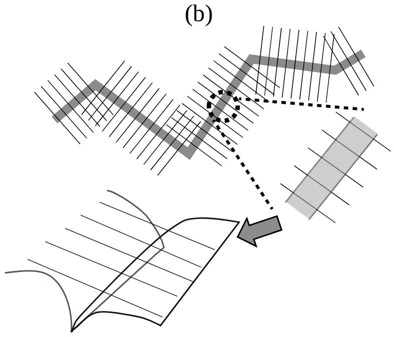
Fig. 4. The traditional crack gauge and a processed result by using AD filter described in Fig. 2 [9][10].



(a)

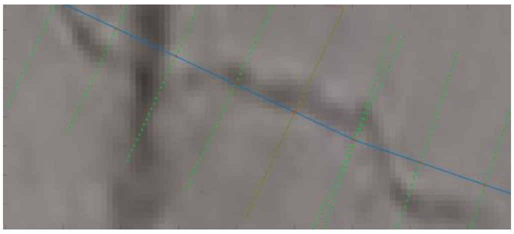


(b)

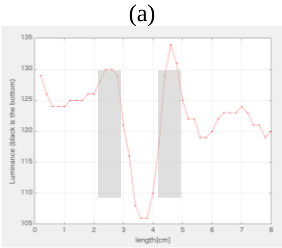


(c)

Fig. 5. The proposed sub-pixel width estimation method for an appropriate measurement.



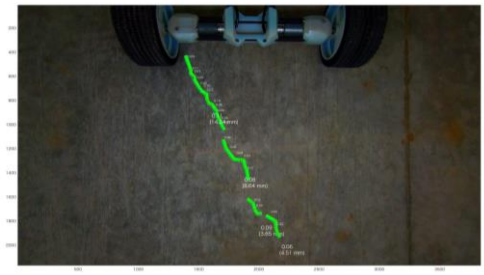
(a)



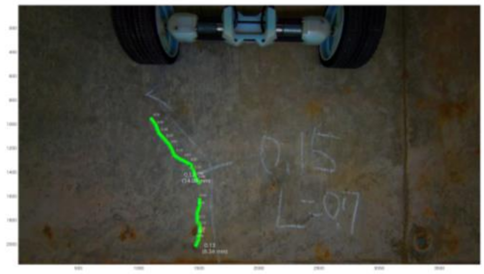
(b)

Fig. 6. An example of the tangential section of the target crack to estimate the width of the crack.

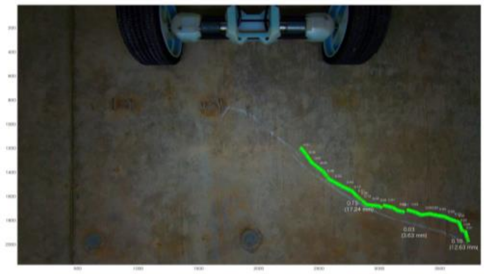
(a)



(b)



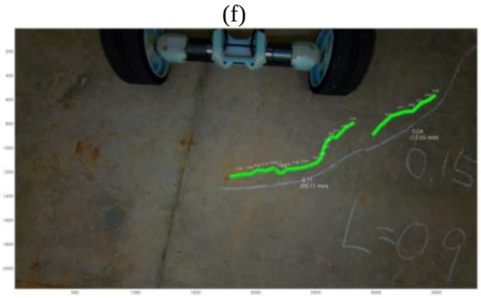
(c)



(d)

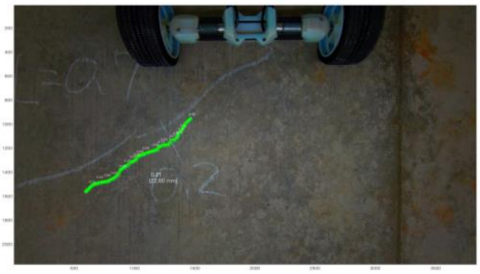
(e)

(f)



(g)

(h)



(i)

Fig. 7. Estimated widths and lengths of cracks superimposed on target images. The human measurement was described by chalk lines.

4. Results

The proposed method was applied to the real environment data that collected by the specialized multicopter [7]. After the crack detection (Fig. 2; Detectors), the gauging of the width and length was done successfully. The evaluation was verified with the human expert data provided by Shin-Nippon Nondestructive Inspection Co., Ltd. The numerical analysis and visualization were derived from the programming code with MATLAB. Fig. 7 showed the results.

5. Conclusion

We demonstrated the length and width estimation of concrete cracks by proposing the sub-pixel estimation method to adapt to the direction of extension of cracks even they were wandering. In further analyses, full automated concrete detection and estimation is expected [11-20].

Acknowledgment

This work was supported in part by SIP Infrastructure Regional Implementation Support Teams: Promoting Innovation in Regional Infrastructure Maintenance. The authors gratefully acknowledge Dr Hideki Wada and Shin-Nippon Nondestructive Inspection Co., Ltd for assistance with understanding of the essential parts in the nondestructive inspection and the correction of the field-work inspection data for the concrete crack detection.

References

1. H.-W. Song, H.-J. Cho, S.-S. Park, K.-J. Byun, K. Maekawa, "Early-age Cracking Resistance Evaluation of Concrete Structures," Concrete Science and Engineering, Vol. 3, No. 10, pp. 62 - 72, 2001
2. A. Bentur, K. Kovler, "Evaluation of early age cracking characteristics in cementitious systems," Materials and Structures, Vol. 36, Article number 183, 2003
3. H. Mihashi, J. P. de B. Leite, "State-of-the-Art Report on Control of Cracking in Early Age Concrete," Journal of Advanced Concrete Technology, Vol. 2, No. 2, pp. 141-154, 2004
4. M. Safiuddin, A.B.M.A. Kaish, C.-O. Woon, S.N. Raman, "Early-Age Cracking in Concrete: Causes, Consequences, Remedial Measures, and

Recommendations," Applied Sciences, Vol. 8, No. 10, Article number 1730. 2018
5. J. Xin, Y. Liu, G. Zhang, Z. Wang, J. Wang, N. Yang, Y. Qiao, "Evaluation of early-age thermal cracking resistance of high w/b, high volume fly ash (HVFA) concrete using temperature stress testing machine," Case Studies in Construction Materials, Vol. 16, Article number e00825, 2022
6. Subcommittee on Promoting Regional Implementation of New Technologies SIP Infrastructure Program Coordination Committee, Organization for Promotion of Civil Engineering Technology, JSCE, "Report of SIP Infrastructure Regional Implementation Support Teams: Promoting Innovation in Regional Infrastructure Maintenance
7. "R&D of a multicopter-based inspection robotic system with visual observation and hammering test devices,"https://www.jst.go.jp/sip/dl/k07/booklet_en/51.pdf
8. J.-L. Starck, Y. Moudden, J. Bobin, M. Elad, and D. L. Donoho "Morphological component analysis", Proc. SPIE 5914, Wavelets XI, 59140Q, 2005
9. A. Dixit, H. Wagatsuma, H., "Investigating the effectiveness of the sobel operator in the MCA-based automatic crack detection," 2018 4th International Conference On Optimization and Applications (ICOA), pp. 1–6, 2018
10. A. Dixit, H. Wagatsuma, H., "Comparison of effectiveness of dual tree complex wavelet transform and anisotropic diffusion in mca for concrete crack detection," 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 2681–2686, 2018
11. T. Yamane, P.-j. Chun, "Crack Detection from a Concrete Surface Image Based on Semantic Segmentation Using Deep Learning," Journal of Advanced Concrete Technology, Vol. 18, No. 9, pp. 493-504, 2020
12. S. Oh , S. Ham, S. Lee, "Drone-Assisted Image Processing Scheme using Frame-Based Location Identification for Crack and Energy Loss Detection in Building Envelopes," Energies, Vol. 14, No. 19, Article number 6359, 2021
13. P.-J. Chun, S. Izumi, T. Yamane, "Automatic detection method of cracks from concrete surface imagery using two-step light gradient boosting machine," Computer-Aided Civil and Infrastructure Engineering, Vol. 36, No. 1, pp. 61-72, 2021
14. Ç.F.Özgenel, A. Gönenç Sorguç "Performance Comparison of Pretrained Convolutional Neural Networks on Crack Detection in Buildings", ISARC 2018, Berlin, pp. 693-700, 2018
15. L. Zhang, F. Yang, Y. D. Zhang, Y. J. Zhu, "Road Crack Detection Using Deep Convolutional Neural Network," 2016 IEEE International Conference on Image Processing (ICIP), 2016
16. Ç. F. Özgenel, "Concrete Crack Images for Classification," https://data.mendeley.com/datasets/5y9wdsg2zt/2
17. Chao Su, W. Wang, "Concrete Cracks Detection Using Convolutional NeuralNetwork Based on Transfer Learning," Artificial Intelligence for Civil Engineering, Vol. 2020, Article ID 7240129, 2020
18. S. Li, and X. Zhao, "Image-Based Concrete Crack Detection Using Convolutional Neural Network and Exhaustive Search Technique," Advances in Civil Engineering, Vol. 2019, Article ID 6520620, 2019
19. S. E. Park, S.-H. Eem, H. Jeon, "Concrete crack detection and quantification using deep learning and structured light," Construction and Building Materials, Vol. 252, Article ID 119096, 2020
20. C. V, Dung, L. D. Anh, "Autonomous concrete crack detection using deep fully convolutional neural network," Automation in Construction, Vol. 99, pp. 52-58, 2019

Authors Introduction

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.

Mr. Wataru Oshiumi

He is currently a fourth grade undergraduate student in Department of Mechanical and Control Engineering, Faculty of Engineering, Kyushu Institute of Technology, Japan. He is interested in image processing systems.

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 Intelli- gence Systems. His research interests include non- linear dynamics and robotics. He is a member of IEEE.

A Systematic Geometric Design Method of Flexible Bars Available for Personalized Knee Orthoses with Spring-Damper Functions

Pancho Dachkinov[1]
[1]Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-Ku, Kitakyushu, 808-0196, Japan

Shintaro Kasai[1]
[1]Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-Ku, Kitakyushu, 808-0196, Japan

Kohei Tanaka[1]
[1]Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-Ku, Kitakyushu, 808-0196, Japan

Hiroaki Wagatsuma[1, 3]
[1]Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-Ku, Kitakyushu, 808-0196, Japan
[3]RIKEN CBS, Japan
E-mail: : dachkinov.pancho-nikolaev608@mail.kyutech.jp, kasai.shintaro660@mail.kyutech.jp, tanaka.kohei886@mail.kyutech.jp, waga@brain.kyutech.ac.jp

Abstract

3D printed compliant mechanisms were recently highlighted not only in the traditional way, but also in the reverse engineering of human joint supportive devices. However, a systematic design principle is unclear for solving of the issue how it can be modifiable to fit to target body requirements in the sense of the personalization. We proposed a geometrical approach providing the target design by using a morphological replacement at the concentration of stress. This concept can be verified in orthoses to improve the knee joint function by 3D printed flexible prototypes. It allows larger deformations of the bars to control joint's motion. Theoretical analysis and experiments demonstrate the flexibility and support during the flexion and extension of the knee. It implies the impact of the geometry in orthosis design.

Keywords: 3D Printing, Flexible Bars, Compliant Mechanisms, Knee Supportive Orthoses, Customized Design

1. Introduction

The advancements in 3D Printing and Additive Manufacturing and the increased interest from a research perspective in the recent years have enabled their applications outside of the prototyping stage and directly into functional parts. Additive Manufacturing has been widely adopted in industries such as Space Engineering, Automotive, Medicine and Stomatology, Jewelry and Fashion, and more. Amongst its advantages in comparison with the conventional technological processes are the ability to control the geometry from inside and therefore the density of the structure and optimizing the mass to strength ratio of the part, realizing complex shapes, and avoiding entirely the assembly process, fabricating mechanisms that rely solely on deformation to execute the motion called compliant mechanisms.

These compliant mechanisms have controllable elasticity (depending on the geometry and the infill density, and pattern) and repetitive cycle motion. Those unique propertied make them strong candidate for orthotic devices applications. In addition, the 3D Printing compatibility of the compliant mechanisms allows deeper customization regarding the individual patient's requirements.

The rest of this article is structured as follows: Section 2 gives the research background, the Methodology for developing various the designs is described in Section 3, Section 4 explains and summarizes the results, Section 5 gives a discussion, Section 6 concludes the paper.

2. Related Work

From a research perspective, compliant mechanisms are not new topic due to their capabilities and behavior. The topology is investigated in [1] and [2] gives the design of parallel-guiding mechanisms. [3] provides the design of large displacement compliant joints. This research works proposes a revolute and a translational solution of joints with large range of motion. In [4] a design based on compliant building block for micro/nano positioning is described. In [5], [6] and [7] the design and applications of more commonly spread type of compliant joint – a cross-spring pivot are analyzed.

In addition to advancements in the compliant joints research, the compatibility of the 3D Printing materials and their properties is crucial to the proper function of the mechanisms. The deformations of 3D Printed structures made from a flexible filament with varied relative density (infill) are explained in [8]. In [9], the authors discuss the properties of the medical grade TPU. Other works show the application of the 3D Printing the Orthopedics for creating orthoses such as [10], [11] and [12]. The methodology of designing such orthotic devices for the human arm is provided in [13].

3. Methodology

3.1. *Flexible Bars and 3D Printing Parameters*

The beams are design in Autodesk Inventor Professional 2021. Initially, divided into four groups (A, B, C and D) based on common geometrical features. For the bars in Group A, geometry is removed from a simple beam annotated with 0. In Groups B, C and D, geometry is added to the beam as a half regular polygon with different number of the sides – square, hexagon and half circle respectively. All bars have same length l=160 mm and a square cross-section a=10 mm. The compliant hinging geometry is varied. Groups with all bar variations are shown on Figure 1. The dimensions that are unique for a given group are marked on the figure. Once 3D Printed, the physical pretotypes of the bars will be attached laterally of the knee. Therefore, the bars have to match certain requirements such as compactness, providing lateral stability of the knee joint and supporting the knee flexion and extension.

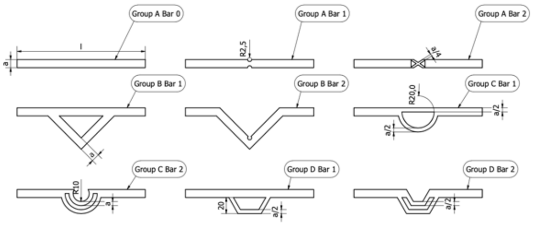


Fig. 1. Bar Groups

All bars are 3D Printed by using the Fused Deposition Modeling (FDM) technology on a 3D printer Anycubic i3 Mega from a flexible Thermoplastic Polyurethane (TPU) filament called Ninjaflex from the company Ninjatek [14]. The specimens are printed under the same conditions and by using the same settings (parameters) as shown in Table 1 configured in the slicing software Ultimaker Cura. The orientation for one of the specimens on the 3D Printing building platform is the same – the models are flat with their most surface in the platform. As shown in Table 1, the samples are printed with infill density of 90% and infill pattern called Cross 3D and according to the description of the slicing software Ultimaker Cura [15] it is a strong infill appropriate for 3D deformations.

Table 1. 3D Printing Processes

Parameter	Values
3D Printer	Anycubic i3 Mega
Slicing Software	Ultimaker Cura
3D Printing Material	Ninjatek Ninja Flex TPU Filament
Printing Temperature, ℃	230 (Recommend for the material)
Platform Temperature, ℃	60
Supporting Material	No
Infill Density, %	90
Infill Type	Cross 3D
Layer Height, mm	0.1
Printing Speed, mm/s	30

The infill density percentage is an important parameter related with stiffness of the parts. In general, the higher the infill density the stiffer the part is. However, it should
be noted that increasing the infill leads to heavier samples and interesting observation shown in Figure 2 is that the time duration of the printing process grows linearly up to 90% and then increases significantly at a 100%. Therefore, 90% was chosen for infill density across all samples.

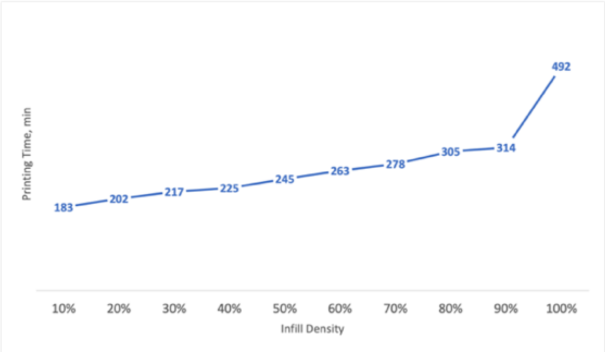


Fig. 2. Relationship between the 3D Printing Infill Density and the Printing Duration

The example shown on Figure 2 is for Group A Bar 0 but applies for all samples.

3.2. *Load Analysis of the Bars*

The main requirement of the beams is to provide the lowest resistance support the motion during knee flexion (in plane bending while at the same time preventing lateral (out of plane) motion. In addition, the full range of motion of the human knee joint must be considered as well – from fully extended knee – 0° to fully flexed knee at 140 ° degrees. The designed compliant beam must provide deformation in this range. However, due to its symmetry, each side of the joint need to deform at about an angle of 70°.
The samples have been tested in non-linear environment using Autodesk Inventor Nastran.
The models illustrated in Figure 3 represent a schematic view of the loading conditions. The members are fixed in one end and at the other a roller support is used. A force is applied at the middle of the models.

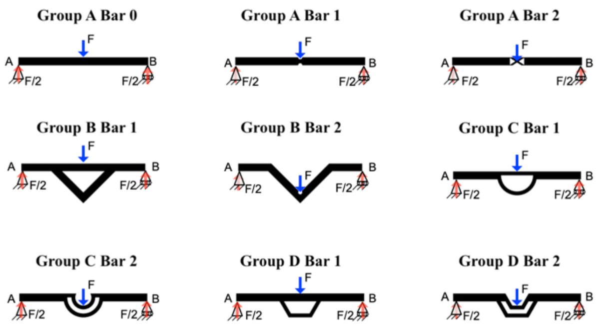


Fig. 3. Loads on the Bars

The results of the FEA analysis will show which bars deform the most under the same boundary conditions. The beams realizing the highest values of the deformations are suitable candidates as a knee assistive device.

4. Results

All beam models were subjected to a Finite Element Analysis in the Non-linear static stress environment of Autodesk Inventor Nastran where a 3-point bending test was conducted on the samples. The study accounts for the large deformations and uses a non-linear elastic material. All beams are loaded with the same force F=150 N in negative Y direction as shown on Figure 4. The mechanical properties of the material are based on the Ninjaflex TPU filament – Young's Modulus is 12 MPa, Tensile Strength 26 MPa and Poisons Ratio is 0.35. The simulation of each sample includes 20 increments. The purpose of this study is to show the beams with the highest deformations while the same loading is applied based on the geometry while accounting for the non-linearities from large deformations of the model and the anisotropic material. Figure 4 shows the setup of the analysis and the results for the models with highest displacements – Group A Bar 2 and Group C Bar 2 respectively.

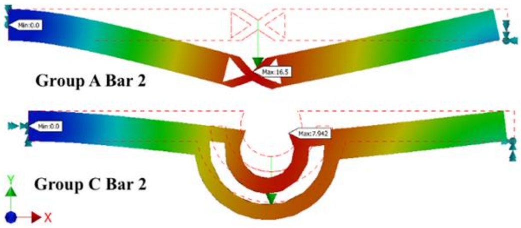


Fig. 4. Non-linear Static Stress Analysis of the Bars with the Largest Deformations

The values of the stress and the displacements from the studies are given in Table 2. The plots on Figure 5 summarize the results from the analysis for the stress and the displacements of the beams during the FEA bending test. The values are arranged from the smallest to the largest.

Table 2. Mechanical Behavior of the Bars

Name	**Stress, σ_{Max}, MPa**	**Displacements, Δ, mm**
Group A Bar 0	33.75	1.684
Group A Bar 1	45.13	6.119
Group A Bar 2	62.73	16.5
Group B Bar 1	30.04	0.7827
Group B Bar 2	39.17	3.996
Group C Bar 1	34.47	1.819
Group C Bar 2	46.58	7.942
Group D Bar 1	35.37	1.806
Group D Bar 2	4.681	0.2599

As it can be observed in Figure 5, the samples that experienced the highest values of the stress and the displacements are from Group A bars 3, 2 and 4. The lowest values belong to samples from Group B – bars 1 and 3. Similarly, the displacements with highest values are again members from Group A Bar 2 and Group C Bar 2 respectively and the lowest – Group D bars 2.
As it is expected, the highest deformations lead to the highest levels of the stress in the samples.

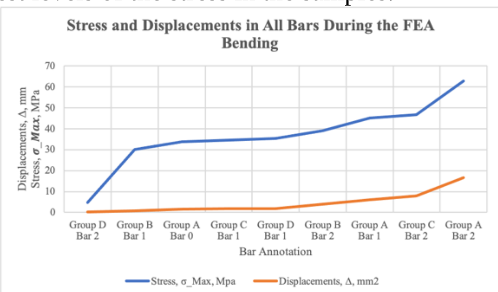


Fig. 5. Mechanical Behaviors of the Bars Based on the Geometry

For the samples to be applicable for an Orthotic device, a good balance of the stress and the strain (displacements) relationship is required. Therefore, possible candidates are Group A Bar 2 and Group C Bar 2.
The force-displacement relationships for all beams after the simulations are illustrated on Figure 6. The non-linear behavior of the models can be observed from the plots. For some beams such as Group A Bar 0, Group B Bar 1 and Group D Bar 2, the load of 150 N was not significant enough to deform them out of the linear elastic region.

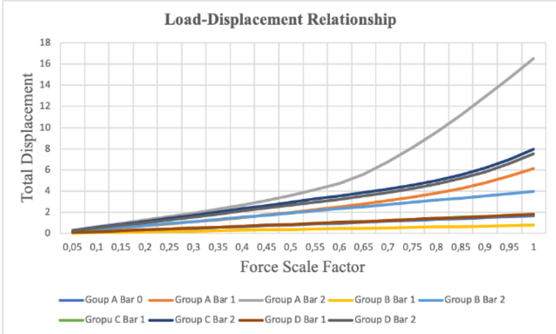


Fig. 6. Non-linear Deformation of the Bars – FEA Results

All 9 specimen were 3D Printed following the with the parameters from Table 1. The physical prototypes made from TPU filament are show on Figure 7.

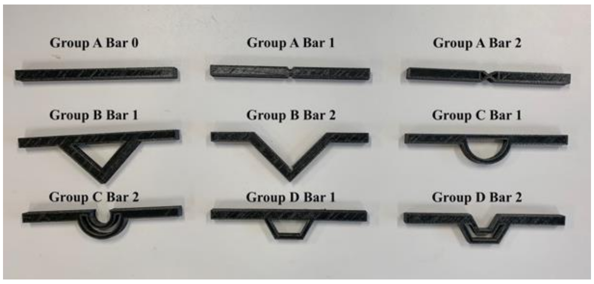


Fig. 7. Sample 3D Printed from Flexible TPU Filament

Figure 8 illustrates the measured 3D Printing time in minutes a) and the measured mass of the physical samples in grams b). It can be observed that Group B Bar 1 takes more 2 hours longer than the other samples. Another important parameter is the mass of the models – due to their attachment laterally of the knee, compactness and low mass are required from the beams. Here again Group B Bar 1is the model with the highest mass – 21 grams, 6 grams heavier than the next beam Group D Bar 1.

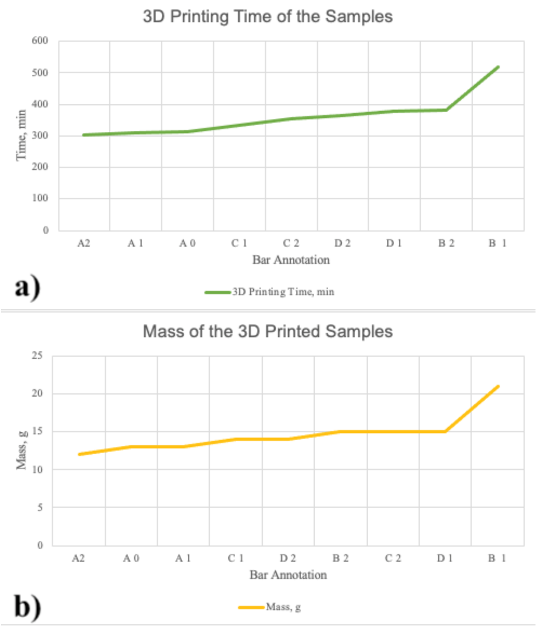


Figure 8: Duration of the 3D Printing and Mass of the Samples

In the process of designing and evaluating the level of applicability of the beams for the task, all of the above parameters have to be carefully balanced. This is reflected in the evaluation Table 3, created for all bars. The evaluation criteria are: high level of the displacements (high flexibility), low level of the stress, low duration of the printing process and mass of the prototype. These four parameters are labeled as “good”, “average” and “poor” and rated with points 1, 0 and -q respectively. After that the points are combined and the beams with most points are the most suitable to be implemented as a compliant hinge of an orthotic device for the knee joint.

Table 3. Evaluation of the Samples

Name	Stress	Displacements	Time	Mass	Score
A 0	poor	poor	good	good	0
A 1	good	good	good	good	4
A 2	good	good	good	good	4
B 1	poor	poor	poor	poor	-4
B 2	avg.	avg.	poor	avg.	-1
C 1	avg.	avg.	avg.	avg.	0
C 2	good	good	avg.	poor	1
D 1	avg.	avg.	poor	poor	-2
D 2	poor	poor	avg.	avg.	-2

As it can be observed from the table, Bars 1 and 2 from Group A and Bar 2 from Group C are the models with most points and therefore most appropriate for the task.

5. Discussion

In this research we investigate the application of the flexible compliant beams fabricated by using the advancements in the Additive Manufacturing in the field of orthopedics as a knee supportive device.
In order to achieve that, 9 models were develop and analyzed in Non-linear studies accounting for the large displacements the supports have to undergo in a real working environment. However, the simulations are not used to provide precise values of the stress-displacement relationships but to limit the models with less flexible properties due to their geometry.
It is important to mention, the out of plane deformations that have to be eliminated to provide a lateral stability of the knee were not included in the FEA studies. They are subject to the future development of the models.
After the evaluation process of the samples, due to its better out of plane stability Bar 2 from Group C was implemented in such a device 3D Printed from a material PPGW. The prototype support is shown on Figure 9 a) extended and b) flexed knee. It allows free motion (bending) and stabilizes the joint during walking.

Fig. 9. Physical Prototype Used in Orthotic Support

As part of the future plans, the lateral deformations have to include in the studies and models that provide greater resistance will be developed.
Examples of improving the lateral stiffness is shown on Figure 10 a) by replacing the square cross-section with a “I” beam cross-section and increasing the range of motion is illustrated in b).
Another unique opportunity provided by the additive processes is the infill density variability allowing local control of the stiffness of the model.
Combining the above parameters will unlock the great potential of the additive manufacturing for customizable solutions depending on the individual requirements.

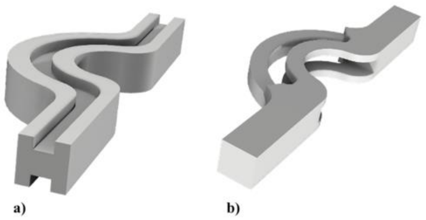


Fig. 10. 3D Representation of the Beam

6. Conclusion

This paper analyses the mechanical behavior of 3D Printed flexible beams with application as a orthotic device. The bars use their compliance to execute the motion and have different hinging mechanism. A non-linear finite element analysis accounting for the anisotropic material and the large deformations of the flexible models was conducted on all beams to determine the most appropriate for the application. In the selection process the 3D Prating time and the mass of the samples were included. Plots of the results were provided. An evaluation table ranking the criteria was created to objectively determine the beams with most applicable behavior.

Advantages of the designed beams are the low-cost, low mass and compactness in comparison to existing solutions that are hardly affordable and complicated. It allows for simple attachment and replacement with a support with different stiffness depending on the stage of the healing process of the patient.

In the future, this table will be extended including more parameters such as various cross-section geometries, infill patterns and densities and multi-material assemblies combining flexible and stiff materials in different areas of the supportive device.

The constantly increasing capabilities of the additive manufacturing processes and 3D Printing more specifically allow for fabricating complex geometries with controlled stiffness for the first time providing an opportunity for meeting the personal requirements in the field of orthoses – the most important link of the patient-doctor relationship aiming for improving the quality of life.

Acknowledgements

This work was supported in part by EU Horizon 2020 MSCA RISE project “Cyber-Physical Systems for PEdagogical Rehabilitation in Special Education (CybSPEED)“, JSPS KAKENHI (16H01616, 17H06383) and the New Energy and Industrial Technology Development Organization (NEDO), 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).

References

1. M., Murphy, A., Midha, L., Howell, “The Topological Synthesis O\of Compliant Mechanism”, Mech. Mach. TheoryVol.31, No. 2, pp. 185-199, 1996, 0094 I14X(95)00055-0.
2. N., T., Pavlovic, N., D., Pavlovic, “Compliant mechanism design for realizing of axial link translation”, Mechanism and Machine Theory 44 (2009) 1082–1091.
3. B., Trease, Y., Moon, S., Kota, “Design of Large-Displacement Compliant Joints”, Journal of Mechanical Design, JUuly 2005, Vol. 127, ASME, pp. 788-798.
4. Y., Jingjun, P., Xu, Z., Shanshan, B., Shushing, Z., Guanghua, “A new large-stroke compliant joint & micro/nano positioner design based on compliant building blocks”, ASME/IFToMM International Conference on Reconfigurable Mechanisms and Robots, pp. 409*41
5. L., Liu, S., Bi, Q., Yang, Y., Wang, “Design and experiment of generalized triple-cross-spring flexure pivots applied to the ultra-precision instruments”, Review of Scientific Instruments 85, 105102 (2014); doi: 10.1063/1.4897271.
6. B., Jensen, L., Howell, “The modeling of cross-axis flexural pivots”, PII: S0094-114X(02)00007-1, Mechanism and Machine Theory 37 (2002) 461–476.
7. K., Marković, S., Zelenika, “Characterization of cross-spring pivots for micropositioning applications”, doi: 10.1117/12.2178490, Proc. of SPIE Vol. 9517 951727-2.

8. Płatek, P., Rajkowski K., Cieplak, K., Sarzyn´ski, M., Małachowski, J., Woz´niak, R., Janiszewski, J., “Deformation Process of 3D Printed Structures Made from Flexible Material with Different Values of Relative Density”, Polymers 2020, 12, 2120; doi:1
9. J., Xiao, Y,. Gao, “The manufacture of 3D printing of medical grade TPU ”, DOI 10.1007/s40964-017-0023-1.
10. R. Chen, L. Chen, B. Tai, Y. Wang, A. Shih, J. Wensman, “Additive Manufacturing of Personalized Ankle-Foot Orthosis” , Proceedings of NAMRI/SME, Vol. 42, 2014.
11. Y., Lin, L., Huang, C., Chen, “Strength Evaluation and Modification of a 3D Printed Anterior Ankle Foot Orthoses” , Appl. Sci. 2020, 10, 7289; doi:10.3390/app10207289.
12. Lin,K.-W.;Chou,L.-W.;Su, Y.-T.; Wei, S.-H.; Chen, C.-S. Biomechanical Effect of 3D-Printed Foot Orthoses in Patients with Knee Osteoarthritis. Appl. Sci. 2021, 11, 4200. https://doi.org/10.3390/ app11094200
13. Chaparro-Rico,B.D.M.; Martinello, K.; Fucile, S.; Cafolla, D. User-Tailored Orthosis Design for 3D Printing with PLACTIVE: A Quick Methodology. Crystals 2021, 11, 561. https://doi.org/10.3390/ cryst11050561
14. Ninjatek, Ninjaflex, TPU 3D Printing Filament, Technical specifications, 2021/11/24, https://ninjatek.com/shop/ninjaflex - print-guidelines.
15. Ultimaker Cura, 3D Printing slicing software, Infill pattern guide, 2021/12/10. https://all3dp.com/2/cura-infill-patterns-all-you-need-to-know/

Authors Introduction

Mr. Pancho Dachkinov

He received his Master’s degree in Engineering in 2018 from the Technical University of Sofia, Bulgaria. He worked in the Institute of Robotics – Bulgarian Academy of Sciences 2016-2019. He is currently a doctoral student in Kyushu Institute of Technology, Japan

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

Mr. Kohei Tanaka

He is currently a fourth grade undergraduate student in Department of Mechanical and Control Engineering, Faculty of Engineering, Kyushu Institute of Technology, Japan. He is interested in 3D printing and compliant mechanisms.

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 Intelli- gence Systems. His research interests include non- linear dynamics and robotics. He is a member of IEEE.

State-space modeling of motion of fingers measured by Leap Motion Controller

Ryuichi Usami
Graduate School of Humanities and Social Sciences, Hiroshima University,
1-1-1 Kagamiyama, Higashi-Hiroshima city, Hiroshima, 739-8524 Japan

Hideyuki Tanaka
Graduate School of Humanities and Social Sciences, Hiroshima University,
1-1-1 Kagamiyama, Higashi-Hiroshima city, Hiroshima, 739-8524 Japan
E-mail: tanakalpha@hiroshima-u.ac.jp

Abstract

This paper considers modeling of the motion of experts in sports to demonstrate their motion to beginners, studying the dynamics of a transient response from one position to another, such as a crouching start of a short-distance race. A modeling algorithm is developed to remove personal habits of experts from their motion and to present simple models to learners according to their learning stages. It is applied to fingers motion captured by the Leap Motion Controller.

Keywords: Leap Motion Controller, state-space representation, step response, modeling of motion of fingers,

1. Introduction

Motion of players in sports has been measured by sensors. For example, it has been reported that motion sensors were used for capturing the movements of the golf swing[1] and that pressure sensor systems were presented for monitoring the ankle supination torque during sport motions[2]. A wearable inertial sensor network and its associated activity recognition algorithm were presented for accurately recognizing human daily and sport activities[3]. Motion capture also has been used for analyzing sports performance; see a systematic review of the recent developments of motion capture systems for the analysis of sport performance[4]. Moreover, a system that is applicable to marker-less sports movement analysis has been presented[5].

Demonstrating models to learners has proven to be particularly effective in enhancing motor learning[6]. Presenting a model of an expert to beginners is therefor expected to be effective in learning the motion. However, in sports, advanced techniques and skills are too complex for beginners, as well as individual habits. We hence believe that a simple model that removes these factors would make it easier for beginners to learn the motion. We therefore introduce a step-by-step simplification of the behavior of an expert in this research.

In this study, we make mathematical models to give a simplified model for beginners and a complicated model for advanced learners. To this end, we introduce state-space representation driven by a step input, often used for dealing with dynamics in control engineering[7]. We use state-space representations, because it can deal with dynamics of multiple input and multiple output, and we can obtain a simplified model by reducing the order of the state-space representation. We can model a transient response from one position to another by means of a step input, e.g. a crouching start of a short-distance race. We will model a single grasping motion of a hand as a simple example of motion that moves from one position to another, because motions of finger joints are related to each other and are suitable for studying modeling a transient response of sports movements.

As a motion capture device, we use the Leap Motion Controller (LMC), an optical hand tracking module that captures the movements of hands[8]. Accuracy and robustness of the LMC has been analyzed[9], and the performance of the LMC was evaluated with the aid of a professional, high-precision, fast motion tracking system[10].

2. Measuring data

We describe how to measure the finger motion. Figs.1 and 2 show the start (the open hand) and the end (the grasped hand) of transient motion, respectively. The LMC can measure 21 finger joints in a three-dimensional cartesian coordinate system, meaning that the LMC obtains 63 elements at a sampling time. Figs. 3 and 4 show the measuring points and the coordinate system, respectively. We define the output $y(t)$ as a vector with 63 rows as

$$y(t) = \left[p_x(t)^\top \quad p_y(t)^\top \quad p_z(t)^\top\right]^\top,$$

where $p_x(t), p_y(t), p_z(t) \in R^{21}$ are positions in the X-Y-Z, coordinate of 21 joints of fingers. The positions of finger joints are measured at the equal sampling time between the time interval from the open hand to the grasped hand. The time interval is around 1 second, and 100 data is sampled in 1 second. We use Processing for measuring the grasping motion[11,12].

Fig.1. start of measurement

Fig.2. end of measurement

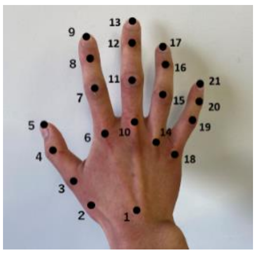


Fig.3. measuring points

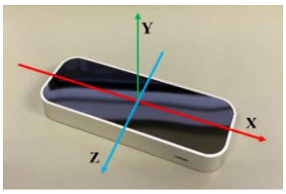


Fig.4. coordinate system

3. Problem setting

Suppose that the sampling time is h and let y_k be the coordinate $y(t)$ at $t = kh$:

$$y_k = y(kh). \tag{1}$$

We model y_k by means of the distance-time state-space representation:

$$x_{k+1} = Ax_k + Bu_k, \tag{2a}$$
$$y_k = Cx_k + \zeta, \tag{2b}$$

where x_k is a state of the system, $x_0 = 0$, ζ is a constant vector, and u_k is as follows:

$$u_k = \begin{cases} 0 & (k = -1, -2, \dots) \\ 1 & (k = 0, 1, 2, \dots) \end{cases} \tag{3}$$

We suppose that $k = 0$ when the transient motion starts and consider the following problems.

Problem 1: Find (A, B, C, ζ) within the degrees of freedom of similarity transformations, given y_k in (2) and (3).

Problem 2: Suppose that the positions of joints fingers are measured as in (1). Find a mathematical model for y_k in (2) and (3).

The complexity of the model of the finger motion in (2) is determined by the size of the matrix A or the order of the system. The lower the order, the simpler the model is. The higher the order, the more accurately the finger motion can be modeled, and the closer motion is generated to the original one.

By choosing the order of the model (2), we can generate samples of the motion of the experts for different level of learners. It is expected that beginners will be able to learn simple motions generated by simpler dynamical models, whereas advanced learners will be able to learn more complex motions. It should be noted that the dynamical system (2) can model the transition from one point to another, though we model finger motion in this study.

4. Solution via deterministic realization

We solve Problem 1 by using the deterministic realization algorithm[13] and apply the algorithm to Problem 2. The deterministic realization algorithm is suitable for the purpose of this study, because the order of the model can be systematically determined by the singular value decomposition (SVD)[7]. Since the deterministic realization algorithm was developed for obtaining a state-space model of an impulse response, we modify the algorithm to obtain a state-space model of the finger motion.

We consider Problem 1. Let us describe y_k for (2) as

$$y_k = \sum_{j=1}^{k} CA^{j-1}B + \zeta, \qquad (4)$$

and let v_k be as follows ($k \geq 1$):

$$v_k = y_k - y_{k-1}. \qquad (5)$$

The signal v_k is the difference of the output y_k. From (4), v_k satisfies the following equation:

$$v_k = CA^{k-1}B. \qquad (6)$$

For positive integers $\tau > n$ and $N > n$, define the Hankel matrix $H \in R^{\tau p \times N}$ from v_k as follows:

$$H = \begin{bmatrix} v_1 & v_2 & \cdots & v_{N+1-\tau} \\ v_2 & v_3 & \cdots & v_{N+2-\tau} \\ \vdots & \vdots & \vdots & \vdots \\ v_\tau & v_{\tau+1} & \cdots & v_N \end{bmatrix}. \qquad (7)$$

We also define the extended observability and reachability matrices as follows:

$$\mathcal{O}_\tau = [C^\top \quad (CA)^\top \quad \cdots \quad (CA^{\tau-1})^\top]^\top, \qquad (8a)$$

$$\mathcal{C}_\mathcal{N} = [B \quad AB \quad \cdots \quad A^{N-1}B]. \qquad (8b)$$

From (6), H satisfies the following equation:

$$H = \mathcal{O}_\tau \mathcal{C}_\mathcal{N}. \qquad (9)$$

Let us compute the SVD of H:

$$H = U\Sigma V^\top \approx [U_1 \quad U_2] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^\top \\ V_2^\top \end{bmatrix} = U_1 \Sigma_1 V_1^\top, \qquad (10)$$

where Σ is a diagonal matrix, and U and V are orthogonal matrices satisfying $U^\top U = I$ and $V^\top V = I$. From (9), $\mathcal{O}_\tau$and $\mathcal{C}_\mathcal{N}$ are expressed as

$$\mathcal{O}_\tau = U_1 \Sigma_1^{\frac{1}{2}}, \qquad \mathcal{C}_\mathcal{N} = \Sigma_1^{\frac{1}{2}} V_1^\top. \qquad (11)$$

From (8), C and B can be obtained as follows:

$$C = \mathcal{O}_\tau(1{:}p,:), \qquad B = \mathcal{C}_\mathcal{N}(:,1), \qquad (12)$$

where we use the colon notation[14]. Let us define $\mathcal{O}_\tau^\downarrow$ and $\mathcal{O}_\tau^\uparrow$ as follows:

$$\mathcal{O}_\tau^\downarrow = \mathcal{O}_\tau(1{:}p(\tau-1),:), \qquad (13a)$$

$$\mathcal{O}_\tau^\uparrow = \mathcal{O}_\tau(\tau+1{:}p\tau,:). \qquad (13b)$$

From (8a), we have

$$\mathcal{O}_\tau^\downarrow = [C^\top \quad (CA^\top) \quad \cdots \quad (CA^{\tau-2})^\top]^\top, \qquad (14a)$$

$$\mathcal{O}_\tau^\uparrow = [(CA)^\top \quad (CA^2)^\top \quad \cdots \quad (CA^{\tau-1})^\top]^\top, \qquad (14b)$$

and hence

$$\mathcal{O}_\tau^\downarrow A = \mathcal{O}_\tau^\uparrow. \qquad (15)$$

We can thus obtain A by solving the least-squares method. Let the estimates of A, B, C and ζ be denoted as $\hat{A}$, $\hat{B}$, $\hat{C}$, and $\hat{\zeta}$, respectively. By setting the initial state as $\hat{x}_0 = 0$, we compute η_k for $k = 0, \cdots, N$ as follows:

$$\hat{x}_{k+1} = \hat{A}\hat{x}_k + \hat{B}u_k, \qquad (16a)$$

$$\eta_k = \hat{C}\hat{x}_k, \qquad (16b)$$

where u_k is the step input in (3). We then have $\eta_k = \sum_{j=1}^{k-1} \hat{C}\hat{A}^{j-1}\hat{B}$ and obtain the following equations from (16):

$$\zeta = y_k - Cx_k = y_k - \textstyle\sum_{j=1}^{k} CA^{j-1}B.$$

We compute an estimate of ζ, by averaging $y_k - \eta_k$:

$$\hat{\zeta} = \frac{1}{N+1}\sum_{k=0}^{N}(y_k - \eta_k)\,. \qquad (17)$$

Thus, estimates of (A, B, C, ζ) for Problem 1 are obtained as $(\hat{A}, \hat{B}, \hat{C}, \hat{\zeta})$, and those of $\hat{y}_k$ in (2) are given by

$$\hat{x}_{k+1} = \hat{A}\hat{x}_k + \hat{B}u_k, \qquad (18a)$$

$$\hat{y}_k = \hat{C}\hat{x}_k + \hat{\zeta}. \qquad (18b)$$

We summarize the above procedure for Problem 1 as the following algorithm:

Step 1: Calculate v_k in (5).
Step 2: Construct the block Hankel matrix H in (7).
Step 3: Compute the SVD in (10).
Step 4: Determine $\mathcal{O}_\tau$ and $\mathcal{C}_\mathcal{N}$ as in (11).
Step 5: Compute the estimate $(\hat{A}, \hat{B}, \hat{C})$ of (A, B, C) from (12) and (15).
Step 6: Obtain an estimate $\hat{\zeta}$ of ζ as (17) and calculate $\hat{y}_k$ in (18).

By applying this algorithm to Problem 2, we have mathematical models for motions of fingers. The model is simplified by the SVD of the Hankel matrix in (10), and the order is determined by the number of the dominant singular values. We hence select the order by choosing the number of the non-zero diagonal elements of H. We thus simplify the state-space model (3) and obtain simplified transient motion from the open hand to the grasped hand. The lower the order of state-space model is, the simpler the transient motion becomes.

5. Experimental results

This section discusses obtained models. We use the first 0.7 seconds of the transition from the open hand to the grasped hand, focusing on the dynamics of the start of the motion and removing the data for the end of 0.3 seconds. The sampling time is $h = 0.01$(s). In Figs. 5-8, we show $p_y(t)$ at the measuring point 13, which is a tip of the middle finger and has one of the largest motions among the measuring points of fingers. The horizontal and vertical axes express the time (s) and the positions (mm) of $p_y(t)$, respectively. Red lines in Figs. 5-8 indicate the position of the original motion, whereas blue ones in Figs. 5-8 show models at the order of 4, 7, 12, and 18, respectively. Figs. 5-8 demonstrate that the output of models (blue lines) become closer to the original motion (red lines) as the order is higher; compare blue lines in Figs. 5 and 8 and notice that the motion drawn by blue line in Fig 5 is much simpler than that in Fig 8. We can thus demonstrate fingers motions with different simplifications.

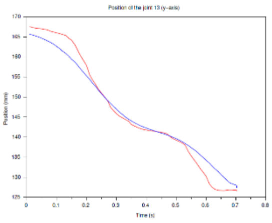


Fig.5. model ($n = 4$)

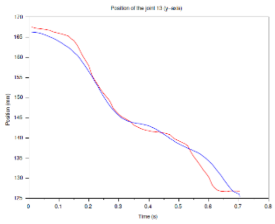


Fig.6. model ($n = 7$)

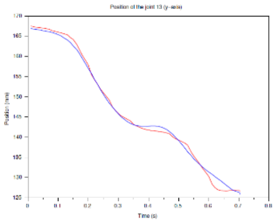


Fig.7. model ($n = 12$)

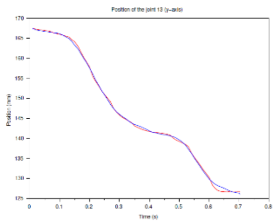


Fig.8. model ($n = 18$)

6. Conclusions

In this study, we have made mathematical models of fingers motions with the aim of simplifying the motion of experts in sports. We solved the problem of modeling fingers motions based on a step response of a state-space representation, by modifying the deterministic realization algorithm. The experimental results showed that the modeling method satisfied the purpose of simplifying complex fingers motions reducing the order of the state-space representation. By choosing the order, we could select the model of finger motion from simple one to more accurate one that is close to the original motion.

References

1. H. Ghasemzadeh, V. Loseu, E. Guenterberg, and R. Jafari, "Sport training using body sensor networks: A statistical approach to measure wrist rotation for golf swing", Proceedings of the Fourth International Conference on Body Area Networks, 2009.
2. D. T.-P. Fong, Y.-Y. Chan, Y. Hong, P. S.-H. Yung, K.-Y. Fung, and K.-M. Chan, "A three-pressure-sensor (3PS) system for monitoring ankle supination torque during sport motions", Journal of Biomechanics, Volume 41, Issue 11, pp. 2562-2566, 2008.
3. Y.-L. Hsu, S.-C. Yang, H.-C. Chang, and H.-C. Lai, "Human daily and sport activity recognition using a wearable inertial sensor network", IEEE Access, vol. 6, pp. 31715-31728, 2018.
4. B. Pueo and J. M. J.-Olmedo, "Application of motion capture technology for sport performance analysis", editor: Federación Española de Asociaciones de Docentes de Educación Física (FEADEF), Retos, 32, pp. 241-247, 2017.
5. B. Rosenhahn, T. Brox, U. G. Kersting, A. W. Smith, J. K. Gurney, and R. Klette, "A system for marker-less motion capture", Künstliche Intelligenz, vol. 1, pp 46-52, 2006.
6. G. H. Sage, "Motor learning and control -A neuro-psychological approach-", Wm. C. Brown Publishers, 1984.
7. T. Katayama, "Subspace methods for system identification", Springer-Verlag London Limited, 2005.
8. ultraleap, "Leap Motion Controller", https://www.ultraleap.com/product/leap-motion-controller/.
9. F. Weichert, D. Bachmann, B. Rudak, and D. Fisseler, "Analysis of the accuracy and robustness of the Leap Motion Controller", Sensors, 13, pp. 6380-6393, 2013.
10. J. Guna, G. Jakus, M. Pogačnik, S. Tomažič, and J. Sodnik, "An analysis of the precision and reliability of the Leap Motion sensor and its suitability for static and dynamic tracking", Sensors, 14, pp. 3702-3720, 2014.
11. Processing, https://processing.org.
12. ultraleap, "API Overview", https://developer-archive.leapmotion.com/documentation/java/devguide/Leap_Overview.html.
13. B. L. Ho and R. E. Kalman, "Effective construction of linear state-variable models from input/output functions", Regelungstechnik, vol. 14, no.12, pp. 545-548, 1966.
14. G. H. Golub and C. F. Van Loan, "Matrix computations, 4th Edition", The Johns Hopkins University Press, 2013.

Authors Introduction

Mr. Ryuichi Usami

He received his bachelor's degree from the School of Education, Hiroshima University, Japan in 2021. He is currently a Master's Program student of Graduate School of Humanities and Social Sciences in Hiroshima University, Japan.

Dr. Hideyuki Tanaka

He graduated master course at graduate school of engineering in Kyoto University and received Dr. (Eng.) from Kyoto University. He is a member of Division of Educational Sciences, Graduate School of Humanities and Social Sciences in Hiroshima University. He is a member of IEEE, SICE, and ISCIE.

Applicability Verification of iWakka Game to Children with Developmental Coordination Disorder

Masakazu Nomura
Department of Electrical and Mechanical Engineering, Graduate School of Engineering
Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya, Aichi 466-8555, Japan

Moe Nishiya
Department of Electrical and Mechanical Engineering, Graduate School of Engineering
Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya, Aichi 466-8555, Japan

Yoshifumi Morita
Department of Electrical and Mechanical Engineering, Graduate School of Engineering
Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya, Aichi 466-8555, Japan

Hideo Yamagiwa
Tokyo Metropolitan Tobu Medical Center, 3-3-25 Shinsuna, Koto-ku, Tokyo, 136-0075, Japan
E-mail: m.nomura.733@nitech.jp, m.nishiya.456@nitech.jp,
morita@nitech.ac.jp, ot_trc@mtrc.jp
www.nitech.ac.jp

Abstract

In our previous study, we developed a testing and training device, "iWakka," of the adjustability for grasping force. Moreover, we developed "iWakka Game," for autistic patients. In this study, we investigated the applicability of the iWakka Game to children with developmental coordination disorder (DCD). We developed a small-sized grasping body and improved the evaluation method. We applied to four children with DCD. Consequently, they completed the evaluation task and the method enabled the extraction of the characteristics of hand dexterity.

Keywords: Rehabilitation, Hand dexterity, Evaluation, Children with developmental coordination disorder

1. Introduction

In daily life, hands and fingers play an important role such as grasping a cup, and handling a chopsticks and forks with appropriate force. The adjustability for grasping force (AGF), which is one of the motor functions of fingers, is the ability to grasp an object with an appropriate force. In our previous study, a training and testing device, "iWakka," for the AGF was developed.[1] iWakka consists of a grasping body, "Wakka," and a measurement device. The measurement device consists of a control box, an iPad, and "iWakka Viewer". When patients grasp Wakka, iWakka Viewer shows the grasping force. They adjust their grasping force to get closer to the target grasping force. The measurement sampling time is 0.1 s, the measurement range is 100–500 g, and the resolution is 1.6 g. Morita et al. observed[2] that adding iWakka to constraint-induced (CI) movement therapy had therapeutic effects for patients after a stroke. To apply iWakka to autistic patients, we also developed "iWakka Game", which is one of iWakka Viewers, by improving the game quality of the original one with the Lodz University of Technology and the Center of Autism Diagnosis and Therapy, Lodz, Poland.[3]

Developmental coordination disorder (DCD) is a common neurodevelopmental disorder characterized by impairments in the development of both fine and gross motor skills.[4] Because grasping is the basic ability for various activities, there is a need for therapists to evaluate the characteristics of hand dexterity and plan more effective training for their hands. However, few studies have focused on the characteristics of hand dexterity. The purpose of this study was to investigate the applicability of the iWakka Game to children with DCD and evaluate their characteristics of hand dexterity.

2. iWakka Game

This game is developed for autistic patients to train their grasping force by improving the quality of the original game. The brown bird in the iWakka Game rises when Wakka is gripped and falls when it is released. When a task begins, the target waveform with red stars moves from right to left. The waveform pauses when the bird hits the cloud placed above and below the bird. Three setting parameters and a target waveform are determined. The three parameters consist of the distance between the bird and the cloud, the measurement range, and the speed at which the waveform flows from right to left. The

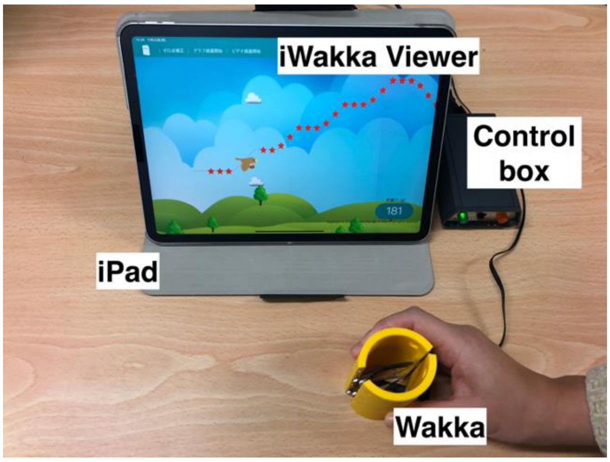


Fig. 1. iWakka.

iWakka Game has nine types of waveforms, such as staircase, slope, and sine wave.

The AGF is evaluated from the mean absolute error between the value of the target waveform and the bird's vertical position that indicates the grasping force by grasping Wakka. However, we could not evaluate the AGF for each isometric muscle activity (IC) to hold an object with a constant grasping force, concentric muscle activity (CC) to grip it while increasing grasping force, and eccentric muscle activity (EC) to release it while decreasing the grasping force.

3. iWakka for Children with DCD

3.1. *iWakka Game*

Figure 1 shows iWakka. Before testing using the iWakka Game, a preliminary experiment was conducted to verify its applicability to children with DCD. We decided the three parameters required to complete the tasks appropriately. The distance between the bird and cloud was 100 g. The measurement range was 100–400 g. The speed was 2.97 mm/s from the right to the left of the screen on an iPad. Notably, a grasping device with a smaller outer diameter and an evaluation method to extract the characteristics were required.

3.2. *Grasping device*

The small-sized grasping body Wakka was developed to reduce the outer diameter by 83%. This was decided based on the requirements of therapists and the ratio of the length from the wrist to the tip of the middle finger. The average ratio is 190–220 mm for adults and 150–190 mm for elementary school children. The relationship between the amount of deformation and the force of the device is linear. Table 1 shows a comparison of the small-sized and original Wakka. The spring constant of the small-sized Wakka was 5.96×10^2 N/m, implying that grasping it with a force of 60.8 g produced a deformation of 1 mm.

Table 1. Comparison of the small-sized and original Wakka.

	Small-sized Wakka	Original Wakka
Spring constant [N/m]	5.96×10^2	4.82×10^2
Outer diameter [mm]	54	65
Height [mm]	80	80
Weight [g]	81	112

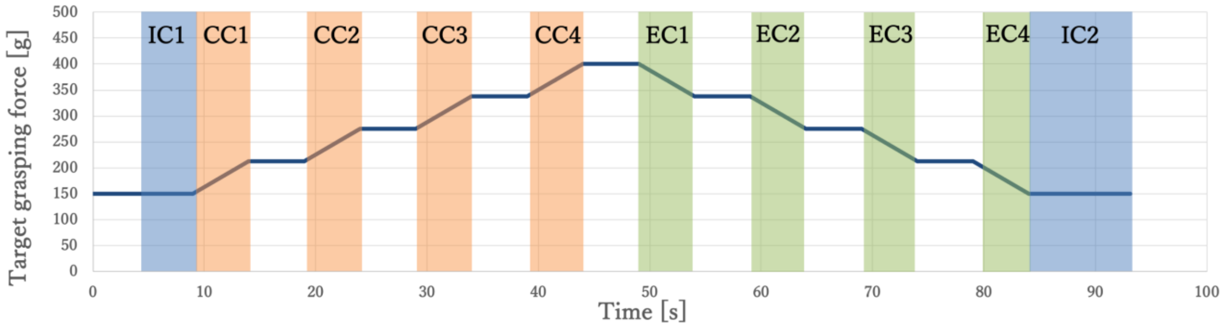


Fig. 2. Target waveform and sections to be evaluated.

3.3. *Evaluation method for hand dexterity*

The waveform shown in Fig. 2 was selected to evaluate the characteristics of hand dexterity. Total time was 93 s. The waveform was subdivided into IC, CC, and EC sections and the ranges of grasping forces. IC was evaluated twice at 150 g. The first evaluation began 3 s after the start. CC and EC were evaluated in four sections of the range of the grasping force. The average absolute error between the grasping force and the target grasping force in each section was calculated as the *AGF*, as shown in Eq. (1).

$$AGF_{(*)} = \frac{1}{T_{(*)}} \sum_{k=1}^{N_{(*)}} |f_d(k) - f(k)| \ [\mathrm{g}] \qquad (1)$$

$(*)$ denotes the section, $f(k)$ denotes the measured grasping force, and $f_d(k)$ denotes the target grasping force. $T_{(*)}$ denotes the time taken for the section, $N_{(*)}$ denotes the number of data for the section. T_{IC1} is 6s, T_{IC2} is 9 s, and T for each of the CC and EC sections is 5 s. Because the measurement sampling time is 0.1 s, N_{IC1} is 60, N_{IC2} is 90, and N for each of the CC and EC sections is 50. However, if the bird hits the cloud, $T_{(*)}$ and $N_{(*)}$ is extended for the section. This means that the time of collision with the cloud is included in the calculation. The target grasping force at the start of the collision continues until the end of the collision. The smaller the *AGF*, the higher is the AGF ability.

Table 2. Subject information.

Sub ject	Age	Sex	Disease	MABC-2 Manual Dexterity
1	6	Male	DCD, and at risk of Dysarthria	6
2	5	Male	DCD	6
3	7	Male	DCD, ASD, and ADHD	10
4	9	Male	At risk of DCD	5

4. Experiments and Results

4.1. *Subjects*

The four patients are listed in Table 2. All patients were diagnosed with DCD or at a risk of developing DCD. iWakka was used with the dominant hand.

4.2. *Characteristics of hand dexterity*

Figure 3 shows the AGF for each section. Comparing CC and EC, Subjects 1 and 4 had lower ability at EC than at CC. All AGF in the EC sections of Subject 1 were greater than all AGF in the CC sections. As for Subject 4, AGF_{EC2}, AGF_{EC3}, and AGF_{EC4} were greater than AGF_{CC3}, AGF_{CC2}, and AGF_{CC1} respectively, at the same grasping ranges. On the other hand, Subjects 2 and 3 had lower ability at CC than at EC. AGF_{CC2} and AGF_{CC3} of Subject 2 and AGF_{CC1}, AGF_{CC2}, and AGF_{CC3} of Subject 3 were relatively high. Furthermore, AGF_{IC2} of Subject 1 was greater than AGF_{IC1}, indicating that he had a low ability

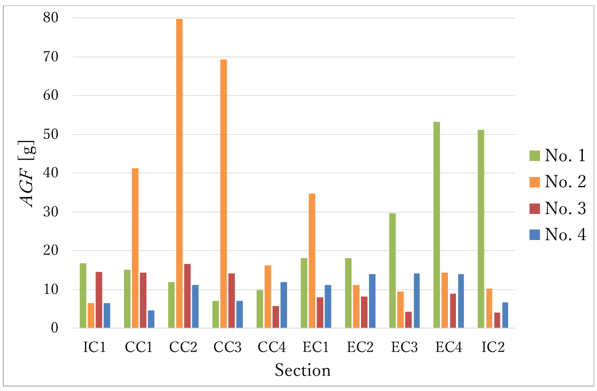


Fig. 3. *AGF* of each section.

to hold at 150 g after decreasing the grasping force. In addition, the AGF_{EC1} of Subject 2 was greater than AGF_{CC4} and AGF_{EC2}, indicating that the ability to switch from gripping to releasing force was low.

The comparison of CC and EC showed the difference in the ability of each muscle activity, and the comparison of neighboring sections showed the ability to switch between muscle activities.

5. Conclusion

We observed that the iWakka Game was applicable not only to autistic patients but also to children with DCD, aged 5–9 years, by using the small-sized grasping device, Wakka, and appropriate setting parameters. Moreover, the proposed evaluation method enabled the extraction of the characteristics of hand dexterity for each child. They had different characteristics of holding, gripping, and releasing in each range of grasping force that they were not good at. Improving each ability can enhance hand movements necessary for daily life. This result leads therapists to plan more effective training for hand dexterity.

References

1. Y. Morita, S. Nakajima, N. Sato, K. Yasaki, “Quantitative Evaluation System of Adjustability for Grasping Force”, Journal of the Japan Society for Welfare Engineering, Vol. 18, Subject 2, pp. 5-10, 2016. (in Japanese)
2. Y. Morita, K. Ando, M. Nomura, K. Toshima, T. Tamaru, “Verification of Usefulness of Testing/Training Device of Adjustability for Grasping Force on Hemiplegic Patients after Stroke”, Transactions of the Society of Instrument and Control Engineers, Vol. 55, Subject 1, pp. 78–84, 2018. (in Japanese)
3. M. Nomura, N. Kucharek, I. Zubrycki, G. Granosik, Y. Morita, “Adjustability for Grasping Force of Patients with Autism by iWakka: A Pilot Study”, 2019 12th International Workshop on Robot Motion and Control (RoMoCo), pp. 50-55, 2019.
4. Dagmar F.A.A. Derikx, Marina M. Schoemaker, “The Nature of Coordination and Control Problems in Children with Developmental Coordination Disorder during Ball Catching: A Systematic Review”, Human Movement Science, Vol. 74, p. 102688, 2020.

Authors Introduction

Mr. Masakazu Nomura

He received his Master's degree from the Department of Electrical and Mechanical Engineering, Graduate School of Engineering, Nagoya Institute of Technology, Japan in 2020. He is currently a Doctoral course student in Nagoya Institute of Technology, Japan.

Ms. Moe Nishiya

She received her bachelor’s degree from the Department of Electrical and Mechanical Engineering, Nagoya Institute of Technology, Japan in 2020. She is currently a master course student in Nagoya Institute of Technology, Japan.

Dr. Yoshifumi Morita

He is a Professor of Department of Electrical and Mechanical Engineering, Graduate School of Engineering, Nagoya Institute of Technology, Japan. His research interests are robot motion control, support control for human-machine cooperative system, and rehabilitation support robot.

Mr. Hideo Yamagiwa

He is a Pediatric Occupational Therapist at Tokyo Metropolitan Tobu Medical Center, Japan. He graduated from Chiba Medical Technology College, Japan.

Biofouling Monitoring Experiments of Underwater Concrete Samples for Offshore Platform Cleaning Robot Development

Keisuke Watanabe
Department of Navigation and Ocean Engineering, Tokai University, 3-20-1 Orido Shimizu-ku Shizuoka, 424-8610, Japan

Hiroki Goda
Department of Civil Engineering and Architect, Kyusyu Institute of Technology, 1-1 Sensuicho, Tobata-ku Kitakyushu, Fukuoka, 804-8550, Japan

Koji Harada
Technical Research Institute, Nishimatu Construction Co.Ltd., Address Aikawa, Kanagawa, 243-0303, Japan
E-mail: keisuke@tsc.u-tokai.ac.jp, goda.hiroki220@mail.kyutech.jp

Abstract

As the Japanese government decided to boost the carbon neutral power source development, offshore wind farm projects are emerging in Japan and hundreds of platforms will be constructed in the near future. Some of these platforms are possible to be floating structures and made of concrete-like material whose biofouling should be limited from the viewpoint of drag force reduction. An autonomous cleaning robot is one of possible solutions to minimize the effect of biofouling. To develop the cleaning device, we started field experiments to study biofouling process. In this paper, we introduce some results of biofouling monitoring experiments using several different concrete-like samples at sea.

Keywords: Offshore Power Plant, Biofouling, Unmanned Maintenance, Concrete Sea Exposure Experiment.

1. Introduction

Offshore wind power projects are becoming more active in the waters near Japan. There are fixed and floating options for offshore platform types. In the case of a floating platform, biofouling may cause an increase in fluid resistance and weight in the floating part, and it is considered desirable to remove it. On a fixed platform, the impacts described above are expected to be relatively small, but from a structural maintenance standpoint, it may be desired to remove marine biofouling. Subsea structures such as subsea mining is also desired to be cleaned in the deep water to maintain its mechanical functions[1]. It is expected that waves will be high in the waters near Japan that are suitable for offshore wind power generation. In such areas, the removal work performed by divers is not only dangerous, but also limits the number of days that can be worked. Therefore, it is desirable to develop a robot that can automatically clean underwater structures such as columns and foundations.

The columns of the structure are currently made of steel or concrete. The steel columns are coated to reduce biofouling. In the development of a cleaning robot, it is necessary to develop a cleaning device so that this coating will not be damaged. On the other hand, in the concrete column, the concrete is exposed to seawater. It has been reported that in conventional harbor structures, biofouling developed to cover the surface of concrete has a positive effect on the life of concrete because it plays a role in preventing seawater components from penetrating into concrete[2].

This research started to investigate a method for removing biofouling in concrete structures, unlike the conventional research that actively promotes biofouling. We envision cases where it is desirable to reduce fluid resistance, such as floating platforms and cooling water pipes in power plants. From the standpoint of eliminating biofouling, the authors think that it is necessary to consider the following three points.

(i) Development of new materials and surface treatment methods that prevent the penetration of seawater components
(ii) Understanding the process of development of biofouling on concrete surfaces
(iii) Removal method by cleaning robot

In order to proceed with the study on these three points, the authors started an experiment to observe the state of biofouling using different types of concrete test samples[3]. In this paper, we will introduce the method of the experiment on our pier, some experimental results, and some aspects in developing the cleaning robot that we are studying at this stage. The experiment is being conducted at the pier of School of Marine Science and Technology, Tokai University in Shimizu port.

2. Some Aspects to be Considered in Developing the Cleaning Robot

The robot is put into the sea from a support ship or a platform. The robot's power is either built-in battery of the required capacity or supplied externally through an umbilical cable. The housing of the robot can be deformed according to the curve of the column to be cleaned, and the robot housing is designed so that cleaning tools such as brushes can always stick to the column. Regarding cleaning equipment consideration, we will proceed its design with the examination while clarifying how the species and biomass change over time through ocean exposure experiments. As concrete is a brittle material, and at the same time, attached organisms such as barnacles can chemically integrate with the concrete material, so forcibly peeling off barnacles can damage the concrete surface. Therefore, it is necessary for us to devise a method different from the cleaning equipment that is being studied for steel materials such as ship-hull cleaning robots.

Table 1 Basic System Configuration of the Robot

Basic System Configuration
Thrusters and Actuators
Cleaning Device
Underwater Camera with Lights
Underwater Position Detection System
Inertia Measurement Unit
Communication Interface
Removed Debris Recovery System
Control System
Power Source

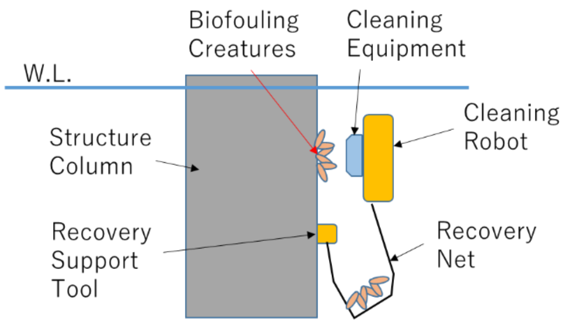


Fig. 1. Conceptual Schematic View of the Cleaning Robot.

In order to reduce the impact on the marine environment, the removed attached organisms must be recovered by robots rather than released into the sea. The mechanism of the cleaning equipment and the recovery device needs to be designed to function as one system, and in order to optimize these, it is necessary to know the formation process of biofouling of concrete material through actual sea experiments.

Table 1 shows the contents of the basic system configuration of the cleaning robot currently planned. Figure 1 shows a schematic diagram of the concept of robot operation. In order to automatically control the robot, it must be possible to detect the underwater position, speed, and attitude angles of the robot. A communication system with the land base is required to capture the cleaning status with an underwater camera and monitor it at the land base. The robot's movement and trajectory tracking control are performed by thrusters, and actuators that operate cleaning equipment are also required. The cleaning operation is performed automatically while following a given trajectory.

3. Biofouling Experiment at Pier

The sea exposure experiment is being conducted at the pier of the School of Marine Science and Technology, Tokai University. As shown in Fig. 2, different types of concrete test pieces are hung. Since it is considered that many attached organisms such as barnacles develop near the water surface, the case near the water surface as shown in Zone B in the figure and the case near the seabed as shown in Zone C were set for comparison. Figure 3 shows how the test pieces are arranged in the pier in Zone B. The height of a piece is adjusted around the mean sea level and it appears above the sea surface when the tide is low, on the other hand, it submerges beneath the sea surface when the tide is high. The pieces in Zone C are always fully submerged near the seabed.

A ring-shaped pile was embedded during concrete molding to hang the test piece. The size of the test piece is about 0.1 m in length, about 0.1 m in width, about 0.2 m in height, and weighs about 49 N. The weight of each test piece was precisely measured as shown in Fig.6. By accurately measuring the weight over time, the amount of attached biomass is estimated.

4. Some Experimental Results

The test piece was first installed on August 26, 2021, but since the metal fittings that hang the test piece corroded and fell, the test piece was newly installed on September 21, 2021, except for the two test pieces. Figure 4 shows a test piece appearance in Zone B before the installation. Figure 5 shows the development of biofouling on the test piece on October 20, 2021, about one month later. Figure 6 shows further development of biofouling of the test piece on November 15, 2021, about three months later.

As shown from Fig. 4 to Fig. 6, barnacles mainly grew on the concrete surface in Zone B. After 1 month, the concrete surface was visible in some places, but after 2 months, the entire surface of the concrete test piece was thickly covered. Figure 7 shows the mass increase rate of the test piece. In the most increased case, a 10% (about 5N) increase was seen in about 80 days. That is, it was suggested that it leads to a non-negligible weight increase in floating offshore structures.

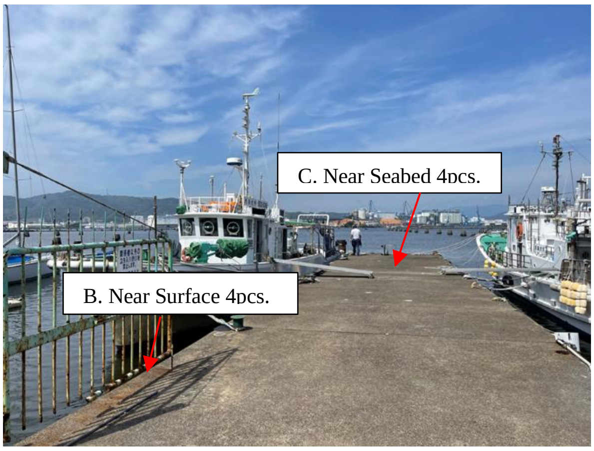


Fig. 2. Zoning of Test Pieces Arrangement at Pier.

Fig. 3. Near Surface Condition in Zone B.

Fig. 4. A Test Piece Appearance before Installation.

5. Conclusion

In this article, we introduced the concept of a robot that automatically removes marine-attached organisms, assuming a concrete column of a floating offshore structure. In order to find effective solution of cleaning

Fig.5 A Test Piece Appearance after One Month.

Fig.6 A Test Piece Appearance after Three Months.

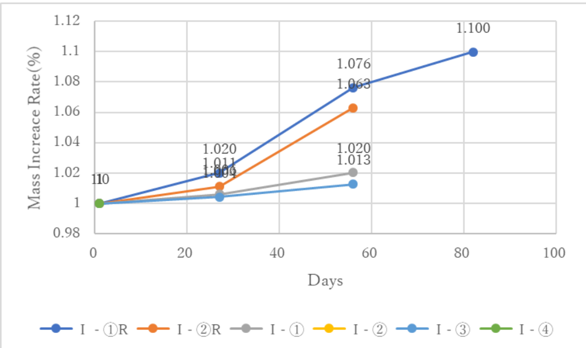


Fig.7 Mass Increase Rate of Test Pieces.

equipment of the robot, we started sea exposure experiment of concrete test pieces. The experimental setups and some early stage results were presented. Although further investigation is needed, the result implies biofouling is possible to influence the structure dynamics and its removal might be meaningful.

Acknowledgements

The authors would like to thank Mr. R. Ino of Kyusyu Institute of Technology for helping the field experiments. This work was supported by JSPS KAKENHI Grant Number 18K04585.

References

1. Y. Watanabe, K. Watanabe, H. Suzuki, "Research and Development on a Self-Walking Vertical Mining System Using DTH for Seafloor Mining and Sampling", Proc. of OMAE2019, 2019
2. T. Maruya, M. Iwanami, "Durability enhancement of RC structures covered with a dense layer formed by marine aquatic fouling organisms", Journal of the Japan Society of Civil Engineers, No.739,pp.61-75,2003(in Japanese)
3. H. Goda, K. Harada, K. Takasu, H. Suyama, M. Hibino, "Flow Characteristics and Strength Characteristics of Geopolymers Using Fly Ash Modified by the Flotation Method", The 9th International Conference of Asian Concrete Federation, 2021

Authors Introduction

Dr. Keisuke Watanabe

He is a Professor of School of Marine Science and Technology, Tokai University in Japan. He received his B.E., M.E. and ph.D. degrees from the Dept. of Naval Architecture and Ocean Engineering, The University of Tokyo, Japan in 1991,1993 and 1996. His research interest is autonomous marine vehicles application.

Dr. Hiroki Goda

He is an Associate Professor of Dept. of Civil Engineering and Architect, Kyushu Institute of Technology in Japan. He received his B.E., M.E. and ph.D. degrees from the Dept. of Civil Engineering, Kyushu University, Japan in 1999, 2001 and 2010. His research interest is sustainable constructive material like geopolymer.

Dr. Koji Harada

He is a Senior Reserher of the Technical Research Institute, Nishimatu Construction Co.Ltd. His research interest is maintenance engineering, impact engineering, and materials engineering. He started esearch on geopolymers in the construction field in Japan.

Fall Risk Notification System using LiDAR Sensor for the Visually Impaired People

Daigo Katayama*, Kazuo Ishii, Shinsuke Yasukawa
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu-ku, Kitakyushu-shi, Fukuoka, 808-0135, Japan

Satoshi Nakadomari, Koichi Wada, Akane Befu, Chikako Yamada
NEXT VISION, Kobe Eye Center 2F, 2-1-8, Minatojima, Minamimachi,
Chuo-ku, Kobe-shi, Hyogo, 650-0047, Japan

E-mail: katayama.daigo696@mail.kyutech.jp, ishii@brain.kyutech.ac.jp, s-yasukawa@brain.kyutech.ac.jp, satoshi.nakadomari@riken.jp, wakichi@nextvision.or.jp, befu@nextvision.or.jp, yamada@nextvision.or.jp.

www.lsse.kyutech.ac.jp

Abstract

We have developed the fall risk notification system using LiDAR sensors to reduce number of fall accidents on platform involving visually impaired people. In this paper, we report the experiment results of the environment recognition algorithm for the fall risk notification system. In this algorithm, height grid map is generated from the depth image from LiDAR sensor and the posture of iPhone. In the experiment, we evaluated the accuracy and responsivity when approaching risky area of falling, such as stairs.

Keywords: Electronic Travel Aids, Negative Obstacle Detection, Point Cloud, LiDAR Sensor.

1. Introduction

Fall Accidents from station platforms are a serious problem in the daily mobility of the visually impaired people. According to statistics[1] from the Ministry of Land, Infrastructure, Transport and Tourism, an average of 76 fall accidents occurs every year, indicating that the number of fall accidents is not small.

To cope with these problems, research and development of devices called electronic walking aids, which convey sensor information in a non-visual way, has been promoted. In recent, smartphones have become a main part of aids, and some have been developed to help people get in line[2] and navigate to their destinations[3]. However, there are few examples of applications that notify the user of the fall risk.

The overview of the electronic walking aid developed in this study is shown in Fig. 1. iPhone 12 Pro processes information about the surrounding environment, and when a warning is needed, it sends a command to the control unit, which vibrates the vibrating device and transmits the information to the user. In this paper, we evaluate the performance of the fall risk notification system using the depth information from LiDAR sensor on the iPhone 12 Pro when the user approaches a risky area of falling.

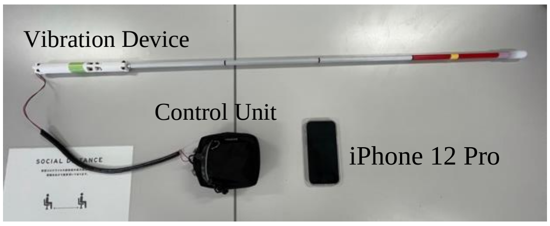


Fig. 1 Overview of the electronic travel aid

2. Methods

2.1. *System Usage Conditions*

This system assumes the usage conditions that the iPhone is tilted by an angle θ and mounted in front of the chest as shown in Fig. 2. Assuming that the height to the user's chest is H_{chest} and the length from the bottom of the iPhone to the camera unit is l_{camera}, the height from the iPhone to the road surface H_{road} is calculated by Eq. (1).

$$H_{\mathrm{road}} = H_{\mathrm{chest}} + l_{\mathrm{camera}} \cos\theta \qquad (1)$$

2.2. *Algorithm of Fall Risk Estimation*

This algorithm calculates the risk of falling, the overview shown in Fig. 3. This risk calculates from the minimum distance from the user to a risky area, such as a stair or a gap, and the occupancy of risky area in the user's close range. The details of each step are shown below.

2.2.1. *Conversion from Depth Image to Point Cloud*

3D points $\boldsymbol{p}_{u,v}$ calculate from the depth $d_{u,v}$ obtained from the LiDAR sensor (u, v are coordinates on the depth image) and the intrinsics of the LiDAR sensor (f_u, f_v, c_u, c_v) as shown in Eq. (2).

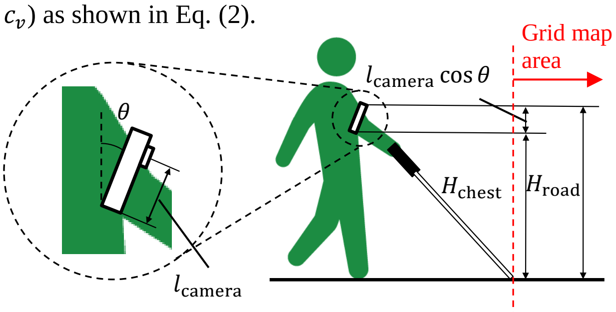


Fig. 2 Mounting position of iPhone

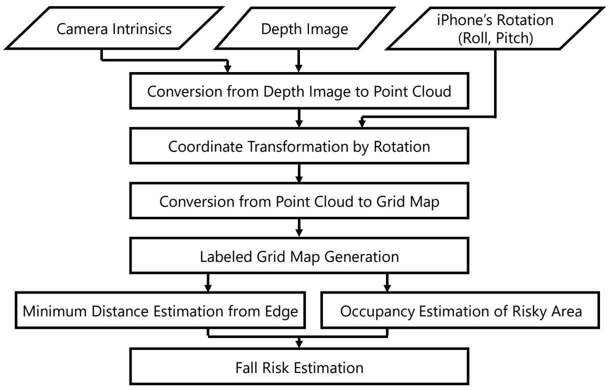


Fig. 3 Overview of fall risk estimation algorithm

$$\boldsymbol{p}_{u,v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} d_{u,v} \\ \frac{u-c_u}{f_u} \\ -\frac{v-c_v}{f_v} \end{bmatrix} \qquad (2)$$

2.2.2. *Coordinate Transformation by Rotation*

This method does the coordinate transformation for the points $\boldsymbol{p}$ (camera coordinates) to $\boldsymbol{P}$, the axis indicating the height direction (Z-axis) is in the reverse direction of the gravity direction. this coordinate transformation as shown in Eq. (3) use the quaternion $\boldsymbol{q}$, which is calculated from the Roll angle and Pitch angle of the iPhone. So that the point $\boldsymbol{p}_{u,v}$ and $\boldsymbol{P}_{u,v}$ is set to $\boldsymbol{p}_{u,v}$ = [0, x, y, z], $\boldsymbol{P}_{u,v}$ = [0, X, Y, Z] in Eq. (3).

$$\boldsymbol{P}_{u,v} = \boldsymbol{q}\boldsymbol{p}_{u,v}\boldsymbol{q}^{-1} \qquad (3)$$

2.2.3. *Conversion from Point Cloud to Grid Map*

The points $\boldsymbol{P}$ is transformed into the height grid map based on the coordinates in Section 2.1.2 due to reduce the noise in the points and simplify the calculation. The grid stores the points which their X-axis and Y-axis position is inside of the grid range. The height $h_{X,Y}$ (X and Y are coordinates on the grid map) in each grid is calculated from the height (Z-axis position) of the stored point $\boldsymbol{P}$ and the number of stored points n as shown in Ep. (4).

$$h_{X,Y} = \frac{\sum_{k=1}^{n} Z_k}{n} \qquad (4)$$

2.2.4. *Labeled Grid Map Generation*

This method generates the labeled grid map by threshold process in order to each grid classify safe and risky areas as shown in Eq. (5). $L_{X,Y}$ is the label of each grid. h_{th} is tolerance of threshold.

$$L_{X,Y} = \begin{cases} Safe\ Area & if\ h_{X,Y} \geq H_{\mathrm{road}} - h_{\mathrm{th}}, \\ Risky\ Area & if\ h_{X,Y} < H_{\mathrm{road}} - h_{\mathrm{th}}. \end{cases} \qquad (5)$$

2.2.5. *Minimum Distance Estimation from Edge*

In this method, the labels of the grid are examined along the depth direction (X-axis direction in the grid map) in the grid map, and the edge is defined as the point where the label changes from safe area to risky area. The minimum distance from the edge is calculated by Eq. (6) to the row with the least number of grids to the edge. n_{min} is the least number of grids to the edge, ΔD is the grid width.

$$D_{min} = \Delta D \cdot n_{min} \qquad (6)$$

2.2.6. *Occupancy Estimation of Risky Area*

This method calculates the occupancy of the risky area O_{risky} and the occupancy of the safe area O_{safe} in the user's close range by Eq. (7) and (8). n_{risky} and n_{safe} are the number of grids labeled risky area and safe area in closed range. n_{close} is number of grids in closed range.

$$O_{risky} = n_{risky} \,/\, n_{close} \quad (7)$$

$$O_{safe} = n_{safe} \,/\, n_{close} \quad (8)$$

2.2.7. *Fall Risk Estimation*

The fall risk level is calculated by the minimum distance from the edge and the occupancy of the risky area. The level is classified into three levels as shown below. D_{close} and D_{far} are distance thresholds.

- High level: $D_{min} < D_{close}$ or $O_{risky} > O_{safe}$
- Medium level: $D_{min} > D_{close}$ and $O_{risky} < O_{safe}$
- Low level: $D_{min} > D_{far}$ and $O_{risky} < O_{safe}$

In medium level, add some distance thresholds to be able to notify finely when the minimum distance is updated.

2.3. *Experiment on Risky Area of Fall*

We did the experiment of the system when entering an area where there is a risk of falling. This experiment was done for evaluation the accuracy of the estimation of the minimum distance and the responsivity of the notification. The details are as follows.

2.3.1. *Procedure*

The task of this experiment was for the user to wear the device and walked straight to the goal as shown in Fig. 4. The conditions are as follows.

(i) Approach angle: 90[deg]
(ii) Approach angle: 45[deg]

Condition (i) was assumed that users are getting on and off the train, and condition (ii) was assumed that user are moving along the platform and approaching the edge of the platform.

The data to be collected during the experiment were depth images, attitude and self-position from iPhone.

2.3.2. *Device Setting*

The iPhone was fixed to the front of the user's chest with a harness-type fixture. The iPhone was also tilted forward 30[deg] on the fixture.

2.3.3. *Experiment Place*

This experiment was done on the outdoor at the Wakamatsu campus of Kyushu Institute of Technology. As shown in Fig. 5, this place had a flat surface with a wide stair (Height of step: 0.14[m], Depth of step: 0.34[m]).

2.3.4. *Analysis Method*

The analysis for evaluation were done on the accuracy of the minimum distance and the responsivity of the fall risk notification. For the accuracy of the minimum distance, the error was calculated as the difference between the distance between the start point and the goal point and the distance calculated from the iPhone's self-position and approach angle. The responsivity of the fall risk notification was analyzed by calculating the difference between the time it takes for each level of fall risk to be first presented and calculated distance applied to the estimation of the fall risk level.

This analysis was done on another PC. The PC used for the analysis was MacBook Pro (Made by Apple Inc., CPU: Intel Core i9 (2.3 [GHz], 8 Core), RAM: 16 [GB]), and the software was MATLAB R2021a.

In the analysis, H_{road} = 1.32[m], θ = 30[deg], D_{min} = 1.35[m], D_{far}= 2.35[m], h_{th} = 0.05[m]. The grid width was 0.05[m] in both the X and Y-axis directions, and 1 distance threshold was set 1.85[m].

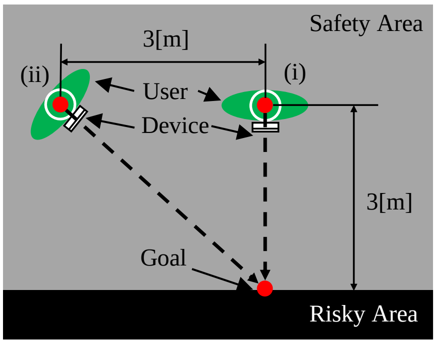


Fig. 5 Conditions of task

Fig. 4 Experiment place

3. Results

3.1. *Accuracy of Minimum Distance from Edge*

The error average of the minimum distance on each condition as shown in Table 1. When approach angle θ_a was 90 [deg], the average errors were less than 5 [%], but when θ_a = 45 [deg], the errors were around 15 [%].

3.2. *Responsivity of Fall Risk Notification*

The difference average in time at the first notification of each fall risk level as shown in Table 2. When the approach angle was $\theta_a = 90$[deg], the differences in time were within the range of ±0.04[s], but when $\theta_a = 45$[deg], the differences were calculated to be around 0.4[s].

4. Discussion

4.1. *Effect of Approach Angle*

In this experiment and analysis, they were shown that the minimum distance from the edge and the time to the first notification of each fall risk level both tended to worsen when the approach angle $\theta_a = 45$[deg].

This cause seems different grid position of the minimum distance from the ideal one when the edges on the grid map do not have a straight-line shape as shown in Fig. 6. This cause effect to the error in the minimum distance by the amount of the grid width, so it is necessary to consider the allowable error including the grid width.

The current algorithm focuses on a single grid point closest to the edge. Therefore, the value of the minimum distance can be unstable, depending on the shape of the edge. The unstable notification can confuse the user[4], we should consider segmentation method such as line fitting for the edge.

Table 1. The error average of the minimum distance

Error average [%]			
Times	1st.	2nd.	3rd.
$\theta_a = 90$ [deg]	5.01	2.20	4.24
$\theta_a = 45$ [deg]	15.0	15.6	15.4

Table 2. The difference in time at the first notification

Difference average in time [s]			
Times	1st.	2nd.	3rd.
$\theta_a = 90$ [deg]	0.034	-0.039	0.028
$\theta_a = 45$ [deg]	0.418	0.445	0.423

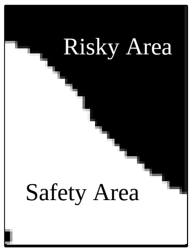


Fig. 6 The shape of the edge on grid map
(In case of $\boldsymbol{\theta_a} = 45$ [deg], 1st.)

4.2. *Challenges in Use by the Visually Impaired*

Based on the results of this experiment, there are two specific issues that need to be considered when considering the actual use of this system by visually impaired people.

The first issue is to stabilize notifications. As mentioned in the previous section, the current algorithm causes unstable notifications when approaching at an angle, i.e., when walking on the platform and approaching the edge of the platform, which may lead to confusion for the user. This issue will be considered in line with the aforementioned.

The second issue is the timing of the notification. As can be seen from Table 2, the timing of notification was found to be later than the ideal one in most cases. Considering the delays caused by the communication with the vibration device and the reaction speed of the user, it is necessary to set a margin for the timing of notification considering the above matters.

5. Conclusion

In this paper, we evaluated an algorithm for the fall risk notification system that visually impaired people of their risk of falling using depth information from LiDAR sensor on iPhone 12 Pro. We mainly evaluated the accuracy of the minimum distance and the responsivity of the notification and found that the above two parameters deteriorated when approaching diagonally to the edge.

In the future, we will try to solve the above problems, and stabilize of notification and implement the notification timing added safety margin.

Acknowledgements

This research was supported by Research Grant of JR–West Relief Foundation in 2020 (Grant No. 20R032).

References

1. MLIT, “Study Group for Improving Safety at Station Platforms”, https://www.mlit.go.jp/tetudo/tetudo_fr7_000015.html. (In Japanese)
2. M. Kuribayashi et al., “LineChaser: A Smartphone-Based Navigation System for Blind People to Stand in Lines”, Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, no. 33, pp. 1-13, 2021.
3. M. C. Rodriguez-Sanchez et al., “Accessible smartphones for blind users: A case study for a wayfinding system”, Expert Systems with Applications, vol. 41, issue 16, pp. 7210-7222, 2014.
4. M. Tauchi, T. Ohkura, “States and Problems of Assistive Technology for the Visually Impaired: Walking Alone”, Journal of the Society of Instrument and Control Engineers, vol. 34, no.2, pp. 140-146, 1995. (In Japanese)

Authors Introduction

Mr. Daigo Katayama

He is a Doctoral Student at Graduate School of Life Science and Systems Engineering, Kyushu Institute of University. He received Master’s degree from Graduate School of Life Science and Systems Engineering, Kyushu Institute of University.

Prof. Kazuo Ishii

He is a Professor at Graduate School of Life Science and Systems Engineering, Kyushu Institute of University. He received Ph.D. from School of Engineering, The University of Tokyo.

Dr. Shinsuke Yasukawa

He is an Associate Professor at Graduate School of Life Science and Systems Engineering, Kyushu Institute of University. He received Ph.D. from Graduate School of Engineering, Osaka University.

Dr. Satoshi Nakadomari

He is a Director at NEXT VISION. He graduated from the Department of Medical, The Jikei University. He received Ph.D. from the Department of Medical, The Jikei University.

Mr. Koichi Wada

He is a Vision Park Information Master at NEXT VISION. He graduated from Acupuncture and Physical Therapy Teacher Training School of University of Tsukuba.

Ms. Akane Befu

She is a Vision Park Information Concierge at NEXT VISION. She graduated from The Department of Domestic Science, Tokushima Bunri University.

Ms. Chikako Yamada

She is an Executive Secretary at NEXT VISION. She graduated Graduate School of Sociology, Bukkyo University. She received Master’s degree from Graduate School of Sociology, Bukkyo University.

Reflection Coefficient Estimation through the Modelling of Ultrasonic Transmission

Ryuugo Mochizuki*
Center for Socio-Robotic Synthesis, Kyushu Institute of Technology, 2-4 Hibikino Wakamatsuku Kitakyushu, 808-0196/Fukuoka, Japan

Yuya Nishida, Kazuo Ishii
Department of Human Intelligence Systems, Kyushu Institute of Technology, 2-4 Hibikino Wakamatsuku Kitakyushu,808-0196/Fukuoka, Japan
E-mail: mochizuki.ryuugo126@mail.kyutech.jp, y-nishida@brain.kyutech.ac.jp,ishii@brain.kyutech.ac.jp
www.lsse.kyutech.ac.jp

(leave space here should be 4 lines between *www.university_name.edu* and **Abstract***)*

Abstract

In food industry, shortage of workers is one of a serious problem. Automation of food handling is a critical isSsue nowadays. To alleviate the damage caused by food picking operation by robotic hand, we propose non-contact acoustic impedance estimation with ultrasonic wave, which should be preceded before the picking for optimization of grasp stiffness. We have the assumption of the correlation between hardness and acoustic impedance. The impedance is calculated by the product of sonic velocity and density of a medium. From the point of view, the harder the medium is, the larger the impedance should be.

We built up ultrasonic transmission model considering attenuation by ultrasonic reflection and absorption, then, made an experiment to estimate internal reflection of samples with two overwrapped media of different acoustic impedance.

Keywords: Acoustic Impedance, Ultrasonic Transmission, Hardness Estimation, Internal Reflection, Absorption

1. Introduction

In Japan, employment opportunities to both sexes have steadily increased up to now. According to the report[1], number of housewife household has been decreased by nearly two million, while, dual-income households have been increased by the same amount. Single households have been increased (for elderly, about three million, for non-elderly, about 1.5 million) for 20 years recently[2]. As the market of prepared dishes, the magnitude has rose up by seven times from 1985 to 2016[3]. From the point of view, the relevance can be assumed between the demand for prepared dishes and increase of single or dual-income households. On the other hand, in food industry, the jobs-to-applicants ratio has rose up from 1.48 to 2.78 for five years from 2013[4], whose ratio is higher than other industry, little applicants are assumed to be willing to work for the industry. If the increase of single or dual-income households continues, the supply of prepared dishes may not be enough, thus, automation of food process and packing is one of a crucial task.

As one of the task in the automation, we aim to realize safe food picking (without damaging food) by a robot hand. To cope with the task, we propose to introduce non-contact hardness estimation with ultrasonic wave to optimize grasp stiffness, which precedes to the picking. We estimated reflection coefficient by overwrapping two samples with different acoustic impedance, considering the value as hardness.

2. Related Work

Machado et. al. have analyzed reflected ultrasonic wave to evaluate wood strength[5]. Cho et al. have measured depth of cheeses with 99.98 % of accuracy, furthermore, succeeded in observe more than 0.9 of correlation between sonic velocity and a mechanical property of the cheese[6]. Satyam et al. estimated varying hardness of tomatoes over a time period from fresh to full ripen cycle of the crops[7].

However, to the best of our knowledge, no report of hardness estimation considering heterogeneity of hardness has been found. The heterogeneity usually orients from uneven heating of food during cooking process, thus, we should consider the case of soft interior with hard surface.

3. Proposed Method

3.1. *Outline*

As the preparation, we estimate attenuation coefficient in the air α_0 and sound pressure of emitted wave P_0, then obtain and treat emitted wave $\psi_{t,0}(t)$ in similar way as basis function for wavelet transform[8]. To analyze received wave, we consider arbitral samples as multiply overlapped media as Fig. 1 shows (z_k denotes the surface of medium k), for this reason, received wave $p_r(t)$ can be regarded as summation of reflected waves $p_{r,k}(t)$ at different surfaces of media. $p_{r,k}(t)$ is expressed in two ways, that is modelled reflection $\hat{p}_{r,k}(t)$, and reflection $\psi_{r,k}(t)$ based on $\psi_{t,0}(t)$. The former expresses relationship between reflection coefficient r_k at the surface of medium k and attenuation coefficient α_k[Np/m] inside medium k based on amplitude, however, waveform information is not considered The latter can express waveform because $\psi_{t,0}(t)$ is stretched or compressed in two direction (time and amplitude) to fit $p_{r,k}(t)$ in this stage. The analyzation process of $p_r(t)$ is shown as Fig. 2. First, the amplitude of $p_r(t)$ is judged through the thresholding by fixed value T_P. If the judgement is true, $\psi_{t,0}(t)$ slides in the time axis direction, then stretched or compressed in the ratio of A_k vertically and B_k horizontally so that the correlation between $\psi_{t,0}(t)$ and $p_r(t)$ is maximized, then, $\psi_{r,k}(t)$ can be

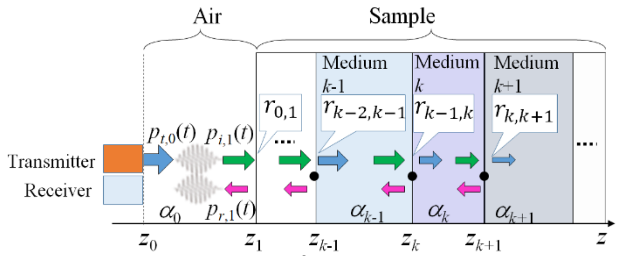


Fig. 1 Assumption of Sample Composition

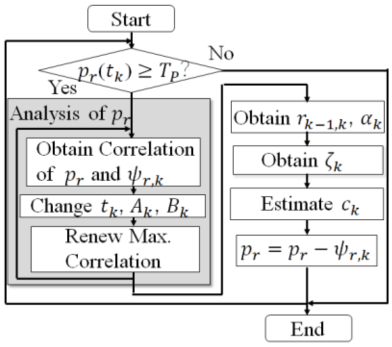


Fig. 2 Process of Analyzing p_r

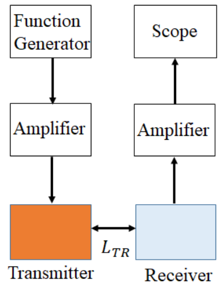


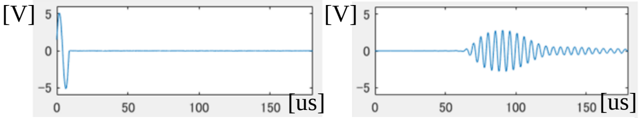
Fig. 3 Calibration Environment

obtained. We assume the amplitude of reflection is larger at closer area of sample surface, then, we limit searching area of the maximization of the correlation to $t_k > t_{k-1}$(Here, t_k denotes estimated time of the reception of $p_{r,k}(t)$). At this stage, t_k should overwrap the maximum amplitude of $\psi_{t,0}(t)$.

In the estimation of $\hat{p}_{r,k}(t)$, α_k and r_k are solved so that the squared error between $\hat{p}_{r,k}(t)$ and $\psi_{r,k}(t)$ is minimized. $\hat{p}_{r,k}(t)$ has another unknown parameter c_k [m/s] (sonic velocity in medium k), which is obtained through the fitting of linear function estimated from the relation between sonic velocity and acoustic impedance referring to a database. Hereafter, the estimation of α_k and $r_{k-1,k}$ proceeds with iterative subtraction by $\psi_{r,k}(t)$ from $p_r(t)$, which ends when the amplitude of p_k is less than T_P.

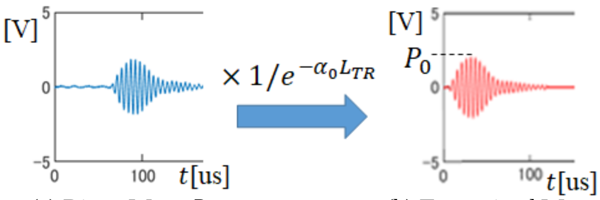
3.2. *Calibration of attenuation coefficient and sound pressure of the emission*

Sound Pressure P_o cannot be estimated in [Pa], then we assume that the pressure is proportional to output voltage [V] of the receiver. We obtain P_o [V] and α_o[Np/m] as the calibration process under the condition (Fig. 3) in which transmitter and receiver are aligned face to face. Function Generator generates one cycle of sign wave (Fig.4(a)) to the transmitter and direct wave is received



(a)Input of Function Generator (b)Amplified Direct Wave

Fig. 4 Transmitted Wave



(a) Direct Wave P_D (b) Transmitted Wave (basis function) $\psi_{i,0}$

Fig. 5 Transmitted Wave

as Fig.4(b) shows. P_o and α_o can be calculated with the amplitudes of directed waves that are observed from a few different distances L_{TR} from the transmitter. When P_o and α_o are obtained, as Fig. 5 shows, cancellation of the attenuation by α_o results in the solution of $\psi_{t,0}(t)$.

3.3. *Expression of the reflected wave*

$p_{r,k}(t)$ is expressed in two ways, based on transmission and attenuation model $\hat{p}_{r,k}(t)$ and $\psi_{r,k}(t)$, which is made from the compression of $\psi_{t,0}(t)$. As $\hat{p}_{r,k}(t)$, we consider the attenuation of ultrasonic wave originates from the reflection (The amplitude of the reflected wave is diminished to the ratio $r_{k-1,k}$ at location z $=z_k$) and transmission (The transmitted wave moves in velocity c_k[m/s] and the amplitude is diminished to the ratio of $e^{-c_k(z_{k+1}-z_k)}$ since its entry at $z=z_k$). After the emission of $p_{t,0}(t)$, the reflection and transmission is repeated 2k-1, 2k times, each other. Thus, $\hat{p}_{r,k}(t)$ is expressed in Eq. (1), (2).

$$\widehat{p_{r,k}}(t) = r_{k-1,k} C_{\Pi,k} \psi_{i,0}(t) e^{-2\alpha_{k-1}(z_k - z_{k-1})}. \quad (1)$$

$$C_{\Pi,k} = \prod_{k'=1}^{k-1} (1 - r_{k'-1,k'}{}^2) e^{-2\alpha_{k'-1}(z_{k'} - z_{k'-1})}. \quad (2)$$

However, waveform is not considered in $\hat{p}_{r,k}(t)$. Even if the Function Generator generates a cycle of sine wave, the transmitted wave is shown in a scope as a group of sine waves whose amplitude changes continuously and whose frequency is equal to the central frequency f_c[Hz] of the transducer (Fig. 5). We stretch $\psi_{t,0}(t)$ A_k times vertically and B_k times horizontally as Eq. (3) shows, in this way, reflected wave based on $\psi_{t,0}(t)$ $\psi_{r,k}(t)$ can be obtained.

$$\widehat{p_{r,k}}(t) = A_k \psi_{i,0}\left(\frac{t - t_0}{B_k}\right) \quad (3)$$

We assume maximum sound pressure is measured at $t = t_0$ in case that transmitter and receiver contacts each other, and the frequency of transmitted wave changes $1/B_k$ times in media k compared to the emission.

4. Experiment

As we mentioned in Chapter 1, hardness s is considered to be acoustic impedance ζ. Here, a medium k is regarded to have a uniform ζ_k then we treat a sample as a medium. For neighboring media k-1 and k, ζ_k have the relationship with r_k shown in Eq. (4) [10].

$$r_{k-1,k} = \frac{\zeta_k - \zeta_{k-1}}{\zeta_k + \zeta_{k-1}}. \quad (4)$$

From the point of view, large $r_{k-1,k}$ can be observed if ζ_k has large difference from ζ_{k-1} . The object of this experiment is to configure the increase of $r_{1,2}$ if ζ_2 increases. As well, ζ_2 and $r_{1,2}$ should have the same relation even if the depth of samples H_S (H_{S1}, H_{S2} for Sample 1, 2 each other) changes. Table 1 shows the pattern of sample combination, here we selected $H_{S1} = 29$[mm] for pattern 1, 2, and $H_{S2} = 35$[mm] for pattern 3, 4. We only consider positive reflection coefficient.

4.1. *Preparation*

We recorded amplitude P_D of direct wave using the environment shown in Fig. 3 and changing $L_{TR} = 10$, $20..,50[mm]$. We expect that P_D attenuates under extended L_{TR} following Eq. (5), with reference of $L_{TR} = L_{TR,1} (= 20[\text{mm}])$, then iteratively estimated α_0 and P_0 for 20 times.

$$P_D(L_{TR}) = P_D(L_{TR,1}) e^{-\alpha_0 (L_{TR} - L_{TR,1})}. \quad (5)$$

$$P_0 = P_D(L_{TR,1}) e^{\alpha_0 L_{TR,1}}. \quad (6)$$

4.2. *Environment*

Fig. 6 shows the environment for reflection estimation, and, Fig. 7 shows used samples. The true value of ζ and c is summarized in Table 1. ζ can be expressed in another way in Eq. (7), with the relationship between ζ and c, here, ρ[kg/m^3] denotes sample density[9].

Table 1 True Value of Sonic Speed c and Acoustic Impedance ζ (at Temperature 30[deg])

Sample Name	H_S [mm]	c [m/s]	ζ [Pa s/m]
Al	10	6420	17.3×10^6
Gel (s=0)	29	2237	2.5×10^6
	35	2233	2.4×10^6
Gel (s=7)	30	2185	2.5×10^6
	36	1643	1.8×10^6

Table 2 Setting for the Experiment

T_P	$P_0/100$		f_c	200[kHz]	
H_E	50[mm]		τ	±40V	(Calib.)21[deg]
θ_1	15[deg]				(Ref.)21[deg]
P_0	±40V	3.07[V]		±50V	(Calib.)21[deg]
	±50V	3.27[V]			(Ref.)24[deg]
α_0	±40V	2.43[Np/m]	a_1	471.1[m /(Pa s^2/m)]	
	±50V	1.52[Np/m]	a_2	349.8[m/s]	

Table 3 Combination of Samples

Pattern	Sample 1		Sample 2	
	Sample Name	H_{S1} [mm]	Sample Name	H_{S2} [mm]
1	Gel (s=0)	29	Gel (s=7)	30
2	Gel (s=0)	29	Al	10
3	Gel (s=0)	35	Gel (s=7)	36
4	Gel (s=0)	35	Al	10

$$\zeta = \rho c. \tag{7}$$

Urethane Gel is produced through the mixture and coagulation of two kinds of liquid. thus, H_S is not uniform. We measured H_S in the environment shown in Fig. 8, by sandwiching the sample by a transmitter and receiver, then subtracting L_2 from L_1, and, obtained true c by measuring elapsed time from wave emission to reception and dividing H_S with the time. As the true ζ, we used Eq. (7) to obtain. However, as ζ of aluminum, we referred to Ref. 11 because the media has extremely large ζ .As the calibration to decide α_0 and P_0 , we repeated for 20 times, and averaged the results of the estimations. Table 2 shows the setting of experiment(central frequency f_c , distance from transmitter and sample surface H_E, stilt of transducers θ_1, temperature τ), and estimation of α, P_0. The amplitude of the input to transmitter was ± 40, ± 50[V]. The coefficient a_1, a_2 of the linear function to find sonic

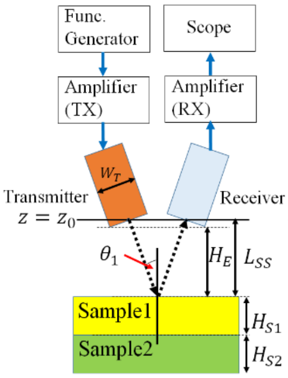
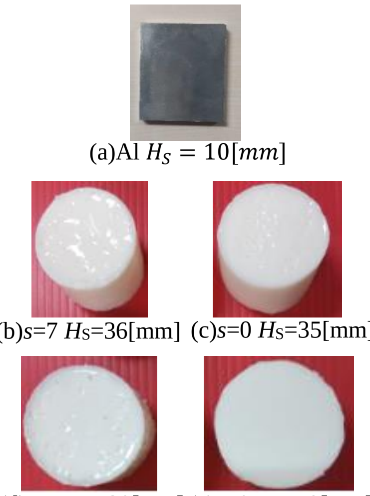


Fig. 6 Experimental Environment



(a)Al $H_S = 10[mm]$
(b)s=7 H_S=36[mm] (c)s=0 H_S=35[mm]
(d)s=7 H_S=30[mm] (e)s=0 H_S=29[mm]

Fig. 7 Samples for the Environment

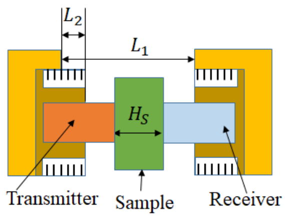


Fig. 8 Measurement of H_S, c, ζ

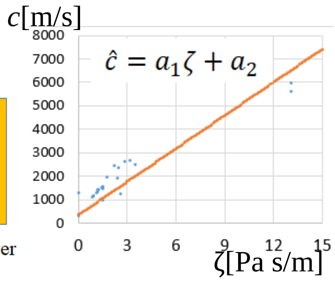


Fig. 9 Estimation of c with input of ζ

velocity is shown in the table as well. a_1 differs from ρ in that a_1 is regarded uniform, while ρ synchronously changes as c changes. Table 3 shows the patterns of the sample combination, that is, material name, s, H_S.

4.3. *Result and Discussion*

Table4 shows the estimation result (average and standard deviation of $r_{0,1}, r_{1,2}$) totally. As under condition ±40[V] is input to transmitter, Fig. 10, Fig. 11 shows the comparison of $r_{1,2}$, the former is of pattern 1 and 2, the latter is of pattern 3 and 4. As under condition

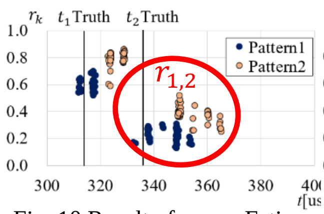


Fig. 10 Result of $r_{k-1,k}$ Estima-Tion for Pattern 1, 2 (Input:±40[V])

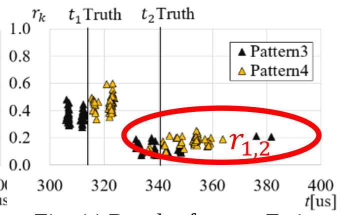


Fig. 11 Result of $r_{k-1,k}$ Estima-tion for Pattern 3, 4 (Input:±40[V])

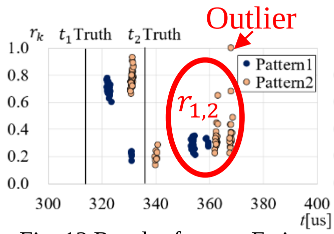


Fig. 12 Result of $r_{k-1,k}$ Estimation for Pattern 1, 2 (Input:±50[V])

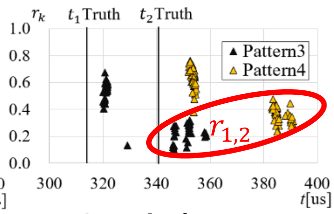


Fig. 13 Result of $r_{k-1,k}$ Estimation for Pattern 3, 4 (Input:±50[V])

Table4 Summary of Reflection Coefficient Estimation Result

		Input ±40[V]	Input ±50[V]
Average of $r_{0,1}$	Pattern1	0.60	0.71
	Pattern2	0.79	0.76
	Pattern3	0.36	0.56
	Pattern4	0.45	0.64
Standard Deviation of $r_{0,1}$	Pattern1	0.05	0.04
	Pattern2	0.05	0.05
	Pattern3	0.05	0.06
	Pattern4	0.05	0.08
Average of $r_{1,2}$	Pattern1	0.22	0.29
	Pattern2	0.37	0.32
	Pattern3	0.15	0.24
	Pattern4	0.18	0.36
Standard Deviation of $r_{1,2}$	Pattern1	0.05	0.04
	Pattern2	0.06	0.10
	Pattern3	0.08	0.06
	Pattern4	0.03	0.05

±50[V] is input to transmitter, Fig. 12, Fig. 13 shows the comparison of $r_{1,2}$, the former is of pattern 1 and 2, the latter is of pattern 3 and 4. where $r_{1,2}$ is surrounded by an oval. As Table 4 shows, pattern 2 shows higher $r_{1,2}$ than pattern 1 and, pattern 4 shows higher $r_{1,2}$ than pattern 3, where the result (increase of ζ_2 resulted in rising $r_{1,2}$.) could be found. However, the 40[us] of delay in t_1 was shown in pattern4 of Fig. 13, which was mainly caused by heterogeneity of sample depth. To reduce the error, bump or inclination of sample surface should be eliminated under coagulation.

4.4. *Conclusion*

As the result of the experiment, we succeeded in observing larger $r_{1,2}$ according to the increase of ζ_2, which may show the possibility of estimating the variety of hardness inside samples.

References

1. "The Japan Institute for Labour Policy and Training, " https://www.jil.go.jp/kokunai/statistics/timeseries/html/g0212.html (in Japanese)
2. "Yahoo News," https://news.yahoo.co.jp/byline/fuwaraizo/20130909-00027589 (in Japanese)
3. "Ministry of Agriculture, Forestry and Fisheries HP," https://www.maff.go.jp/j/wpaper/w_maff/h29/h29_h/trend/part1/chap2/c2_6_00.html
4. "Ministry of Agriculture, Forestry and Fisheries HP," https://www.maff.go.jp/j/press/shokusan/seizo/attach/pdf/190711-2.pdf
5. J Machado, P. Palma, S. Simoes, "Ultrasonic Indirect Method for Evaluating Clear Wood Strength and Stiffness", Non-Destructive Testing in Civil Engineering, pp. 1-6, 2009,
6. B. Cho and J. Irudayararaj, "A Noncontact Ultrasound Approach for Mechanical Property Determination of Cheeses", Food Engineering and Properties, Vol.68, No. 7, pp. 2243-2247, 2003.
7. S. Srivastava, S. Vaddadi and S. Sadistap, "Non-Contact Ultrasonic Based Stiffness Evaluation System for Tomatoes during Shelf-Life Storage", Journal of Nutrition and Food Sciences, Vol. 4, Issue 3, pp 1-6, 2014.
8. R.K.Young,N.K.Bose,"Wavelet Theory and Its Applications",Kluwer Academic Publishers,1993.
9. J. Carlson, J. Deventer, A. Scolan and C. Carlander, "Frequency and Temperature Dependence of Acoustic Properties of Polymers Used in Pulse-Echo Systems", Proceedings in IEEE Symposium on Ultrasonics, pp. 885-888, 2003.
10. S. Takeuchi, Tutorial on the underwater or medical ultrasound transducer, (In Japanese) J. Acoust. Soc. of Japan, Vol. 72, No. 5, pp. 264-272, 2016.
11. Honda Electronics Co., Ltd. "Ultrasound Handbook," (In Japanese), 2008.

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.

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.

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

Evaluation of Maps Constructed by Crawler-type Agricultural Robot in Different Farms

Takuya Fujinaga
Department of Electronics Engineering and Computer Science, Fukuoka University,
8-19-1 Nanakuma, Jonan-ku, Fukuoka 814-0180, Japan

Tsuneo Nakanishi
Department of Electronics Engineering and Computer Science, Fukuoka University,
8-19-1 Nanakuma, Jonan-ku, Fukuoka 814-0180, Japan

E-mail: t.fujinaga.wz@fukuoka-u.ac.jp, tun@ fukuoka-u.ac.jp

Abstract

Various studies have been carried out with the aim of realizing smart agriculture. Many of them have targeted at large field and large-scale facility that are relatively easy to verify and implement robots and IoT devices etc. On the other hand, although the ratio of small-scale facility is high in Japanese agriculture, there are few cases of study targeting small-scale facility. We aim to develop an agricultural robot for supporting agricultural work that can move autonomously in the field for small-scale greenhouse. This paper evaluates maps of three different environments (strawberry farm, herb farm and vegetable farm) constructed using the developed agricultural robot. In addition, through verification experiments and evaluation results, we describe the requirements and concerns for operating robot in small-scale greenhouse.

Keywords: Agricultural robot, Environmental map, Smart Agriculture, SLAM

1. Introduction

While the world population is increasing, the population in Japan is decreasing [1], and so is the agricultural working population [2]. Therefore, it is predicted that Japan, which relies on imports for a large amount of food, will depend more on foreign countries for food. The Ministry of Agriculture, Forestry and Fisheries drew a blueprint of the Agricultural DX (Digital Transformation) concept as a solution to resolve problems brought by aging and shortage of agricultural workers [3]. The initiative aims for farming with high production efficiency using data by utilizing technologies such as robots, artificial intelligence (AI), and the Internet of Things (IoT) in the field. In addition, it aims to provide agricultural products and foods with capturing the needs of consumers shown by the data.

Research and development such as autonomy of agricultural machinery [4], automation of agricultural work (harvesting and monitoring) [5], measurement of growth state of agricultural products [6] have been carried out toward the realization of agricultural DX and smart agriculture. However, only less than 20% of agricultural management bodies practice agriculture based on data. While the focus is on smarter large-scale agriculture, there are few research and development cases targeting small-scale agriculture. In Japanese facility horticulture, the proportion of small-scale facility of 1 ha or less is high [7]. Moreover, many farmers have given up on farming. There are issues such as whether cost-effectiveness can be obtained when introducing robots, AI, IoT, etc. into small-scale agriculture. However, in order to quantitatively evaluate the effect, it is necessary to develop agricultural robot and system for small-scale agriculture and verify them in the actual environment. Also, such development and verification not only can be expected to reduce the workload, but also contribute to the promotion of understanding and introduction of agricultural DX to farmers.

This study defines a farm with a planted area of 1 ha or less as a small scale and focus on agricultural robots for small-scale facility that produces crops. In small facility, few farmers perform multiple farming works (pesticide spraying, monitoring, harvesting, etc.). Therefore, Multi-functional agricultural robots are desired rather than mono-functional agricultural robots for large-scale facility (for example, harvesting robot and transport robot). In an outdoor farm with a vast site area, it is possible to autonomously drive agricultural machinery such as tractors using GPS [4]. In a plant factory, rails are laid to adjust the temperature in the greenhouse, reduce the workload, and improve efficiency, and it has been proposed to move the robots on the rails for harvesting and monitoring [5]. On the other hand, in small-scale facility, it is difficult to perform high-precision positioning by GPS compared to completely outdoors. In addition, the moving environment differs depending on the crops cultivated, and the behavior of robots is limited.

We have developed a prototype for the purpose of developing and operating a general-purpose agricultural work support robot for small-scale facility. Also, aiming at autonomous operation of the robot, as the first step, we verified the map construction of the facility using the prototype. This paper describes the outline of the prototype and the results of verification experiments for map construction conducted in three different farms (strawberry farm, herb farm, and vegetable farm). We also discuss requirements and concerns for operating the robot in small-scale facility obtained from the verification experiments are described.

2. Development of agricultural work support robot

First, the development requirements for a general-purpose agricultural work support are described. Agricultural work includes pinching, harvesting, pesticide spraying, and monitoring growth state. For robot to perform such agricultural work, the robot needs a moving mechanism to move in the facility, sensors to

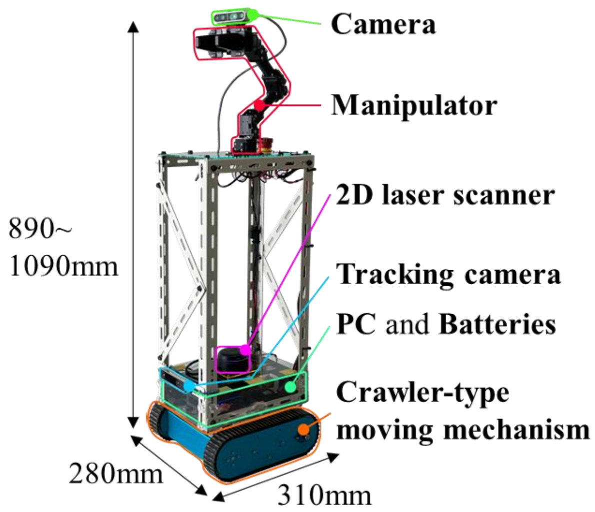


Fig. 1. Prototype developed for agricultural robot.

detect environment, plants and harvested objects, and manipulator or work machine to carry out agricultural work. The following functions are also required: map construction of an environment, self-position estimation, path planning, object detection, control of a manipulator or a working machine, and a decision-making system that integrates them. In addition, the robot needs to be sized so that it can move between ridges or cultivation beds in the facility and approach the target crops.

In this study, we have developed a prototype to meet the above requirements (Fig. 1). The prototype consists of the following elements.

- Crawler-type moving mechanism (NEXUS ROBOT, New Tracked Mobile Tank Robot Kit 10022) that allows the facility to move on soil
- 2D laser scanner (SLAMTEC, RPLIDAR A3) for map construction and self-position estimation
- 4DOF vertical articulated manipulator (ROBOTIS, OpenMANIPULATOR-X) for carrying out agricultural work (It has an end effector for gripping agricultural products.)
- RGB-D Camera (Intel RealSense Depth Camera D435i) for detecting plants
- Tracking camera (Intel, RealSense Tracking Camera T265) for estimating the self-position without depending on the ground environment (The odometry using the encoder mounted on the moving mechanism causes an estimation error due to slipping during moving.)
- PC for control (CPU: 3.5Ghz, Intel Core-i7, Memory: 32GB, OS: Ubuntu 16.04), and the software development was used ROS (Robot Operating System).

- Batteries (two types): one (12 V, 18 Ah) for manipulators and moving mechanisms, and the other (19 V, 20Ah) for PC

The size of the prototype is 310 mm in length, 280 mm in width, 890 to 1090 mm in height (variable depending on the operation of the manipulator), and the weight is 12.5 kg.

3. Verification of map construction

3.1. *Experimental method*

We conducted verification experiments to construct maps under different environments, with the aim of autonomously operating a prototype developed for the operation of a general-purpose agricultural work support robot in a small-scale facility. The experiments were carried out at three farms (greenhouses): *ichigokirari* which grows strawberries, *Kojima farm* which grows herbs, and greenhouse for practicing agricultural IoT [8] which is an experimental facility for introducing robot, IoT and AI technologies in the agriculture (Various vegetables such as tomato, asparagus, and cucumber are cultivated in this greenhouse.).

In this paper, each farm is called a strawberry farm, an herb farm, and a vegetable farm. The outlines and photographs taken at entrances of the farms are shown in Figs. 2 and 3, respectively. The strawberry farm and the herb farm do elevated cultivation, their grounds are covered with the general agricultural sheet, and the vegetable farm does soil cultivation. Note that equipment and tools for agricultural work are installed and temporarily put in each farm, and no consideration is given to operating the robot.

The experiments were conducted on August 31, 2021, September 1 and November 30, at the strawberry farm, on September 14 at the herb farm, and on September 3 at the vegetable farm. In the strawberry farm in August and September, strawberries were not cultivated in the cultivation beds as shown in Fig. 4(a) before planting the seedlings. On the other hand, in November, seedlings have been planted and flowering has begun as shown in Fig. 4(b). Fig. 4(a) and (b) are photographs of the inside of area B in Fig. 2(a). As can be seen from these photographs, a white sheet is wrapped around the cultivation beds in Fig. 4(b). This sheet has the role of keeping the culture medium warm. It is wrapped around the cultivation beds in November when

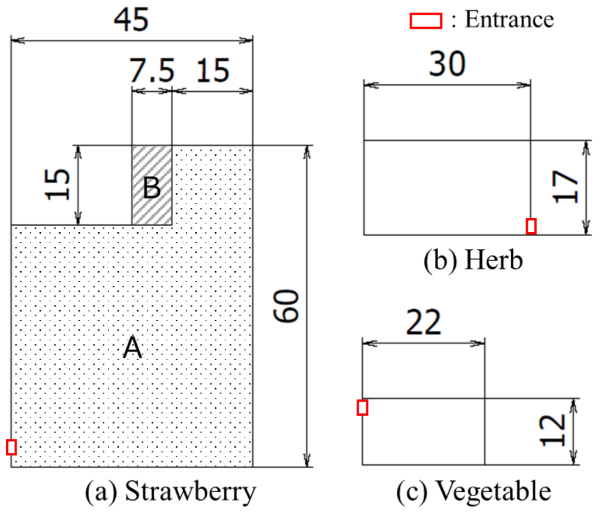


Fig. 2. Outlines of the farms where the verification experiments were conducted (The unit is meter.).

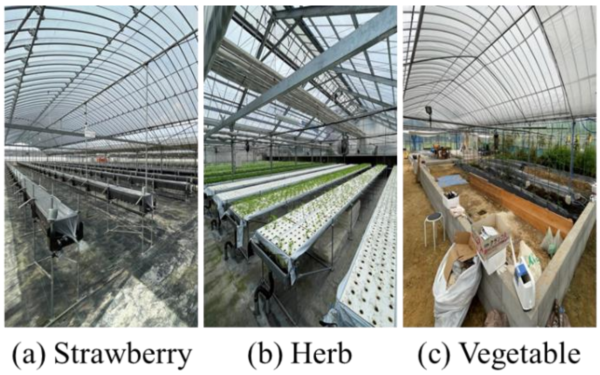


Fig. 3. Photographs of the farms where the verification experiments were conducted.

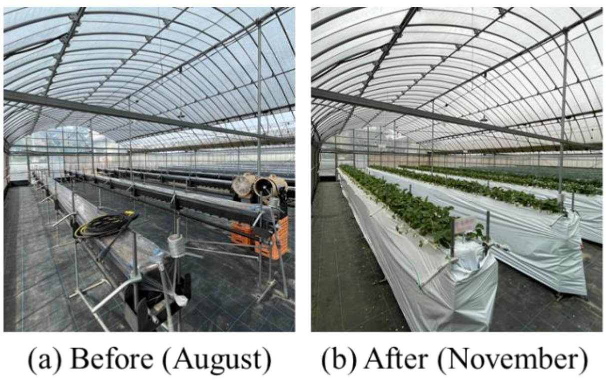


Fig.4. Photographs before and after planting in the strawberry farm.

the temperature begins to drop after planting is carried out in September. In this paper, in addition to the evaluation of maps under different environments, we focus on area B in Fig. 2(a) and compare the constructed maps before and after the sheet is wrapped.

In the experiments, the prototype was visually controlled remotely using a remote controller, and data for constructing an environmental map was acquired. However, if the prototype could not move due to equipment and tools for agricultural work, human intervention was required. Using the data of the 2D laser scanner and the tracking camera, we constructed maps by gmapping that can construct a 2D map with the existing package for SLAM (Simultaneous Localization and Mapping) of ROS.

3.2. *Results*

The environmental maps constructed in three different farms are shown in Fig. 5 and Fig. 6. Fig. 5(a) is the area A of the strawberry farm, (b) is the herb farm, and (c) is the vegetable farm. Fig. 6(a) is shown before planting in the strawberry farm area B, and (b) is shown after planting. In Fig. 5 and Fig. 6, the size of the farm measured from the constructed maps and its actual size are shown (The actual size in parentheses. The unit is meter.). Regarding the accuracy, the error was +2 m for the strawberry farm, the error was +1 m for the herb farm, and the error was +1 m for the vegetable farm compared to the actual size.

In the strawberry farm and the herb farm, the cultivation beds are supported by the props as shown in Fig. 3 (a) and (b). The props can be confirmed from Fig. 5(a), (b) and Fig. 6(a). On the other hand, in the map Fig. 6(b) constructed in the area B of the strawberry farm after planting, the props of the cultivation beds cannot be confirmed. The cultivation beds are covered with sheet, and the sheet can be confirmed in Fig. 6(b). Therefore, the cultivation area is clear. In addition, each farm connects one greenhouse (single building), and there are beams at the connecting part (the part that becomes the end of the single building). The beams can also be confirmed in Fig. 5(a) and (b). There are two areas in the vegetable farm. One is the area where vegetables are cultivated and becomes soil (The plant area in Fig. 5(c)),

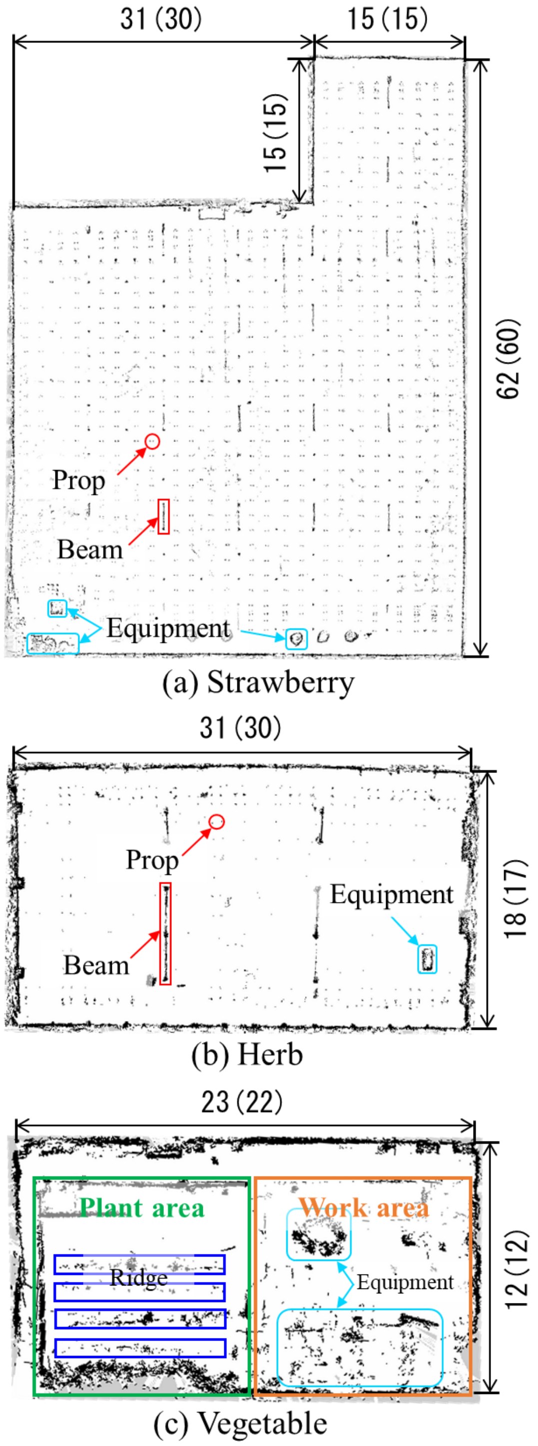


Fig. 5. Maps constructed in verification experiments (The unit is meter.).

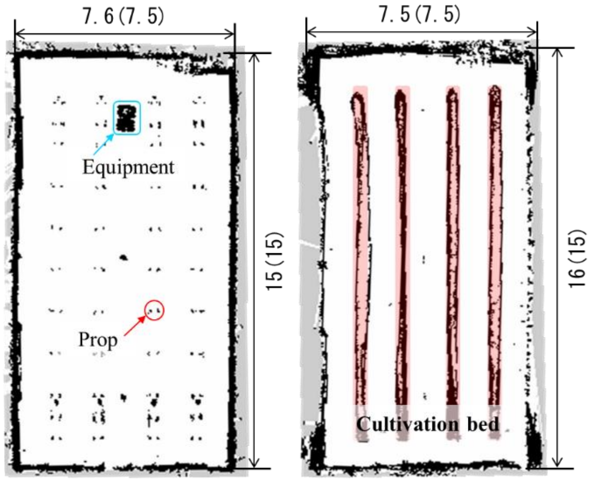
and the other is the area where boards are laid for work (The work area in Fig. 5(c)). Ridges can be seen in the plant area and work desks can be seen in the work area from Fig. 5(c). Agricultural tools installed or temporarily put in the farm are also confirmed from each map.

4. Discussion

4.1. Verification results of map construction

For robot to move autonomously in the farm, it is necessary to detect obstacles in the farm and avoid them. From the results of the verification experiments, the outer shape of the farm can be grasped. Although the props and beams in the farm are confirmed, the robot need to detect and avoid those obstacles. There is noise in the constructed maps, it is necessary for the robot to distinguish them only from the geometrical features such as props and beams. It is thought that mounting an additional camera or 3D laser scanner on the robot will make it easier to distinguish between them and avoid obstacles but adding sensors will increase costs.

On the other hand, it is considered that autonomous moving is possible with only the minimum sensors and processing by changing the farm environment from the verification results. The sheet was not wrapped around the cultivation beds in the experiment conducted before planting. When the map Fig. 6(a) constructed in that environment is used, it is difficult for the robot to move autonomously in the farm without detect the cultivation beds. If so, the cultivation beds are connected as shown in Fig. 4(a), but in Fig. 6(a), only the props of the cultivation beds can be confirmed. Therefore, when trying to find the optimum solution without some conditions in the path planning during autonomous moving, the path in the middle of the connected cultivation beds will be a candidate. However, in the map Fig. 6(b) constructed in an environment where the sheet is wrapped around the cultivation beds as shown in Fig. 4(b), the row of connected cultivation beds can be confirmed and the area where the robot can move becomes clear. Note that in map construction in such an environment, the data from 2D laser scanner will be two parallel lines representing the cultivation beds. To deal with the degeneracy in map construction, this prototype used information of point cloud data from 2D laser scanner and odometry from tracking camera.



(a) Before (August) (b) After (November)

Fig.6. Maps before and after planting in the strawberry farm.

We target small-scale facility, and farmers are concerned about the cost-effectiveness of introducing robots, AI, IoT, and so on. At present, the prototype has the configuration that is considered necessary to support agricultural work. However, instead of simply adding sensors, it is a future task to clarify the requirements for agricultural work support robots and the farm environmental conditions for the robots by repeating verification experiments on various farms. The requirements and concerns in the robot operation obtained from the verification experiments are described in the next section.

4.2. *Operation of robot in small-scale facility*

This prototype was designed to a size that allows it to move between ridges or cultivation beds (The width is 280 mm). It was able to move at the farms conducted in the verification experiments. However, there were places where it was not possible to move due to equipment and agricultural tools. In such places, it is necessary to take measures such as avoiding obstacles by the robot or making the environment easy for the robot to move.

There are undulating spot and places where watering hoses and agricultural tools are temporarily put. For autonomy, moving performance for overcoming undulations and hoses, and algorithm for detecting agricultural tools, etc. are required. The strawberry farm and herb farm have relatively sturdy agricultural sheet laid on the ground. On the other hand, the vegetable farm use film-shaped sheet for covering ridges, and there were cases where the prototype involved the sheet. There were also muddy spots, it is difficult for the prototype to move to such spots. In robot operation, after understanding the cultivation method of the farm and the characteristics of the facility, it is necessary to clarify the objects that need to be detect and make a path plan for autonomous movement.

Small-scale facility is managed by fewer workers than larger facilities. So, small-scale facility has a small cultivation area, and it is easy for workers to move within the farm. Taking advantage of this point, an operation method in which a person helps a robot can be considered as an example. For example, the robot grasps the map of the farm and the location of static equipment (props, beams, etc.), detect dynamic objects (hoses, agricultural tools, etc.). When a problem such as unavoidable occurs, the robot stops and warns the farmer, then the problem is solved by human intervention.

The operation of robot in small-scale facility needs to be customized according to the scale of farmer's management. In some cases, complete autonomy may be desired by adding sensors, and in other cases, it may be desirable to have minimum components and cooperate with humans. We aim to develop a farm work support robot that can handle any situation.

5. Conclusion

This paper described the outline of prototype for a general-purpose agricultural work support robot and the experiments for map construction carried out in three farms. The outer shape of each farm was able to be grasped from the constructed map. The accuracy was a maximum error of +2 m, and some equipment were confirmed. In addition, we showed that changing the farm environment may be effective for autonomous moving of robot. The requirements and concerns in robot operation obtained through verification experiments were described. As the next step, we aim for autonomous moving in the farm based on the constructed map.

Acknowledgements

In the experiments, we received much cooperation from Mr. Seigo Tanaka of *ichigokirari*, Mr. Hiromitsu Kojima of *Kojima farm*, and Prof. Kazuo Ishii of Kyushu Institute of Technology. We would like to express our sincere gratitude for their kind assistance. This work was supported by JSPS KAKENHI Grant Number JP21K20604.

References

1. Statistics Bureau of Japan.
2. Ministry of Agriculture, Forestry and Fisheries. (In Japan ese)
3. Ministry of Agriculture, Forestry and Fisheries.
4. N. Noguchi, "Agriculture Vehicle Robot", Journal of Robotics and Mechatronics, 30(2), pp. 165-172 (2018).
5. T. Fujinaga, S. Yasukawa, K. Ishii, "Tomato Growth State Map for the Automation of Monitoring and Harvesting", Journal of Robotics and Mechatronics, 32(6), pp. 1279-1291 (2020).
6. K. Shimomoto, K. Takayama, N. Takahashi, H. Nishina, K. Inaba, Y. Isoyama, S. OH, "Real-time Monitoring of Photosynthesis and Tran-spiration of a Fully-grown Tomato Plant in Green-house", Environmental Control in Biology, 58(3), pp. 51-57 (2020).
7. Ministry of Agriculture, Forestry and Fisheries.
8. enPiT-Pro, Practical laboratory exercises on agriculture,

A Network of Underwater Sensors Estimating Feeding Behavior for Digitizing Optimized Feeding Decisions in Marine Aquaculture

Dominic B. Solpico*
Department of Life Science and Systems Engineering, Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku Kitakyushu, Fukuoka 808-0135, Japan

Kota Mishima
Department of Human Intelligence Systems, Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku Kitakyushu, Fukuoka 808-0135, Japan

Yuya Nishida
Department of Human Intelligence Systems, Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku Kitakyushu, Fukuoka 808-0135, Japan

Kazuo Ishii
Department of Human Intelligence Systems, Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku Kitakyushu, Fukuoka 808-0135, Japan
E-mail: solpico.dominic-bautista806@mail.kyutech.jp, mishima.kota928@mail.kyutech.jp, y-nishida@brain.kyutech.ac.jp, ishii@brain.kyutech.ac.jp
www.kyutech.ac.jp

Abstract

Optimization of fish feeding in marine aquaculture has relied on an expert farmer's decision-making based on subjective experience. This paper presents the development of a network of underwater current, imaging and IMU sensors for estimating fish feeding behavior for digitizing expert feeding decision-making. We constructed the sensor units and deployed them in fish cages and collected measurements during feeding activities. Experiment results indicate that currents were highest at the surface within the duration of the feeding activity.

Keywords: aquaculture, feeding, sensor, network, currents

1. Introduction

Marine aquaculture is one of the largest contributors to fish production in Japan, accounting for around 22% of the country's production volume as of 2019.[1] However, production has remained stagnant over the years. In addition, with higher costs in feeds and fish meal, the average income of farmers has been decreasing. Efficient fish feeding conventionally depends on the fish farmer's judgement of the fishes' feeding behavior, which is based on his/her intuition and subjective experience. This makes efficient feeding more difficult to achieve for less experienced farmers. By applying digital transformation (DX) to experts' knowledge to optimize fish feeding, data from fish cages could help less skilled farmers improve their decision-making, raise more high-quality fishes, and increase their income by reducing the excess feeding costs from uneaten feeds, which also reduces the pollution these feeds release to the environment. This makes the industry more sustainable.

So far, research efforts to digitize expert decision making have been made in recirculating aquaculture systems, applying artificial intelligence on data from water quality and imaging sensors collected in fish tanks.[2,3,4] Such systems have not yet been implemented in open sea cages, where environment conditions are uncontrolled. Measurement of currents generated by fishes has been proposed to estimate their feeding behavior.[5] Although this has already been proposed as part of a monitoring system for marine cages, only simulation of real-world information has been made so far.

This paper presents the development of a modular network of sensors that measures underwater currents, and record fish activities in sea cages during feeding activities. Such data can help less experienced farmers gain further insight on how experts optimally dispense feeds to fish cages, helping them make better feeding decisions in their own cages and improve their farming operations. This paper includes the deployment of the sensor network on fish cages and the results from the measurements during feeding activities.

2. Modular Sensor Network

2.1. *Network Architecture*

The sensor network is composed of one or more sensor units that are attached to the sides of the fish cages (Fig. 1). Each unit consists of a top module and multiple sensor modules. The top module houses the batteries that power the sensor modules as well as the bridge router that relays the data from the sensors to the main router, and then to the user computer, both aboard the feeding boat.

Each sensor module consists of at least two sensors: a modified flow sensor for measuring underwater currents; a network camera to record fish activity underwater. Some of the modules have an inertial measurement unit (IMU) to measure any movement of the sensor frame cause by currents. A microcontroller collects and calculates readings from both the flow sensor and IMU and sends them by serial to a device server.

Both the device server and the network camera in the module connect to an Ethernet switching hub, which forms a daisy chain with the other sensor modules within the unit, connecting to the bridge router. With all sensor

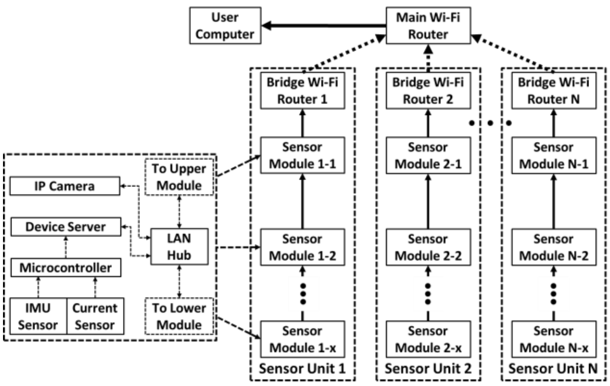


Fig. 1. Topology of the modular sensor network for estimating fish feeding behavior.

unit routers connecting in bridge mode to the main router, the user computer has access to all sensors, collecting data from all of them. With this design, all sensors are well synchronized, with all collected data timestamped more accurately.

2.2. *Prototype Sensor Unit Development*

All electronic components of each sensor module were placed inside a waterproof enclosure mounted on an aluminum frame, except for the current sensor, which was waterproofed and was mounted on the frame above the enclosure. The enclosure featured underwater connectors, to link the power and data of the internal electronics to the current sensor and to the other modules.

Eight sensor modules and four top modules were constructed for the fish cage experiment. Depending on the experiment plan, each sensor unit can have up to three sensor modules. This number was due to the power limitation of constructed cable, although more modules could be connected had a larger gauge been used for the cables' power lines. The unconnected connectors of the lowest sensor modules were sealed.

Modular metal frames were designed and constructed especially for the sensor modules for flexibility in the arrangement of sensors in the fish cages. Each sensor module is fixed in a modular frame in such a way that when multiple frames are attached, the modules would be properly spaced as desired. Each sensor unit has one frame with a sensor module and a top module.

A video management software (VMS) was used in the user computer to display and record the views from the cameras. Data from the current sensors and the IMUs were recorded by accessing the device servers via Telnet using a terminal emulator with logging function, storing accurately timestamped data into files for later analysis.

3. Fish Farm Experiment

3.1. *Overview*

In November 2021, a two-day experiment was made using this sensor network in three fish cages in the town of Shin-Kamigotō, in Nagasaki prefecture. On the first day, two sensor units with three sensor modules each were deployed, one each for two fish cages. Both units were attached to the center south side of the cages. On the second day, four units with two modules each were deployed in one cage, two on the north side and two south side. Two units were mounted on the center of opposite sides. The other two were supposed to be mounted on the center of the east and west sides of the cage, but these are where the feeding boat would dock on. Therefore, they were attached to the northwest and southeastern corners instead.

The sensors were deployed on 11.7 x 11.7 m^2 square cages with depth of five meters, as shown in Fig. 2. All cages contained Japanese amberjack (*Seriola quinqueradiata*) that have been raised for around half a year. There were 11,000 fishes in each of the two cages on the first day, all caught from the wild as fries. They were fed from the boat using a dispensing machine. The cage on the second day contained 8,000 fishes raised from an artificial hatchery. They were fed manually by the farmer from the platform on the center of the cage.

At the start of each data collection, the farmer was requested to wait for five minutes before starting to feed the fishes. This was done to get baseline data before feeding began. The sensors were made to collect until five minutes had passed from the end of the feeding activity to get data when the fishes returned to the same state prior to feeding.

The sensor units were attached to the fish cages on the afternoon of the day before the measurement, as feeding was performed in the morning. It was decided by the research group to fix all modules two meters apart from each other, with the topmost sensor module positioned 0.5 m from the surface. The lower sensors were therefore fixed at 2.5- and 4.5-meter depths.

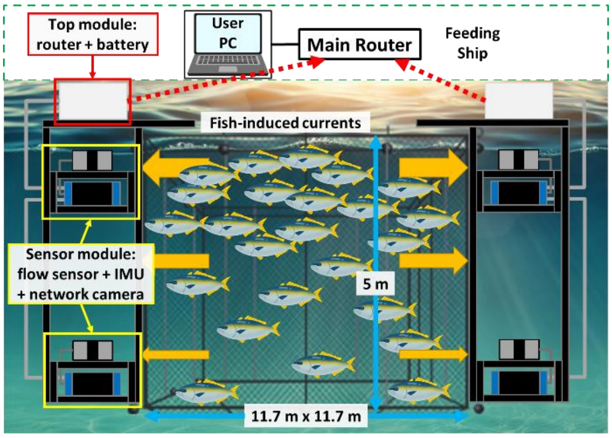


Fig. 2. Side view of the fish cage experiment with two sensor units fixed on opposing sides. This setup was made on the third cage on the second day.

The starting and ending times of dispensing feeds were noted for all measurements. The feeding activity in the third cage was also recorded on video, from which the progression of the amount of feed given was estimated. For easier analysis of data, moving averages of current and IMU measurements within 15 seconds were calculated to reduce the noise, thus smoothening the plot. Noted events during feeding and observations from the underwater cameras were compared with the sensor data.

3.2. *Results and Discussion*

3.2.1. *First Day Measurements*

In the first cage, measurements were collected only at 0.5 and 2.5 m, as connection with the sensor module at 4.5 m could not be made due to internal loose connection. Data at all three depths in the second cage was collected.

The feeding in the first cage lasted around nine minutes, as shown in Fig. 3a. Before feeding started, cameras did not capture fishes swimming at 0.5 m. Feeds were immediately dispensed continuously at the start. At around this time, cameras recorded fishes swimming very fast towards the surface to feed. View at 0.5 m started to blur a minute after due to the bubbles flowing towards the camera.

Although vigorous splashing was already present, surface current measurements did not start to rise above 4 cm/s until around the fourth minute. Current peaked at the sixth minute at around 21 cm/s, which gradually declined. At the end of feeding, current was at 12 cm/s.

Video recordings showed that fishes started going down to the bottom of the cage and surface gradually became as calm and clear as before feeding started. Current dropped to zero in less than a minute. Despite presence of fishes during feeding, only zero readings were collected at 2.5 m.

Shown in Fig. 3b, feeding in the second cage lasted around seven minutes, starting gradually for a minute before dispensing feeds continuously. Even before feeding began, surface currents up to around 10 cm/s was already measured. Some fishes could be seen at the bottom of the video from 0.5 m. Although most of the fishes could be seen from the cameras at 2.5 and 4.5 m around this time, current measurements were very small.

Surface current temporarily dropped to zero during gradual feeding. Fish activity was still the same, as seen in Fig. 4. And then, fishes started swimming upward very fast at around this time to feed. This current drastically rose to peak at around 37 cm/s when continuous feeding started. Video started blurring from the flowing bubbles at around this time.

Fewer fishes could be seen at 4.5 m during feeding as most were at the surface. Peak surface readings gradually decreased as time progressed. Substantial currents, reaching to 10 cm/s, were measured at 2.5 m towards the end of feeding. By the time feeding was stopped, surface reading was at around 14 cm/s and eventually declined to zero within the same minute. At around this time, surface became calm again, and fishes started gathering at the bottom of the cage.

The amount given to the fishes in the first cage was 240 kg while only 230 kg was given to the fishes in the second cage, as 10 kg of feed remained when the farmer decided to stop the feeding machine. The farmer assessed that the fishes were feeding less actively compared to those in the first. However, surface measurements from both cages seem to suggest that feeding activity was greater in the second cage.

3.2.2. *Second Day Measurements*

For the third cage on the second day, due to physical connection problems, no data was collected at 2.5 m on the northeast corner while incomplete data was collected at the same depth on the southwest corner. IMU data was available from the other three modules at 2.5 m. Most of the sensor data from the units at the corners were negligibly small. Therefore, the focus of discussion will be on the data from sensors on the center north and south sides.

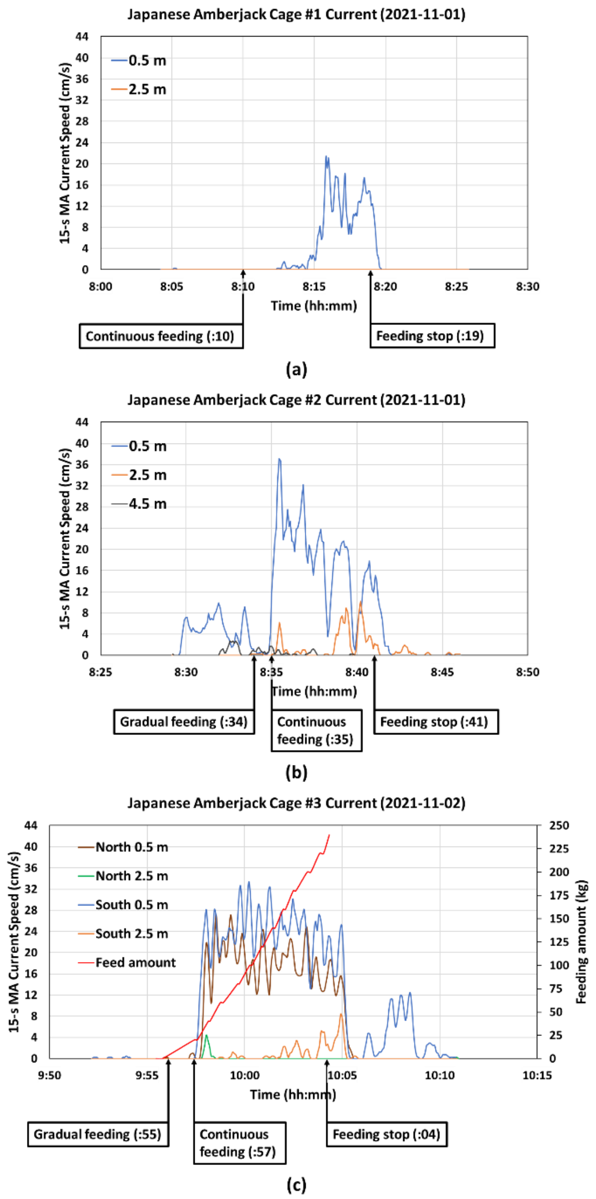


Fig. 3. Current measurements from the first (a) and second (b) fish cages on the first day, and from the third fish cage on the second day (c).

Feeding in the third cage lasted nine minutes, same as in the first cage, as seen in Fig. 3c. Readings were at zero before feeding. Like in the first cage, cameras at 0.5 m captured no fishes at surface. At 2.5 m, while no fishes were seen on the south side, they were seen swimming on the north side. Around the time gradual feeding started,

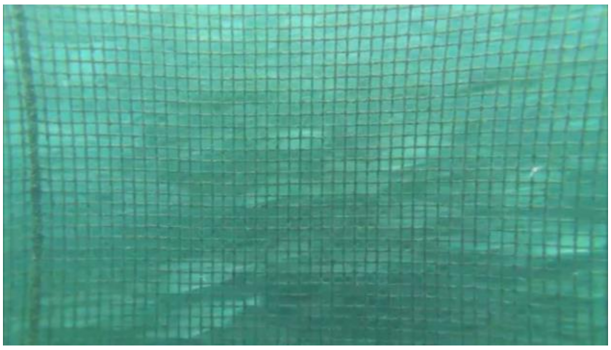


Fig. 4. Camera view at 4.5 m at the second cage during gradual feeding. They were still schooling at the bottom of the cage.

fishes started swimming to the surface fast, like those in the previous cages. At 2.5 m of north side, small readings were detected after feeding started.

Surface measurements started to rise drastically only after continuous feeding started, as measured by sensors on both sides almost at the same time. These peaked at 33 cm/s, with the following peak currents declining until the end of feeding, at which they dropped to zero simultaneously on both sides. After feeding stopped, currents up to 12 cm/s were still read by the sensor on the south side. On the south side at 2.5 m, a pattern of currents like that from the second cage was seen towards the end of feeding. Fish activity from then onward was same as in the beginning.

3.2.2. *IMU Measurements*

Due to a flawed design decision, only four out of the eight sensor modules had an IMU included. Therefore, it was decided to deploy two modules with IMU at 2.5 and 4.5 m in the first cage (although no connection at 4.5 m), and at 0.5 and 4.5 m in the second cage. In the third cage, all IMUs, except for the one on the northeast unit, collected at 2.5 m.

Offsets were evident in the acceleration and rotation data in various axes, as shown in data from one of the sensor modules in Fig. 5, because the IMUs were uncalibrated when they measured data. While fluctuation patterns in acceleration and rotation could be seen within the feeding duration, these changes were at very small scales. Such was the case for all collected IMU data. Accelerations were less than 0.1 m/s^2 (Fig. 5a & 5b), while rotations were less than .01 rad/sec or around 0.6 degrees (Fig. 5c). Data suggest that movement in frames were very small even during fish activities, although further analyses are needed.

4. Conclusion

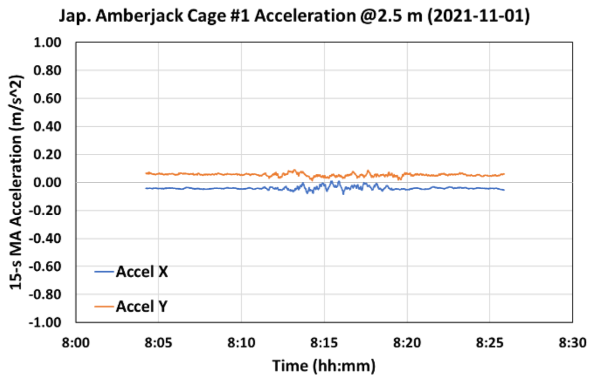
In this paper, a network of sensors was developed to estimate fish feeding behavior, which could be useful for optimizing feeding control. It consisted of current sensors, cameras and IMUs, collecting data at different depths and sides of a fish cage. This system was deployed in three fish cages and collected measurements during feeding activities. Experiment results suggest that changes in current measurements are related to the fish feeding behavior, as recorded. However, data needs to be analyzed further to understand these relationships better.

Acknowledgements

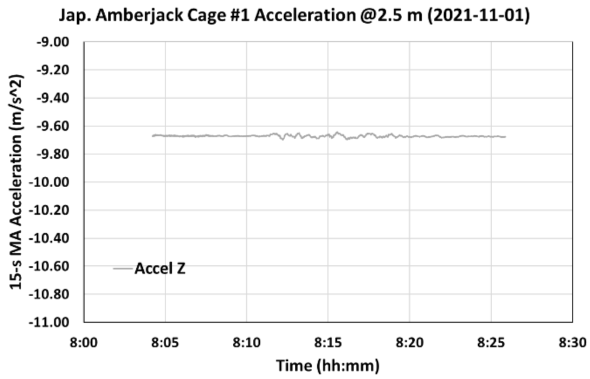
This research is funded by the Japan Science and Technology Agency through the Adaptable and Seamless Technology transfer Program through Target-driven R&D (A-STEP) research grant (No. JPMJTM20GS). We would like to acknowledge the support of our partner company, Belltechne Co., Ltd., in this collaboration for the inputs on the system design, for constructing the mounting frames for the sensors, and for all the support in the execution of the experiments. We would also like to thank Tokumaru Co., Ltd. for providing the access to the fish cages for the experiment as well as for all the support in the sensor deployment. D. B. S. acknowledges the support of the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT) through the Monbukagakusho: MEXT scholarship.

References

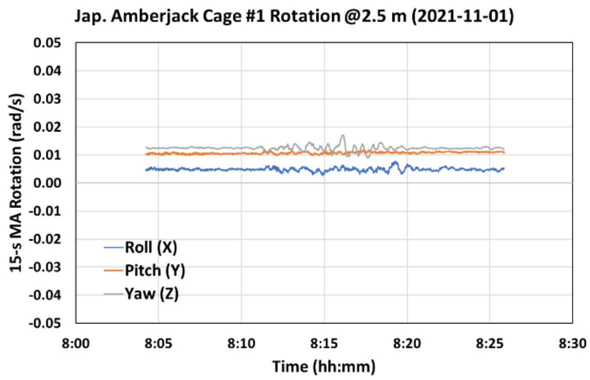
1. Japan Fisheries Agency, "FY2020 Trends in Fisheries & FY2021 Fisheries Policy Summary," Japan Fisheries Agency, Tokyo, Japan, 2020 [Online]. Available: https://www.jfa.maff.go.jp/j/kikaku/wpaper/attach/pdf/index-7.pdf
2. C. Zhou et al., "Near infrared computer vision and neuro-fuzzy model-based feeding decision system for fish in aquaculture," *Comput. and Electron. in Agri.*, vol. 146, pp. 114-124, 2018.
3. C. Zhou, et al., "Evaluation of fish feeding intensity in aquaculture using a convolutional neural network and machine vision," *Aquac.*, vol. 507, pp. 457-465, 2019.
4. T. Wu, Y. Huang, and J. Chen, "Development of an adaptive neural-based fuzzy inference system for feeding decision-making assessment in silver perch (*Bidyanus*



(a)



(b)



(c)

Fig. 5. Acceleration measurements at X and Y axes (a) and at Z axis (b), and rotation measurements at all axes (c) at 2.5 m in the first cage.

bidyanus) culture," *Aquacult. Eng.*, vol. 66, pp. 41-51, 2015.

5. M. Garcia, S. Sendra, G. Lloret, and J. Lloret, "Monitoring and control sensor system for fish feeding in marine fish farms," *IET Commun.*, vol. 5, no. 12, pp.1682-1690, Aug. 2011.

Authors Introduction

Mr. Dominic B. Solpico

Mr. Solpico obtained his M. S. degree in Electronics Engineering at the Ateneo de Manila University, Philippines in 2015. He is currently taking his Doctoral degree course at Kyushu Institute of Technology, Japan. His research interests are in the fields of Intelligent Aquaculture, Wireless Sensor Networks (WSNs), and Field Robotics.

Mr. Kota Mishima

Mr. Mishima obtained his B. S. degree in Engineering at Kyushu Institute of Technology, Japan in 2021 and is currently taking his Master's degree course at the same university. His research interests are in the fields of Fluid Mechanics and Dynamics.

Dr. Yuya Nishida

Dr. Nishida is currently an Associate Professor at the Department of Human Intelligence Systems, Kyushu Institute of Technology, Japan. He obtained his M. 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.

Prof. Kazuo Ishii

Prof. Ishii 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.

Proposal of a Swarm Intelligent Underwater Glider System for Long-term Three-dimensional Wide-area Ocean Observation

Kanako Kobatake, Masakazu Arima*
Department of Marine System Engineering, Osaka Prefecture University
Sakai, Osaka, 599-8531, Japan
E-mail: dd102004@edu.osakafu-u.ac.jp, arima@marine.osakafu-u.ac.jp
http://arima-labo.jp

Abstract

The ocean is facing serious threats due to global warming and ocean acidification. In order to conserve the ocean, it is very important to understand the ocean environmental situations correctly over a long term and wide area. In this research, the authors propose a swarm intelligent underwater glider system for ocean observation. The operation simulation within Japan's Exclusive Economic Zone (EEZ) verified the validity and effectiveness of the proposed system, and the number of underwater gliders was estimated for requiring to realise continuous and even observation of Japan's EEZ.

Keywords: Autonomous Underwater Glider, Swarm intelligent underwater vehicle, Path search, SOARER

1. Introduction

One of the Sustainable Development Goals (SDGs) adopted by United Nations in 2015 is sustainable management and protection of marine and coastal ecosystem from pollution in goal 14: LIFE BELOW WATER. In order to maintain marine and coastal ecosystem and to comprehend ocean condition accurately, it is necessary to establish a long-term marine survey system for the entire sea area. However, it is not easy to investigate in special underwater environments. In conventional survey method, research ships, buoys, floats, Remotely Operated Vehicles (ROVs), and Autonomous Underwater Vehicles (AUVs). However, manned research submersible and ROV are not suitable for wide area ocean survey because the movable range is limited. AUVs can cruise freely under water, but it has limited uptime due to the large battery consumption.

It is widely known that Autonomous Underwater Glider (AUG) has an ability of long-term and wide-area operation by consuming less energy because it propels using fluid dynamic force acting on fuselage and wings by controlling underwater weight and attitude angle while diving and surfacing. In this research, we propose a three-dimensional wide-area ocean observation system using swarm intelligent underwater vehicles centred on multiple AUGs suitable for realising a long-term and continuous operation. It is necessary to operate a huge number of AUGs during the same term expanding over the entire observatory sea area. Some methods of operation while patrolling the set route[1)] and survey using some vehicles while following the leader vehicle[2)] have been proposed.

In this paper, the authors propose a method of searching the movement path of own vehicle. Furthermore, swarm

intelligence for AUGs is introduced for realising the reluctant cooperative behaviour.
In addition, we try to verify the effectiveness of the proposed method by operative simulation within the Japan's Exclusive Economic Zone (EEZ), and to estimate the number of vehicles required to achieve observation of Japan's EEZ. At first, the autonomous underwater glider "SOARER" used in this paper is introduced in section 2. Next, the methods of optimal path search and cooperative behaviour are described in section 3. The operative simulation in Japan's EEZ is demonstrated in section 4. Finally, conclusions are described in section 5.

2. Autonomous Underwater Glider with Independently Controllable Main Wings

2.1. *SOARER*

AUG is an underwater research vehicle suitable for a long term operation. In this research, the autonomous underwater glider with independently controllable main wings named "SOARER"[3, 4] which is developed with the aim of improving maneuverability was used. SOARER has characteristics which can be changed roll angle and gliding angle appropriately by controlling its main wings. The snapshot of the SOARER glider is shown in Fig.1 and specifications are listed in Table 1.
SOARER is mounted three sensors, six actuators and an environmental observation profiler. Fig. 2 shows system configuration of SOARER. The mounted sensors are IMU (Inertial Measurement Unit), GPS (Global

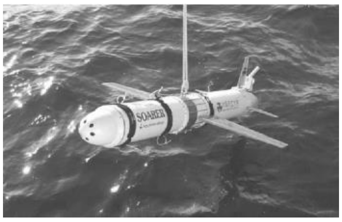


Fig. 1 Snapshot of SOARER.

Table 1 Specifications of SOARER.

Items	Value
Length	2.416 m
Breadth	1.460 m
Height	0.505 m
Mass	Abt. 80 kg
Wing profile	NACA0009
Wing chord at tip/root	0.13 m/0.26 m
Volume of ballast tank	2.01 L
Maximum operational depth	1500 m

Positioning System) and a depth sensor. The actuators are

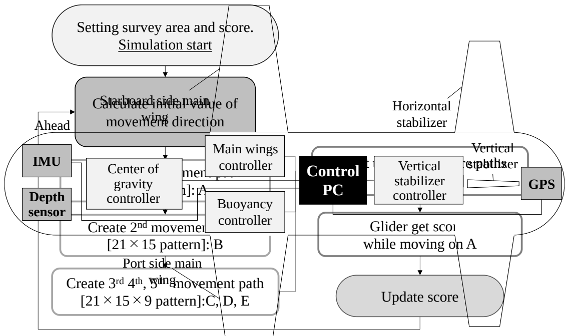


Fig. 4 Outline of path search method.
Fig. 2 System configuration of SOARER.

consisted of a centre-of-gravity controller, a buoyancy controller and four wing controllers. The centre-of-gravity controller controls SOARER's pitch angle by shifting the movable weight to ahead and astern. The buoyancy controller controls underwater weight by injecting and draining water to dive or surface. Wing controllers are vertical and horizontal stabiliser, starboard and port main wings controller. Vertical and horizontal stabilisers mounted behind the vehicle controls yaw and pitch angle by controlling the vertical and horizontal wing like rudder and elevator. Right and left main wings mounted on starboard and port side of the centre of vehicle control its gliding angle and bank angle. For controlling its gliding angle, both main wings can be rotated in the same phase to change its hydrodynamic force around wings. Furthermore, the bank angle can be controlled by rotating in the opposite phase to change its hydrodynamic force acting on each wing.
The environmental observation profiler "RINKO-Profiler" that measures dissolved oxygen (DO), depth, temperature, electrical conductivity, freshwater EC, salinity, chlorophyll, turbidity in seawater is mounted on the SOARER glider.

2.2. *Disturbance compensation underwater*

AUG cannot be positioning underwater using GPS, because the radio wave is not transmitted in the water. Therefore, tracking error increases because AUG cannot recognise disturbance in the sea. In this research, underwater position and disturbance compensation are estimated using the dead reckoning which estimates the position from the heading angle and the speed of the vehicle.
The position is estimated by dead reckoning, and the disturbance compensation uses this estimated vector as shown in Fig. 3. AUG dives to shallow depth for

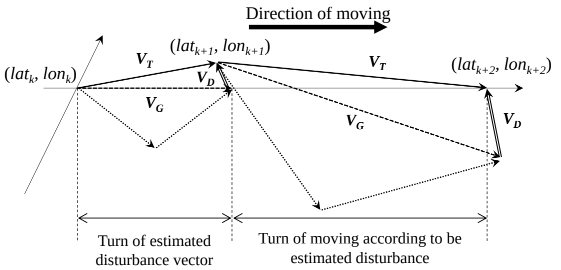


Fig. 3 Estimated disturbance vector.

estimating *disturbance vector*. This *disturbance vector* is calculated from *guessed vector* obtained from moving direction and speed of movement, and *true vector* indicating actual moving direction. In accordance with the estimated disturbance, AUG moves to next surfacing point considering *disturbance vector* during diving to operational depth.

Disturbance vector is estimated as follows:

$$\boldsymbol{V_T} = (dlat, dlon) \tag{1}$$

$$\begin{aligned} dlat &= (lat_{k+1} - lat_k)\cdot 60\cdot 1852 \\ dlon &= (lon_{k+1} - lon_k)\cdot cos(m.l)\cdot 60\cdot 1852 \\ m.l. &= (lat_{k+1} + lat_k)/2 \end{aligned} \tag{2}$$

$$\boldsymbol{V_D} = \boldsymbol{V_T} - \boldsymbol{V_G} \tag{3}$$

where, $\boldsymbol{V_G}$ is *guessed vector*, $\boldsymbol{V_T}$ is *true vector*, $\boldsymbol{V_D}$ is *disturbance vector*, (lat_k, lon_k) is position of starting gliding, (lat_{k+1}, lon_{k+1}) is the estimated surfacing position. *True vector* is calculated by middle-latitude diving as following Eq. (2), and *guessed vector* is estimated by dead reckoning. *Disturbance vector* is calculated by Eq. (3).

3. Optimal Path Search Algorithm

3.1. *Outline of optimal path search algorithm*

For realising wide-area ocean observation with multiple AUGs, it is necessary to select and determine their cruising route through sea areas. Fig. 4 shows outline of optimal path search algorithm. In this research, the optimal path search is realised by iterational local search method which repeats optimal solution search on neighborhood and random transformation of initial value. Every vehicle moves while correcting path planning every surfacing to the sea surface, because AUG cannot surface at the planned point. This algorithm is consisted of three parts, setting survey area part, path search part, move along the optimal route and updating score part. The optimal path is searched by repeating these parts.

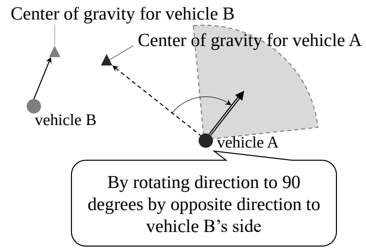


Fig. 6 Reluctant cooperation by vehicle A.

The first part is setting the survey area. The survey area is divided into cubes which are given the score according to importance of the area. The survey area is divided every 1 degree in horizontal X-Y direction, and depth in Z direction is divided as shown in Table 2.

The second part is creating some paths and searching the optimal path from the created paths. *Initial value of movement direction* that is the direction to create path is calculated from the centre of area score which is sum of 9 cubes' vertical scores as follow in Table 2 around the vehicle. Therefore, candidate paths are created toward *initial value of movement direction* from start point as shown in Fig. 5.

Next, paths are created from start point. In this algorithm, 1st movement paths are created from start point, and next movement paths are created from the tip of previous movement paths. Direction of created path is not changed while cruising in the water. Therefore, paths are created in a straight line and increasing on every ascend point like a branch. However, it is high calculation cost that paths

Table 2 Divided depth of the survey cube.

No	Depth[m]	No	Depth[m]	No	Depth[m]
1	0-50	2	50-100	3	100-200
4	200-400	5	400-600	6	600-800
7	800-1000	8	1000-1250	9	1250-1500

Table 3 Yaw angle path pattern.

Number of gliding	Yaw angle						
1:	–90	–60	–30	0	+30	+60	+90
2:		–90	–45	0	+45	+90	
3, 4, 5:			–90	0	+90		

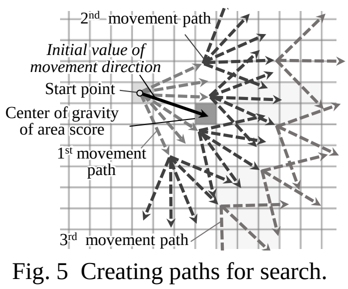


Fig. 5 Creating paths for search.

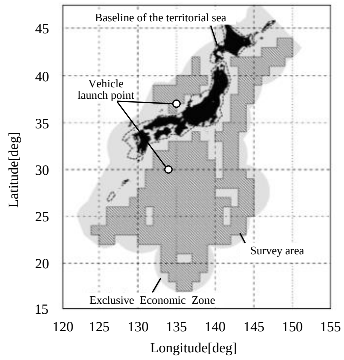


Fig. 7 Survey area based on Japan's EEZ.

are created in large quantities, and calculation power of the mounted PC in AUG is not high to reduce energy consumption. In this research, paths are created for 5 diving times, and the number of created paths is changed gliding angle and yaw angle for every diving times. The number of created path pattern are 1st movement is 3 gliding angle patterns and 7 yaw angle patterns, 2nd movement is 3 gliding angle patterns and 5 yaw angle patterns, and 3.4.5th movement is 3 gliding angle patterns and 3 yaw angle patterns. Yaw angle patterns are changed for every diving times as shown in Table 3, and gliding angle of path is changed to 8, 15, 20 degrees. When the paths have be created, the got scores as passing while 5 diving times is calculated by simulation, and it is assumed that paths while 5 diving times with the highest got score is the optimal paths as most important route.

The final part is moving on the optimal route and updating score of surveyed cubes. AUG moves on the only 1st movement path of optimal paths selected in the second part. When the movement of AUG to the 1st movement path is completed, score of cubes passed through AUG is updated as surveyed cubes. Lastly, the algorithm returns to the second part and searches optimal path again. By the above method, the algorithm search for optimal route that preferentially passes through survey cubes with the highest scores without high accuracy position.

3.2. *Cooperative behaviour*

A lot of AUGs need to be operated efficiently to survey for a wide range. It is necessary to select a route to maintain an appropriate distance between vehicles to avoid collision or expand the survey area. In this research, a simple rule of the reluctant cooperative behaviour is adopted. AUGs are required "not to disturb other vehicles".

When an AUG approaches another vehicle, each AUG estimates the expecting score to get and decides which one should give way. If the distance between the vehicle and centre of area score position is shorter, many high score cubes are near the vehicle. By contrast, if this distance is longer, many high score cubes are far away from the vehicle. In this method, the vehicle with the shorter distance between the vehicle and centre of area score position keeps route. Furthermore, the vehicle with the longer distance gives way by rotating *initial value of movement direction* which is direction to search path method to 90 degrees in opposite side to other vehicle's position.

The outline of cooperative behaviour is shown in Fig. 6. Vehicle A and vehicle B are shown as circle, and the centre of area score is shown as triangle. When two vehicles are approaching within a certain distance as shown in Fig. 6, two vehicles calculate the distance of centre of area score. In this situation, vehicle A gives way to vehicle B because vehicle A's distance of the centre of area score is longer than the vehicle B's distance. Vehicle A gives way by rotating *initial of movement direction* which is direction to search path method to 90 degrees in the direction opposite side to position of vehicle B, and search optimal path toward the direction after rotation.

Furthermore, when more than two vehicles are approaching, *initial value of movement direction* of all approaching vehicles are changed to the outside like a repulsing. If n vehicles are approaching a certain distance, i th vehicle's *initial value of movement direction* which is $Co.[i]$ is shown as follows:

$$Co.[i] = (360/n)\cdot(i-1) \quad (3)$$

4. Simulation of Underwater Operation

4.1. *Setting of simulation*

In this section, operative simulation of many AUGs for Japan's EEZ is conducted to verify the effectiveness of the proposed method and to estimate the number of vehicles required to achieve continuous and even observation of Japan's EEZ.

Survey area is set based on EEZ as shown in Fig. 7. The survey area consists of two unconnected areas, "Japan sea area" and "Pacific Ocean area". These areas are set removing baseline of the territorial sea and island from EEZ. The vehicle's launch positions are shown circle in Fig. 7.

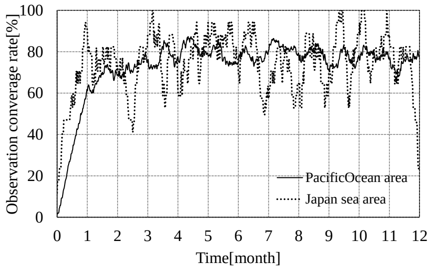


Fig. 11 Observation coverage rate.

The assigned score to the area and latest surfacing positions of all vehicles are managed by land base station. AUGs communicate to the land base station while sea surface to receive the position of other AUGs and send the data of survey seawater. The land station summarises received data of passing cubes and assigns score to area. This assigning score is reassigned after a certain period which is set 2 weeks after the score obtained in this simulation. The operational depth is set 1,500m, and the allowable approach distance between AUGs is set 50km. The simulation runs for a year, and the result of simulation is evaluated by an indicator of the observation coverage rate which is calculated by the number of areas passed within past 2 weeks.

4.2. *Result of simulation*

Figs. 8-9 show simulation results of Pacific Ocean area and Japan Sea area with every number of vehicles. Dashed line is the average observation coverage rate, plot of circle is the differential of observation coverage rate in Fig. 8. Fig. 10 shows tracks of simulation using 30 vehicles in Pacific Ocean area and 2 vehicles in Japan Sea area, for 1st -3rd month, 4th -6th month, 7th -9th month, 10th -12th month from the left. Leftmost figure which is tracks for 1st -3rd month shows that many vehicles expand to whole area over 3 months. Therefore, average observation coverage rate is calculated as an average value between 3rd -12th month after vehicles expanding whole area. Figs. 8-9 shows that observation coverage rate is increasing in proportion to increase the number of vehicles, and the rate converges to 90%.

In Japan Sea area, the rate converges to 90% when the number of vehicles more than 2 vehicles. Furthermore, it is necessary that the operating vehicle is more than 1 vehicle because when the vehicle fail, the survey of the area stops. From the above, the number of vehicles required for the survey of Japan Sea area is 2 vehicles. The average observation coverage rate is 77.2% using 2 vehicles in Japan Sea area.

In Pacific Ocean area, the rate converges to 90% when the number of vehicles more than 45 vehicles. The differential of observation coverage rate is wide variation between 30 and 35 vehicles as shown in Fig. 8. This variation show observation coverage rate decreasing significantly more than 30 vehicles.

From the above, the number of vehicles required to obtain a higher observation coverage rate with a smaller

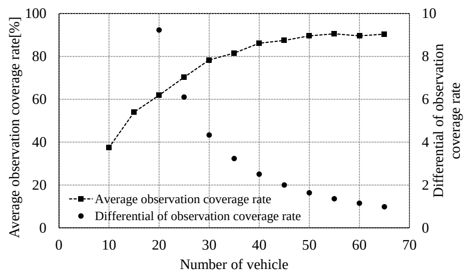


Fig. 8 Observation coverage rate of Pacific Ocean area.

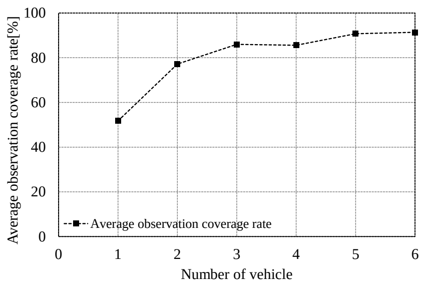


Fig. 9 Observation coverage rate of Japan Sea area.

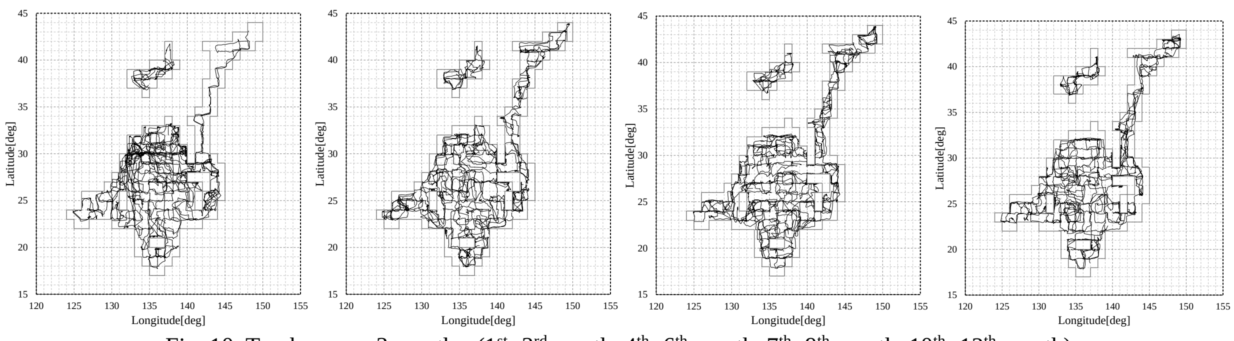


Fig. 10 Tracks every 3 months. (1st -3rd month, 4th -6th month, 7th -9th month, 10th -12th month)

number of vehicles is 30 vehicles. The average observation coverage rate is 78.3% using 30 vehicles in Pacific Ocean area.

Survey resolution is 1degree as shown Fig. 10 in this simulation. Observation coverage rate for a year is shown in Fig. 11. The survey has been continuously expanded to whole survey area, and it has confirmed that the constructed wide area observation system is effectiveness.

5. Conclusions

In this research, an observation strategy of a long-term wide-area sea area using a number of AUGs was proposed. Route search method from the surrounding situation was examined and wide-area ocean observation with AUGs was simulated to verify the effectiveness of the proposed method. The route search method was consisted of optimal path search algorithm and cooperative behaviour. The optimal path search algorithm was developed so that the AUGs pass preferentially through survey cubes with the highest scores without high accuracy position. High survey efficiency was achieved without the vehicle approaching excessively to adopt reluctant cooperative behaviour as swarm intelligence.

Finally, the effectiveness of the proposed method was verified, and the number of vehicles required for observation in Japan's EEZ was estimated by operative simulation.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 21H01552 and the Sasagawa Scientific Research Grant from the Japan Science Society.

References

1. Yan Huang, Jiancheng Yu, Wenmimg Jin, Yeteng Luo, Yan Li, "A Practical Path Tracking Method for Autonomous Underwater Glider Using Iterative Algorithm", Proceedings of OCEANS 2015-MTS/IEEE Washington, DOI: 10.23919/ OCEANS. 2015. 7404523, 2015
2. Edward Fiorelli, Pradeep Bhatta, Naomi Ehrich Leopnard, "Adaptive Sampling Using Feedback Control of an Autonomous Underwater Glider Fleet", Proceedings of 13th Int. Symp. on Unmanned Untethered Submersible Technology, 2003.
3. M. Arima, W. Sumino, A. Toyoda, A. Katsui, and T. Okuno, "Feasibility Study of an Underwater Glider with Independently Controllable Main Wings (1st repot) -Development of an Experimental Underwater Glider-", Journal of the Japan Society of Naval Architects and Ocean Engineers, vol. 4, pp.31-37, 2006 (in Japanese).
4. K. Kobatake, M. Arima, "Development of the Autonomous Control System for Autonomous Underwater Glider with Independently Controllable Main Wings", Journal of the Japan Society of Naval Architects and Ocean Engineers, vol.33, pp.191-197, 2021 (in Japanese).

Authors Introduction

Ms. Kanako Kobatake

She received her Master's degree from the Department of Marine Science and Technology, Tokyo University of Marine Science and Technology, Japan in 2019. She is currently a Doctoral course student in Osaka Prefecture University, Japan.

Dr. Masakazu Arima

He is a Professor of Osaka Prefecture 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.

Underwater Acoustic Communication using QPSK Modulation Method

Yuya Nishida[*1]
Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196, Japan

Yuichiro Uemura[*2]
Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196, Japan

Rikuto Tanaka[*3]
Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196, Japan

Kazuo Ishii[*4]
Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196
*E-mail: y-nishida[*1], ishii[*4]@brain.kyutech.ac.jp, uemura.yuuichirou544[*2], tanaka.rikuto288[*3]@mail.kyutech.ac.jp*
https://www.kyutech.ac.jp

Abstract

Acoustic communication has a lot of influence of noises, and difficulties in sending and receiving data correctly because of the noise from robot's thrusters and other acoustic devices. To improve the acoustic communication performance, this research developed acoustic communication system against the thruster's noise. Communication message includes not only data (payloads) but also sync signal and error correction bytes, and the message is converted to acoustic wave by QPSK modulation which four phase data represent two bits.

Keywords: Acoustic communication, QPSK modulation, Error correction, Noise canceling.

1. Introduction

Acoustic transmission which its baud rate is only 1/1000 of conventional Wi-Fi communication has less attenuation than those of radio or light and is communication between underwater robots. However, acoustic communication has a lot of influence of noises, and difficulties in sending and receiving data correctly because of the noise from robots and surroundings. The noise includes ambient noise from the wave and marine organisms and self-noise from the robot's acoustic devices and thrusters. Ambient noise level is relatively low because underwater robots are operated only in calm sea conditions of WMO sea state code 3 or less, and not many marine organisms make sounds that interfere with communication. High frequencies sound of several hundred kHz is used for acoustic devices such as observation devices and sensors due to require high resolution measurement at a short distance. The noise from the acoustic devices gives little influence on acoustic transmission using sound of several tens of kHz that because the frequency band of the oscillator for acoustic transmission is narrow. On the other hand, the noise from the thruster is high value in a wide frequency

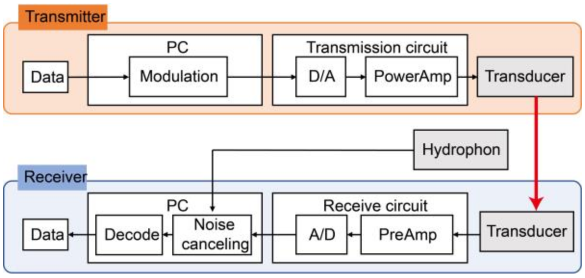


Fig.1 System architecture for acoustic communication

band on the order of several kHz to several tens of kHz and acoustic communication may fail due to the noise from thruster [1].

Acoustic communication with error correction has been developed for improving communication success rate [2-3], but robots is difficult to receive acoustic command when its cruises by full thrust. For reliable and high-speed acoustic communication in noisy situations, this research developed acoustic communication using QPSK modulation method and noise canceling method. This paper explains our acoustic communication and shows experimental results using our system.

2. Acoustic communication system

2.1. *Hardware configuration*

Our acoustic communication system consists of transmitter and receiver units, and the units have a transducer which can transmits and receives acoustic wave. The transducer consists of a PC for modulating data, a circuit for converting digital data to amplified analog data, and a transducer for transmission. The receiver has a PC for recording data, a circuit for amplifying and converting analog data to digital data, a transducer, and a hydrophone for receiving thruster noise. Transceiver and receiver units can have the same structure by using the transducer, but the unit is separated into two structures to make its smaller. Thus, our acoustic communication system can transmit in only one direction.

Sonar head of FURUNO fish finder is used as transducer for transmitter and receiver units. The frequency characteristic is measured because nobody knows it specification. Figure 2 shows frequency characteristics of sonar head for transducer. The transducer has the lowest impedance at 50kHz and 200kHz, and the phase at that frequency was a practical value. Acoustic communication of higher frequency has high baud rate,

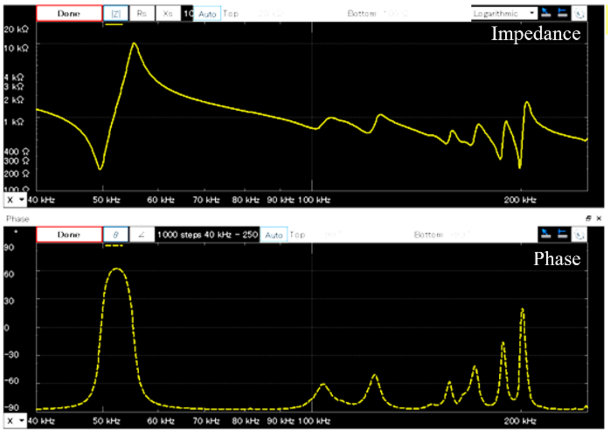
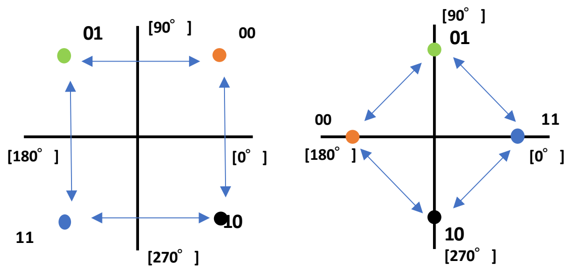


Fig. 2 Frequency characteristics of sonar head for transducer



(a) QPSK (b) π/4 shift QPSK

Fig.3 Phases for QPSK

but its acoustic wave doesn't reach far. To develop a practical acoustic communication that can transmit up to several hundred meters, this research used frequency of 50kHz.

2.2. *Communication format using QPSK*

Communication message in our format is separated into header, payload, error correction code. In this research, synchronization of acoustic wave is very important due to that baud rate of acoustic communication is improved by modulation. The header in the message is represented chirp signal for synchronization. Accuracy of synchronization by time correlation is decreased by the noise when normal sine wave is used as synchronization signal. Even if a part of acoustic wave is transformed by the noise, accuracy of time correlation doesn't decrease so much because frequency of chirp signal changes over time. Payload in the message is used for data transmission, and it is modulated by QPSK (**Q**uadrature **P**hase **S**hift **K**eying) modulation which two bits are

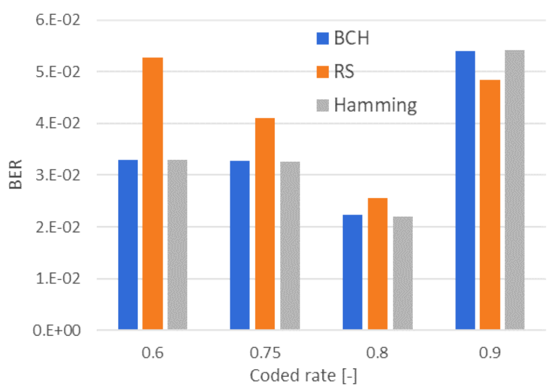


Fig.4 Bit-error rate of each correction method

represented in four quadrants, as shown in Fig.3. QPSK modulation is more robust to the noise than ASK (**A**mplitude **S**hift **K**eying) and can code by single frequency unlike FSK (**F**requency **S**hift **K**eying). This research used $\pi/4$ shift QPSK which uses phases of 0, 90, 180, 270 deg. Even if acoustic wave is perfectly synchronized, decoded payload may include error bit by the noise. Error correction code is added to the end of the message to improve acoustic communication quality. When code rate which is ratio of correction code bit to data bit is high, error correction code corrects many data bit, but communication baud rate is low. The correction performance depends on the correction method. This research evaluates correction performance of BCH code, RS (Reed-Solomon) code and Hamming code by the simulation using additive white gaussian noise. Figure 4 shows BTR (**B**it-**E**rror **R**ate) after error correction by each correction methods. The smaller code rate other than 0.9, the smaller BTR of each method. Correction performance at code rate of 0.9 was low because correction code bit includes most error bits, and there was almost no difference in performance each method at code rate of 0.8 and 0.9. This research used RS code which is high performance against burst error, and RS code was used for communication between underwater robot [4-5].

2.3. *Noise canceling method*

To improve decoding accuracy at receiver unit, the noise from robot's thruster is removed from received acoustic wave by LMS (Least Mean Squares Filter) adaptive filter, as shown Fig.5. The filter adapts the transfer function according to optimization algorithm. In this research, the sound data from the hydrophone for listening thruster noise enters to adaptive optimization block, and thruster noise is removed using filter factor from adaptive algorithm. Figure 6 shows received acoustic wave with thruster noise and acoustic wave after filtering by LMS adaptive filter. The received acoustic wave has a waveform close to a sine wave by applying the filter.

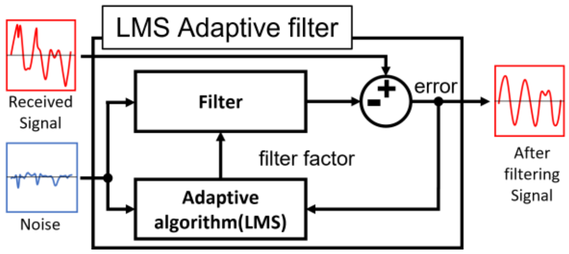


Fig.5 LMS adaptive filter

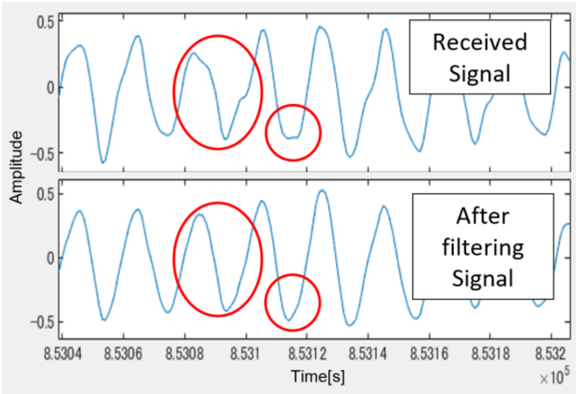


Fig.6 Received acoustic wave before and after filtering

3. Wet test

3.1. *Thruster noise measurement*

To verify the effect of thruster noise on the acoustic communication, the noise during thruster drive was measured. RoboPlus Hibikino thruster which was located on small experimental tank was used for the measurement, and the noise was measured when three levels of thrust to the maximum output were generated. Figure 7 shows power spectrum obtained by FFT analysis of measured noise. Sound pressure of the noise depends on the thrust from the thruster, and the noise is in the frequencies from 5kHz to 35kHz. The thruster made the noise of 110dB at the frequency of 14kHz when it produced the thrust of 24N. The frequency of maximum noise from the thruster is different from the frequency of

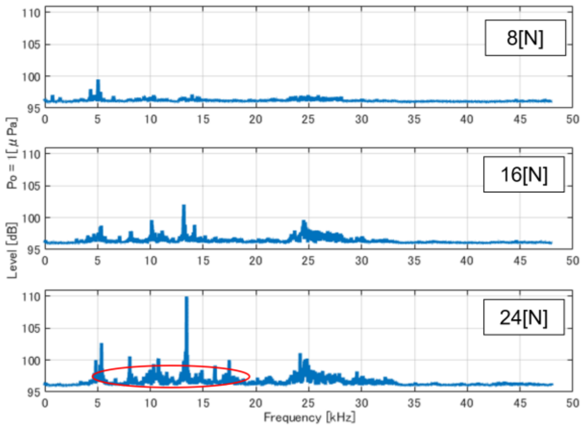


Fig.7 Power spectrum of thruster noise

the acoustic communication (50kHz), but the waveform of acoustic communication may be deformed by the peak noise and the wideband noise. If the waveform is deformed, decoding results of QPSK modulated acoustic wave includes several error bits.

3.2. *Communication performance evaluation*

To evaluate the performance of developed the acoustic communication system, BER of the communication in the presence of the noise when the thruster produced the thrust of 24N was measured. At this time, the sound pressure of acoustic wave produced by the transceiver was set that SNRs (**S**ignal **N**oise **R**atio) to the thruster noise were 2dB, 5dB, 10dB, 15dB and 20dB. The message that its payload were the data of 5,001bits and the code rate of RS is 0.9 is used for the evaluation. Figure 9 shows the BERs of each communication vs. SNR. Our communication method that used QPSK with error correction and noise cancel had a BER of 1/3 compared to the normal QPSK modulated communication, at the SNR of 2dB. At SNR above 10dB, there was almost no BER in the communications with normal and the error correction, and the communication of our method included some BER. This means that the noise canceling method using LMS adaptive filter reduces communication performance a little in an environment with almost no noise. However, the increase in BER by our method is very small, and there can be no environment with almost no noise when the underwater robot is used. Thus, the communication using QPSK with

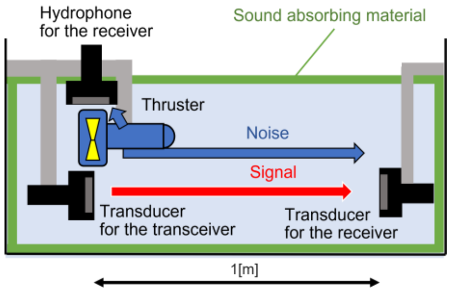


Fig.8 Setup for communication performance evaluation

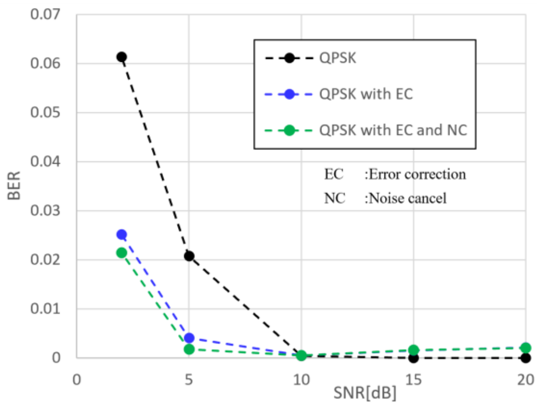


Fig.9 the BER of each communication vs. SNR

error correction and noise cancelling method is the effective communication for the underwater robots.

3.3. *Image transmission test*

To verify whether a large size data can be transmission, the photo image was transmitted by using our acoustic communication system. Transducers for receiver and transceiver was installed on the tank which its diameter is 6m and height is 1.5m, the distance between the transducers was 2m and 5m. Artificial turf was installed behind the thrusters to reduce the effect of multipath of the acoustic wave. This test didn't use noise cancelling method by using LMS adaptive filter because there was no noise resource in the tank. Figure 10 shows the original image and images transmitted 2m and 5m ahead. The communication in the distance of 2m had no BER and decoded image was the same as the original image. Decoded image from the communication data in the distance of 5m had the BER of 0.20 and included different colors on some pixels. The reason why the BER increased in the communication at the distance of 5m is considered to be low sound pressure of transmitted acoustic wave. Signal level of the acoustic

Fig.10 Images before and after image transmission

communication closed the noise level from acoustic wave multipath, due to attenuation of the acoustic wave. Then, the decoder made some error bits due to a decrease in the SNR.

4. Conclusions

This paper proposed the acoustic communication using QPSK modulation to improve the communication performance in noise situation. The message in our communication consists of chirp signal for synchronization of acoustic wave, payload for user and error correction code based on Read Solomon code. The payload was modulated by QPSK modulation which two bits are represented in four quadrants before transmission. For decreasing error bit during decoding, the noise from the thruster is removed by LMS adaptive filter from received acoustic wave before. Our communication had a BER of 1/3 compared to the normal QPSK modulated communication at the SNR of 2dB, and can transmit photo image to 2m away without data loss.

References

1. Jason D. Holmes and William M. Carey, An overview of unmanned underwater vehicle noise in the low to mid frequencies bands, Proc. of Meeting on Acoustic, Vol.9, Issue 1, August 2010.
2. Yuya Nishida, et al., Virtual mooring buoy "ABA" for multiple autonomous underwater vehicles operation, Journal of Robotics and Mechatronics, Vol.28, No.1, pp.86-90, 2016
3. Ohki, T., et al., Development and testing of an unmanned surface towing system for autonomous transport of multiple heterogeneous underwater vehicles for seafloor survey, Marine Technology Society Journal, Vol.54, No.5, pp.61-71, 2020
4. Yuya Nishida, et al., Development of an autonomous buoy system for AUV, Proc. of MTS/IEEE OCEANS, 10.1109/OCEANS-Genova.2015.7271614, Genova, 2015
5. Yuya Nishida, et al., Unmanned surface vehicle for managing parallel cruising of a multiple AUVs, Proc. of MTS/IEEE Techno-Oceans, OCEANS, 10.1109/OCEANSKOBE.2018.8559319, Kobe, 2018

Authors Introduction

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.

Mr. Yuichiro Uemura

He graduated from master's program at Graduate School of Life Science and System Engineering, Kyushu Institute of Technology, Japan in 2021. His reseach area is application for underwater robot.

Mr. Rikuto Tanaka

He completed the production and electrical system technology course at Shikoku Polytechnic 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.

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

Ultrasonic Cleaner using Two Transducers for Ship Hull Cleaning Robot

Yuya Nishida
Dept. of Human Intelligence Systems, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka, 808-0196, Japan
E-mail: y-nishida@brain.kyutech.ac.jp

Toshihiro Matsumura
Dept. of Human Intelligence Systems, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka, 808-0196, Japan
E-mail: matsumura.toshimune399@mail.kyutech.jp

Kazuo Ishii
Dept. of Human Intelligence Systems, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka, 808-0196, Japan
E-mail: ishii@brain.kyutech.ac.jp

Abstract

Fuel consumption of the ship gets worse by barnacles and stains put on its bottom, so that the ship hull should be cleaned regularly. Usual methods such as cleaning by special divers and cleaning after pulling up to the dock are not used frequently due to those high cost and the heavy burden on cleaners. We proposed a cleaning method using underwater vehicle with brushes, however there is a possibility to remove paints on the hull. To improve cleaning performance of the underwater vehicle, the ultrasonic cleaner is developed by using cavitation occurred by sound waves of the acoustic transduces. The cleaner generates sound waves with sound pressure of 0.2MPa or more required to occur cavitation at the point where sound waves from two transducers overlap. In experimental results, the cleaner occurred cavitation enough to make two holes in the aluminum foil located 50mm away from transducers.

Keywords: Ultrasonic Cleaner, Ship Hull Cleaning, Cavitation

1. Introduction

Various marine organisms such as sea lettuce and barnacles adhere to ships and make problems to ships that the weight and water flow resistance of the ship increase, which reduces thrust and increases fuel consumption. It is reported that if the bottom of the ship is not cleaned, the fuel efficiency will deteriorate by about 20% in one year after leaving the dock, and if it is not cleaned after that, the fuel efficiency will deteriorate by up to about 50% [1]. To prevent the adhesion of barnacles and stains put on its bottom, the ship hull should be cleaned regularly. Usual methods such as cleaning by special divers and cleaning after pulling up to the dock are not used frequently due to those high cost and the heavy burden on cleaners. We proposed a cleaning method using underwater vehicle with brushes [2], however there is a possibility to remove paints on the hull. To improve cleaning performance of the underwater vehicle, the ultrasonic cleaner is developed by using cavitation occurred by sound waves of the acoustic transduces.

In this study, we propose ship bottom cleaning using ultrasonic cleaning in order to improve the cleaning ability. Ultrasonic cleaning can control the cleaning position and cleaning strength, and since it is non-contact cleaning, it is considered that the coating is less likely to peel off than the method using a brush. However, ultrasonic cleaning for cleaning the bottom of the ship has not been performed, therefore, the purpose of this study is to investigate ultrasonic cleaning and evaluate its performance for adapting to ship bottom cleaning. We investigate the occurrence of cavitation using an ultrasonic array and discuss the possibility of the ultrasonic cleaning of ship hulls.

2. Ultrasonic cleaner for ship hull

2.1. *Cavitation*

When the pressure of water is lowered with keeping constant temperature, the pressure drops below the saturated vapor pressure and the state changes to a gas. The phenomenon bubbles generated at this time is called cavitation, occurs especially in rotating marine propeller and pump impeller [3][4]. In order to generate cavitation by ultrasonic waves, it is necessary to give water a sound pressure that is lower than the saturated vapor pressure and negative sound pressure is generated in the half cycle of ultrasonic decompression and cavitation occurs (see Fig. 1).

In order to generate cavitation in water, the saturated vapor pressure at an arbitrary temperature is needed. The saturated vapor pressure can be approximated by Wagner's vapor pressure equation in eq. (1) and (2).

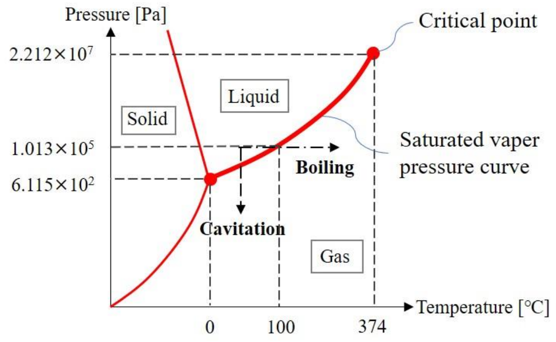
$$p_{ws} = p_c \cdot \exp\left(\frac{A\cdot x + B\cdot x^{1.5} + C\cdot x^3 + D\cdot x^6}{1-x}\right) \quad (1)$$

$$x = 1 - \left(\frac{T}{Tc}\right) \quad (2)$$

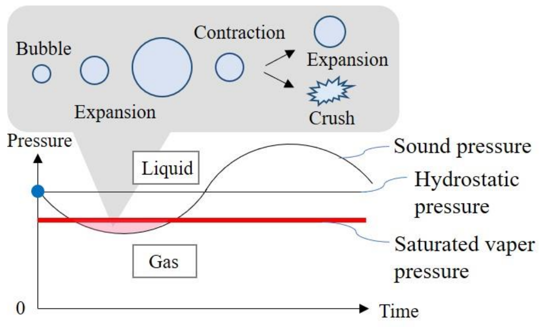
Here, p_{ws} : Wagner's strict vapor pressure [Pa], p_c: critical pressure [Pa], T: temperature [K], T_c: critical temperature [K], A: -7.6451, B:1.4583,C: -2.7758,D: -1.2330 are coefficients for approximation curve. This equation shows an approximate equation of the saturated vapor curve by changing the temperature starting from the pressure at the critical point and the critical temperature. Based on eq. (1), sound pressure with an amplitude of about 0.2 [MPa] is required for cavitation to occur under water at a standard temperature of 25 [C]. In addition, the sound pressure required for cavitation increases as the water depth increases.

2.2. *System architecture of ultrasonic cleaner*

In order to generate cavitation on the surface of the bottom of the ship, we propose a cleaning system that generates cavitation at surface positions with synthetic



(a) Water phase diagram



(b) Cavitation generation process

Fig.1 Water phase diagram and cavitation generation process using ultrasonic wave.

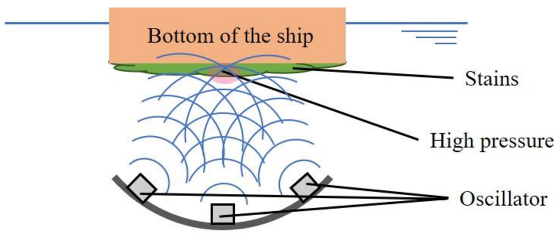


Fig.2 Concept of ultrasonic cleaner for shup hull. Ultrasonic transducer array is used to enhance the power

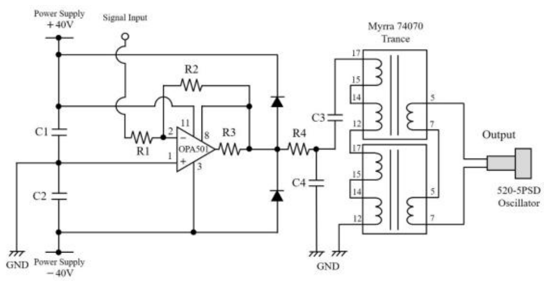


Fig.3 Amplifier circuit for ultrasonic transducers. The voltage of transducers are 300 [Vp-p].

waves using multiple oscillators as shown in Fig. 2. This ultrasonic array does not require a standing wave, and it is possible to obtain as high a sound pressure as the number of oscillators.

FURUNO's 520-5 PSD is used to configure the ultrasonic cleaning system. The specifications of the oscillator are 50-200 [kHz], 600 [W]. We used mainly 50 [kHz], which frequency is within the band 28 to 100 [kHz] often used in general ultrasonic cleaning. The amplifier circuit shown in Fig. 3 was designed to emit high-voltage and high-frequency sound waves using the oscillator. Two regulated power supplies of ± 40 [V] were used as the power supply, and the Tektronix AFG3022 function generator was used as the control voltage signal. To amplify the voltage, TI OPA541 amplifiers and Myrra 74070 trans with 50 [kHz] were used in series. Capacitors C1, C2 and C3 were set to 470 [μF], C4 was set to 0.1 [μF], resistance R1 was set to 1 [kΩ], R2 was set to 10 [kΩ], and R3 was set to 0.1 [Ω]. By using this amplifier circuit, it is possible to obtain an output with a maximum current consumption of 0.9 [A] and 300 [Vp-p] with an input signal of up to 9 [Vp-p].

3. Evaluation experiments of developed cleaner

3.1. *Basic property of ultrasonic transducer*

In order to investigate the basic property of the 520-5 PSD transducer, an experiment shown in Fig. 4 was conducted. The test tank has a size of 1650 × 540 × 600 [mm] and the water level at 550 [mm], and an artificial lawn was laid on the wall surface to suppress the influence of echo. The sound pressure was measured using the TC4013 hydrophone, and the hydrophone and transducer were set at a water depth of 250 [mm]. The measured distance L was set to be at 25, 50, 75, 100, 200, 300 [mm], respectively. The frequency of the transducer is set to 50 [kHz], and the input voltage is 50, 150, 200, 250, 300 [Vp-p] at each length.

Figure 5 shows the measurement results of attenuation characteristics for each input voltage. The higher the input voltage, the higher the sound pressure, and the sound pressure decreased as the distance L increased for all input voltages. The sound pressure required to generate cavitation exceeds 0.2 [MPa] when the input voltage is 300 [Vp-p] with L of 25, 50 [mm], and the input voltage is 250, 200 [Vp-p] with the L was 25 [mm].

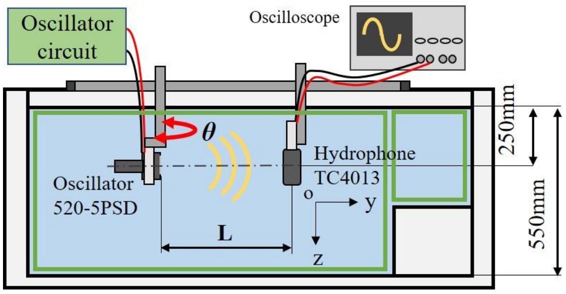


Fig. 4 Arrangement for basic property measurement of the transducer 520-5PSD.

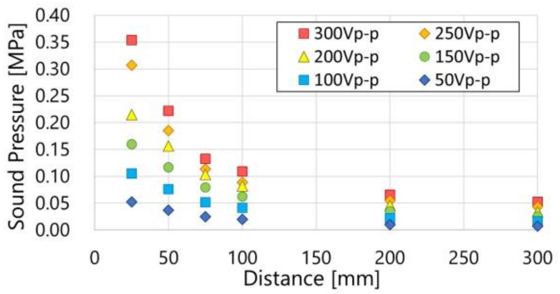


Fig. 5 The relationship between the length L and sound pressure by changing the input voltage.

3.2. *Cavitation using an ultrasonic transducer*

In the experiment in 3.1, it was found that the sound pressure at which cavitation occurs exceeds 0.2 [MPa] when the input voltage is 300, 250, 200 [Vp-p]. In order to verify the cavitation effect using single transducer, an experiment shown in Fig. 6 was carried out. The aluminum foil is fixed to a jig so that the vibration surface of the transducer could be seen vertically. The area of the aluminum foil was 68 x 55 [mm] to secure the diameter of the vibrating surface and the distance to 50 [mm], and colored red to check cavitation effect easier. The installation water depth of the oscillator with the aluminum foil fixed was set to 250 [mm]. The input voltage of the oscillator was 150, 200, 250, 300 [Vp-p], and the transducer was operated for 3 minutes each to qualitatively observe the state of the aluminum foil.

The state of the aluminum foil is shown in Fig. 7. When the input voltage over 200 [Vp-p] was applied, torn holes could be observed, but not in 150 [Vp-p]. The hole lengths generated by the input voltage of 200, 250, 300 [Vp-p] are 14, 24, 33 [mm], respectively, and it can be seen that the higher the input voltage, the larger the holes. It can be seen that the holes were found when the aluminum foil was affected by cavitation and possible to generate sufficiently strong cavitation even with a single oscillator.

From the basic property of the transducer in Fig.5, the sound pressure exceeding 0.2 [MPa] at the measurement point of 25 [mm] occurs with the input voltages of 200, 250, 300 [Vp-p]. From the experimental results in Fig.7, the cavitation occurs when the input voltage was 200, 250, 300 [Vp-p], which shows good agreement.

In this research, an ultrasonic cleaning system with the combined waves using an ultrasonic array is proposed, so that the directivity of the transducer is also important. In order to investigate the directivity of the oscillator, an experiment was conducted with the condition that transducer was fixed with a jig to rotate around the vibration surface, and the distance L was 50, 75, 100 [mm]. The sound pressure was measured with the angles between the y-z plane 0, 15, 30, 45, 60, 75, 90 [deg] with input voltage of 300 [Vp-p], respectively.

Figure 8 shows a directivity of the sound pressure for each angle. It can also be seen that the sound pressure decreased as the angle increased and the sound pressure with the surface angle 30 [deg] becomes half pressure of 0 [deg] (-6 [dB]), which it is almost the same performance as the directivity angle of the transducer specifications. When the system consists of an ultrasonic array with two oscillators, the sound pressure generated by each oscillator should be at least 0.1 [MPa] at a distance of 50 [mm], ± 30 [deg]. At 75 [mm], cavitation will occur within the range of ± 15 [deg].

3.3. *Cavitation using two ultrasonic transducers*

In order to estimate the cavitation strength of an ultrasonic array, the simulations using the property obtained in the previous section were carried out. Figure 9 (a) shows the sound pressure distribution of a single transducer at the input voltage of 300 [Vp-p]. Figures 9 (b) to (d) are the pressure distributions by combined waves when the distance between the centers of the two transducers is 100 [mm] and the relative angles are 30, 60, 90 [deg] and the phase difference was 0 [deg]. From the simulation results, the maximum positions of the combined wave at angles 30, 60, and 90 [deg] were 104, 70, and 47 [mm] and the sound pressure at those positions exceeded 0.2 [MPa], therefore, there is the possibility to generate cavitation by combined waves.

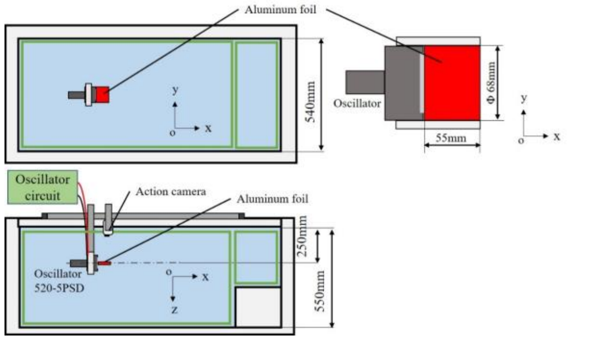
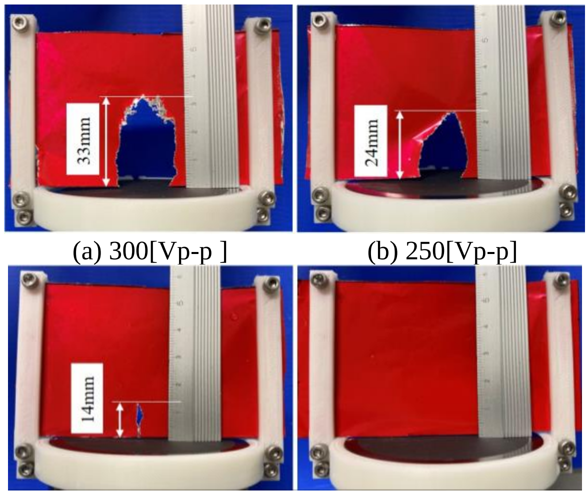


Fig. 6 The cavitation experiment using single transducer.



(a) 300[Vp-p] (b) 250[Vp-p]

(c) 200[Vp-p] (d) 150[Vp-p]

Fig.7 The results of cavitation experiments in the test tank, when the input voltage changes from 150[V] to 300 [V].

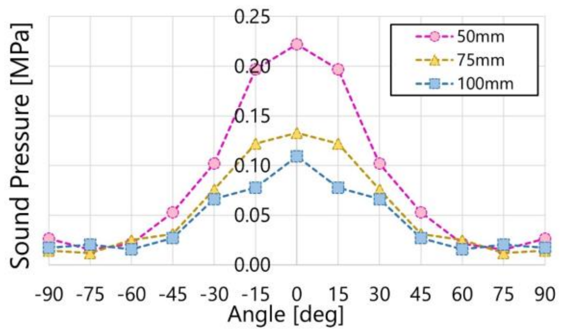
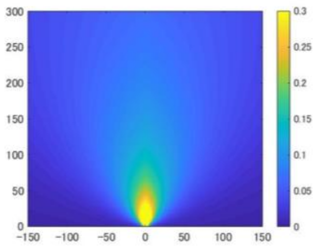


Fig.8 The directivity property of the transducer 520-5PSD with the angle θ from 0 [deg] to +/- 90 [deg]. With the angle of 30 [deg], the pressure becomes half.

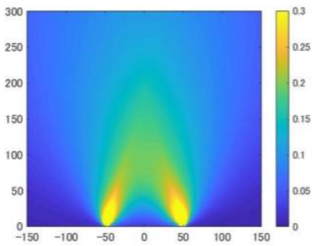
In order to verify cavitation with an ultrasonic array using two oscillators, an experiment was conducted with a configuration as shown in Fig. 10. The two amplifier circuits are used to control the phase of each transducer. The arrangement of the ultrasonic array was such that the distance between the centers of the transducers was constant at 100 [mm], and the angles θ between the oscillators were 30, 60, and 90 [deg], respectively. The distance L along the x-axis from the center of the distance between the centers of the two oscillators was set to be constant at 50, 75, 86, 100, 187 [mm]. The distances 50, 86, 187 [mm] are the focal lengths geometrically determined at the angles 90, 60, 30 [deg], respectively. The sound with the frequency of each oscillator was 50 [kHz] and the input voltage 300 [Vp-p] is used.

Figure 11 shows the results in which the sound pressure distributions of the combined sound waves along x-axis for each angle between transducers 30, 60, 90 [deg], respectively. The phase-adjusted sound pressure was higher than no phase-adjusted. The sound pressure required for cavitation generation exceeded 0.2 [MPa] between 50 and 100 [mm] in the phase-adjusted sound pressure of the three angles.

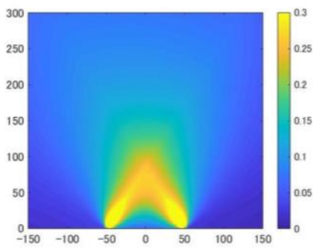
In order to verify the cavitation effect by the ultrasonic array, we conducted an experiment using aluminum fiols. The aluminum foil was fixed so that the vibration surface



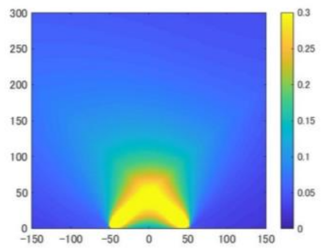
(a) One transducer



(b) Two transducers, 30[deg]



(c) Two transducers, 60[deg]



(d) Two transducers, 90[deg]

Fig. 9 The simulation results are compared using the property obtained from basic experiments.

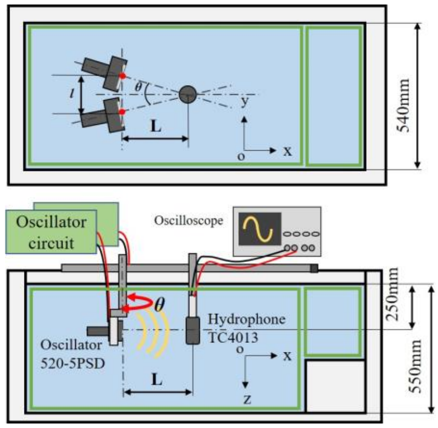
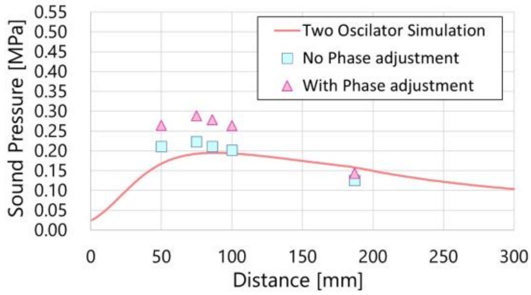
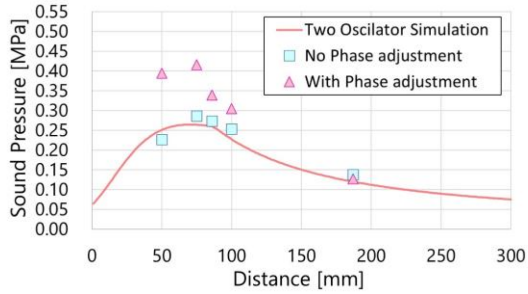


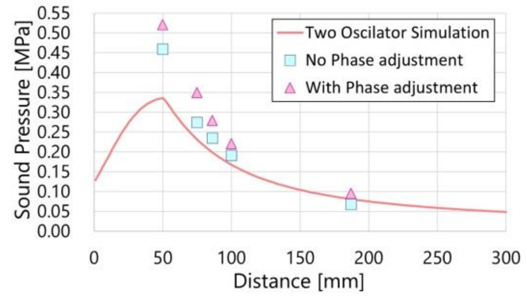
Fig. 10 The experimental setup for the transducer array.



(a) Relative angle: 30 [deg]



(b) Relative angle: 60 [deg]



(c) Relative angle: 90 [deg]

Fig. 11 Experimental results using two transducers with relative angles 30, 60, 90 [deg].

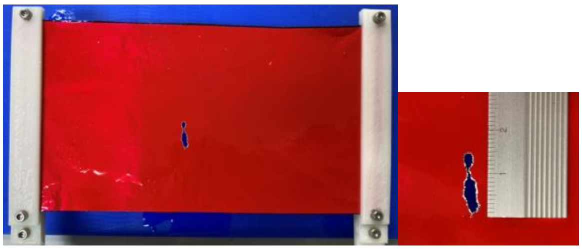


Fig. 12 Cavitation effect using two transducers.

of the oscillator becomes vertically. The area of the aluminum foil was 170 x 100 [mm] so as to cover the arrangement of the two oscillators. The position of the aluminum foil was set 30 [mm] away, the water depth of the aluminum foil was set to 250 [mm], which was the same as the arrangement of the transducers. The arrangement of the transducers was such that the distance between the centers of the transducers was constant at 100 [mm], and the relative angles θ were 30, 60, and 90 [deg], respectively. The input voltage of the oscillator was 300 [Vp-p] and given for 3 minutes.

Figure 12 shows the state of the aluminum foil when the arrangement angle of the oscillator is changed. The holes and cracks were confirmed in the aluminum foil at a position of about 35 [mm] from the center of rotation of the oscillator at all angles. However, since the holes and cracks of 30, 60 [deg] are far from the center of the aluminum foil, it is considered that they were generated by the cavitation effect of a single transducer. The crack of angle 90 [deg] is near the center of the aluminum foil, it is considered that it was generated by the combined wave of the ultrasonic array. The sound pressure of 0.2 [MPa] where cavitation occurs between 50 and 100 [mm] at every angle exceeded 0.2 [MPa], but in this experimental result, there is no holes or cracks.

4. Conclusions

In this paper, we have described the ultrasonic cleaner for the ship hull. The cleaner generates sound waves with sound pressure of 0.2MPa or more required to occur cavitation at the point where sound waves from two transducers overlap. In experimental results, the cleaner occurred cavitation in 50mm away from transducers.

References

1. Yokoi Koji, “On the Influence of Ship's Bottom Fouling upon Speed Performance”, Research studies of Toyama National College of Maritime Technology (37) , p.17～27, 2004
2. A.A.F.Nassiraei, T. Sonoda, K. Ishii, “Development of Ship Hull Cleaning Underwater Robot”, 2012 Fifth International Conference on Emerging Trends in Engineering and Technology, pp. 157-162, 2012.
3. K. Nakamura, Ultrasonic wave, Lecture note on Ultrasonic engineering (8), pp.182-183, Corona, 2001 (in Japanese).
4. Hammitt F.G., Cavitation and multiphase flow phenomena, Mc Graw Hill Inc., 1980.

Acknowledgement

This research is supported by Grant-in-Aid for Scientific Research(B) 18H01643.

Authors Introduction

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.

Mr. Toshihiro Matsumura

He received Master degree 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. 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

Motion Control of a Ship Hull Cleaning Robot

Hyoga Yamamoto
ROBO SCIENCE CO.,LTD.
1-8 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka, 808-0135, Japan
E-mail: n898012h@mail.kyutech.jp

Yuya Nishida
Dept. of Human Intelligence Systems, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka, 808-0196, Japan
E-mail: y-nishida@brain.kyutech.ac.jp

Takayuki Matsuo
National Institute of Technology, Kitakyushu College
5-20-1 Shii, Kokuraminami, Kitakyushu, Fukuoka, 802-0985, Japan
E-mail: matsuo@kct.ac.jp

Kazuo Ishii
Dept. of Human Intelligence Systems, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka, 808-0196, Japan
E-mail: ishii@brain.kyutech.ac.jp

Abstract

As the fuel cost will rise continuously, the reduction of fuel consumption and CO2 emissions are required more severely in the transportation industry. To achieve more efficient ship transportation, the prevention of marine biofouling is inevitable. Anti-fouling paint is effective to prevent biofouling to the ship hulls like barnacles, however, even the painted hull acquires slime-like biofouling caused by marine alga on its surface easily. In general, the cleaning of the ship hull is carried out during inspection in dockyard once a year or by divers, which are high costs and high risk task. Frequent cleaning during the ship berthing is desirable to keep good fuel efficiency, and if possible, ships can keep good fuel efficiency that contributes the reduction of costs and CO_2 production. As a solution for this issue, we have been introducing underwater robots for ship hull cleaning. In this paper, the motion control system of the robot is described.

Keywords: Under Water Robot, Ship Hull Cleaning, Attitude Control

1. Introduction

Global warming caused by greenhouse gases is getting worse, and in recent years there has been a growing demand for reduction of greenhouse effects in various fields. In addition to reducing CO_2 emissions, there is an increasing demand for reductions in fuel consumption due to soaring fuel prices.

In order to satisfy these demands, it is also necessary to make the ship transportation efficient and reduce the ship resistance.

The hull of the ship is always in the sea, and various marine organisms adhere to it for long time. The adhesion of living organisms increases ship resistance and affects the fuel efficiency of ships. It is reported that the fuel efficiency of organisms attached to the hull of a ship deteriorates by about 20% in the first year, and when left uncleaned, it deteriorates by up to about 50% in worst cases [1]. Antifouling paint is effective in preventing the adhesion of organisms such as barnacles, but slime-like organisms due to algae adhere to the surface of the painted hull. In addition, the effect of such paints diminishes over time, so it is necessary to reapply them regularly.

Usually, the method of cleaning the hull of the ship is to lift the ship on land like a dockyard, or to clean it by a diver.

Since docking a ship to the special facility like the dockyard is a large-scale work, it is generally performed about once a year, where cleaning and inspection are performed. However, in order to maintain and improve efficient fuel consumption, it is desirable to clean frequently even while moored. The underwater cleaning operation in the ports by divers is dangerous and the cleaning cost is high.

To solve these problems, we have been developing the robot that can regularly clean the hull of ship in water. This paper describes the motion control of the ship hull cleaning robot.

2. Ship Hull Cleaning Robot

2.1. *Features of the robot*

We have developed ship hull cleaning robots [1]-[3], and the latest robot is shown in Fig.1. The robot has a length of 0.8 [m], a height of 0.4 [m], a width of 0.3 [m], and a weight of about 30 [kg]. The robot has two pressure-resistant cylinders arranged on both sides of the brush unit in the middle. These are arranged so that every equipment is fit inside the metal frame including 6 thrusters. The robot is remotely operated vehicle (ROV) and designed to be the center of gravity and buoyancy as close as possible in order to adjust the posture of robot to the target ship hull angle, and the maximum operation depth is 30 [m]. The main power cable of the robot is AC200 [V] and data is transmitted using PLC (Power Line Communication).

2.2. *System architecture*

The robot system architecture is shown in Fig. 2. PLC is used for communication between the robot and the remote computer for control. The conventional PLC

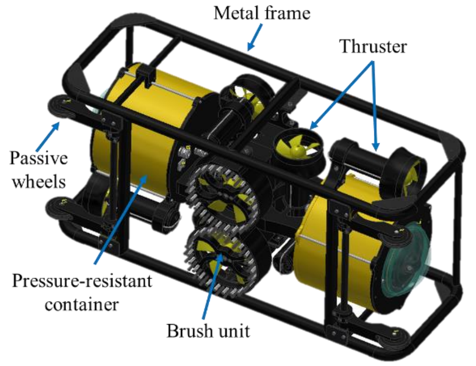


Fig. 1 The ship hull cleaning robot (4th version)

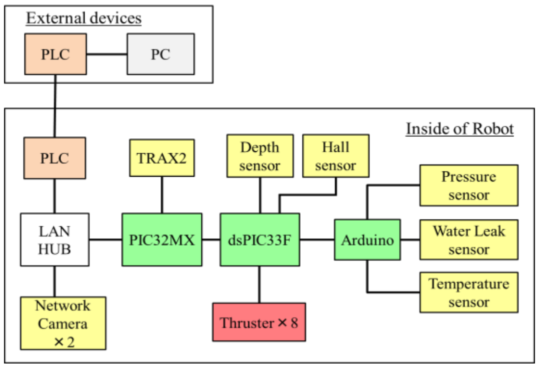


Fig. 2 System architecture

accept 200 [m] length cable and communicate at 200 [Mbps] that is enough speed to transmit two video camera images in the front and rear cylinders of the robot to the remote computer at the speed of 30 [fps].

The pressure-resistant cylinder in the front is mainly composed of power supply circuits such as AC/DCs and DC/DCs. The rear cylinder of the robot consists of sensors and circuits for communication and control. TRAX2 is used as AHRS (Attitude and Heading Reference System) for attitude detection and control of the robot. Magnets are mounted on the front and rear passive wheels, and the hall sensor inside the robot determines whether the passive wheels in the brush side are touched to the hull of the ship. In addition, there is a temperature sensor, barometric pressure sensor, and leak detection sensor inside the cylinders for monitoring whether there is any abnormality inside the robot.

Two MPUs, PIC32MX and dsPIC33F, are used for the communication and control boards. The PIC32MX is mainly used for communication with the ground PC, acquisition of sensor values from the TRAX2, and determination of control and operation quantities. The dsPIC33F mainly acquires other sensor values and generates PWM for each thruster. In addition, data is exchanged between the two by serial communication.

Each of the front and rear cylinders is equipped with a network camera.

3. Kinematics

Figure 3 shows the arrangement and positive thrust direction of six thrusters. The two thrusters (1) and (6) are attached in parallel to the X-axis mainly used for Surge and Yaw motion control. The two thrusters (2) and (5) are parallel to the Z-axis for Heave and Pitch motion control. The remain two thrusters (3) and (4) are installed parallel to the Y-axis for Sway and Roll motion control.

The thrust in each mode of motion, Surge direction F^S, Sway direction F^W, and Heave direction F^H, and the moments in Roll direction N^R, Pitch direction N^P, and Yaw direction N^Y by these thrusters can be expressed by the following equation (1).

$$F_T = \begin{bmatrix} F^S \\ F^W \\ F^H \\ N^R \\ N^P \\ N^Y \end{bmatrix} = \begin{bmatrix} T_1 - T_6 \\ -T_3 - T_4 \\ -T_2 + T_5 \\ d_{3Z}T_3 - d_{4Z}T_4 \\ -d_{1Z}T_1 - d_{6Z}T_6 + d_{2X}T_2 + d_{5X}T_5 \\ d_{1Y}T_1 + d_{6Y}T_6 \end{bmatrix} \quad (1)$$

The main symbols used in this paper are shown below.

F^S, F^W, F^H : Total translational force due to thrusters in the Surge, Sway, and Heave directions [N]

N^R, N^P, N^Y : Total moment due to thrusters in Roll, Pitch, and Yaw directions [Nm]

$T_1, T_2, T_3, T_4, T_5, T_6$: Thrust of each thruster [N]

d_{1Z}, d_{6Z} : Distance from the center of gravity to thrusters 1 and 2 in the Z-axis direction [m]

d_{2X}, d_{5X} : Distance from the center of gravity to thrusters 2 and 3 in the X-axis direction [m]

d_{3Z}, d_{5Z} : Distance from the center of gravity to the thrusters 3 and 4 in the Z-axis direction [m]

d_{1Y}, d_{6Y} : Distance from the center of gravity to thrusters 2 and 5 in the Y-axis direction [m]s

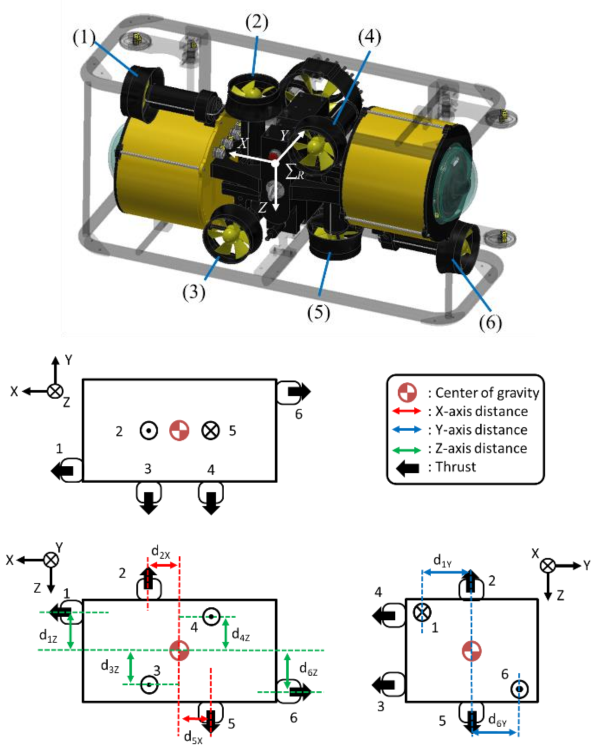


Fig. 3 Arrangement of thrusters with corresponding thrust directions

4. Cleaning Experiments

The heave, roll and pitch data in experimental results are shown in Figs. 4 and 5. During the cleaning work, the operator gives the target depth and target thrust to the robot, and the robot operates by PD control. When the robot cleans the hull of the ship, Heave and Pitch are the control targets.

In this experiment, a thrust force of 20 N is applied in the Surge direction of the robot, and the target depth is changed to 0.6 [m] and 0.9 [m]. From the results in Fig. 4, an overshoot of 0.5 [m] at the maximum appears for each target depth, and the oscillational motion happened around the target value. This result is due to insufficient design of the robot control system. In addition, it is considered that such an operation was performed because the force generated by the brush unit during cleaning and the motion characteristics in the Pitch direction were not taken into consideration.

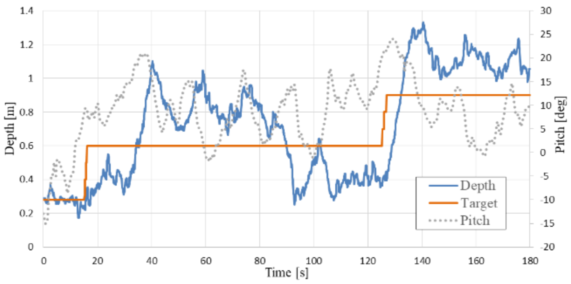


Fig. 4 Results of the experiment (Depth, Pitch)

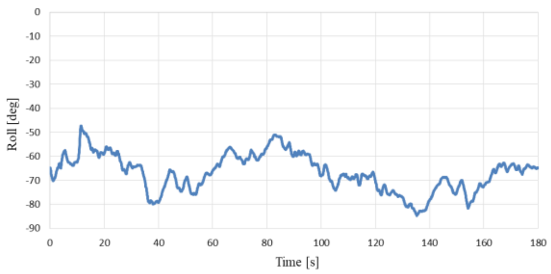


Fig. 5 Results of the experiment (Roll)

Figure 5 shows the attitude change in the direction of the robot Roll. It can be seen that the Roll angle of the robot changes according to the shape of the target hull.

5. Conclusions

In this paper, we have described the new ship hull cleaning robot and the ship hull cleaning experiment using the ship. Experiments have confirmed that the robot can clean and move along the shape of the target hull while changing the depth with respect to the target.

Therefore, it is necessary to design a more stable control system by considering the interaction between thrusters and coupling effects of motion control modes.

6. References

1. K. Ishii, A.A.F.Nassiraei, T. Takemitsu, T. Tasaka, “Control of Underwater Robot for Ship Hull Cleaning”, jsmemovic pp. 501-505, 2011.
2. A.A.F.Nassiraei, T. Sonoda, K. Ishii, “Development of Ship Hull Cleaning Underwater Robot”, 2012 Fifth International Conference on Emerging Trends in Engineering and Technology, pp. 157-162, 2012.
3. T. Takemitsu, T. Tasaka, A.A.F.Nassiraei, K. Ishii, “Development of an Underwater System to Improve Ship Fuel Efficiency -Second Report: Attitude Control of Ship−Wall Cleaning Robot-”, Proceedings of the 2011 JSME Conference on Robotics and Mechatronics. Okayama , Japan, 1A2-M09, 2011.

Acknowledgement

This research is supported by Grant-in-Aid for Scientific Research(B) 18H01643.

Authors Introduction

Mr. Hyoga Yamamoto

He received his Master's degree from the Department of Brain Science and Engineering, Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, in 2013. He is an Engineer at ROBO SCIENCE Co.,Ltd., Japan.

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. 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. 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

Development of a USV Testbed and Its System Check Experiments at Sea

Keisuke Watanabe*
Department of Navigation and Ocean Engineering, Tokai University, 3-20-1 Orido Shimzu-ku Shizuoka, 424-8610, Japan†

Masatoshi Shimpo
Department of Navigation and Ocean Engineering, Tokai University, 3-20-1 Orido Shimzu-ku Shizuoka, 424-8610, Japan‡
E-mail: keisuke@tsc.u-tokai.ac.jp, shimpo-masatoshi@tsc.u-tokai.ac.jp
www.u-tokai.ac.jp

Abstract

The catamaran type USV is useful as an autonomous platform for various activities in the field of marine engineering, such as collection of marine waste in ports, ocean observation platforms, automated vessels for offshore wind farm maintenance, or installation of a subsea drilling machine and so on. The authors have been developing an experimental testbed of USV platform to conduct basic studies on control algorithms by simulation as well as hardware systems to conduct sea experiments. In this paper, we report on some simulation results of LQR based control algorithm and introduce the testbed system with its system check experiment results near our university pier.

Keywords: Offshore wind farm maintenance, USV, Sea experiment, Subsea mining.

1. Introduction

Catamaran-type USVs are useful in various marine projects that have been increasing in recent years, such as the collection of marine plastic waste, wide area observation networks of ocean environment, offshore wind power platform maintenance, or installation of a subsea machinery and so on. If the size of the ship is increased, it is possible to apply it to the installation of the deep seafloor drilling equipment in the development of subsea mining. These USVs are required to have the ability to follow a given trajectory and carry out missions in unknown disturbances such as wind and tidal currents while avoiding other ships and obstacles at sea.

Autonomous navigation algorithms such as following a given trajectory, reaching a target point, avoiding obstacles, etc. are essential functions of a USV, although there are differences in actual individual operations depending on a given mission.

So far, the authors have constructed a water tank experimental system consisting of a catamaran and its position detection system, and verified the obstacle avoidance algorithm[1]. In addition, we have conducted a basic study of the trajectory tracking control algorithm by computer simulation. Although each of these studies was useful as an initial study method, the following problems were recognized as the limitations of the study in tank experiments and computer simulations.

Tank experiment:

(1) It is difficult to introduce unknown disturbances such as waves and wind.

(2) The navigable range of USV is limited.

Computer simulation:

(3) It is not easy to simulate communication problems between a USV and a land base.

In order to clear these issues, it is necessary to repeat the experiment in the actual sea area. For this purpose, the authors are constructing an actual sea area experiment system assuming that the experiment will be conducted in a relatively quiet harbor as the first step. In this paper, we will introduce the outline of the experimental system design and the simulation results on the effectiveness of the optimal control algorithm for trajectory tracking under disturbances.

2. Testbed System Configuration

Figure 1 shows the configuration of the USV testbed. The length is around 1.3m and the width is around 0.8m. Its draft varies depending on its weight. Figure 2 shows the control system architecture placed inside the watertight compartment. We implemented two board computers. The one is for motion control and the other is for image processing. They communicate through a LAN cable. Usually, a ship testbed has one thruster and one rudder and the control problem becomes under-actuated[2]. However, considering some simulation results, we implemented three thrusters to precisely control 3 DOF of (x,y,yaw). The global position of the USV is measured by a GNSS, whose position detection accuracy has been greatly improved recently, as within 0.1m radius without RTK. Yaw angle can be estimated by the GNSS once the USV started cruising, however, when it stops the estimation becomes degrading. So we implemented 9DOF IMU module mainly to detect its yaw angle when the USV isn't cruising. We use wireless LAN of the board computer to monitor its inner status if the testbed is near our land computer base. But the distance we can communicate is limited. As the testbed is supposed to cruise around 1~2km off the shore, we implemented a wireless communication device which can transmit RS232 signals for long distance communication. Through this, we can not only monitor its inner control situation but also control it manually from land PC in case the automatic control system doesn't work well or when we want it to obey our order during experiments.

A forward camera is mounted on the deck to detect obstacles on its trajectory during navigation. The camera is also supposed to be used for AI image processing experiments for inspection of port structures.

Fig.1 Configuration of USV Testbed.

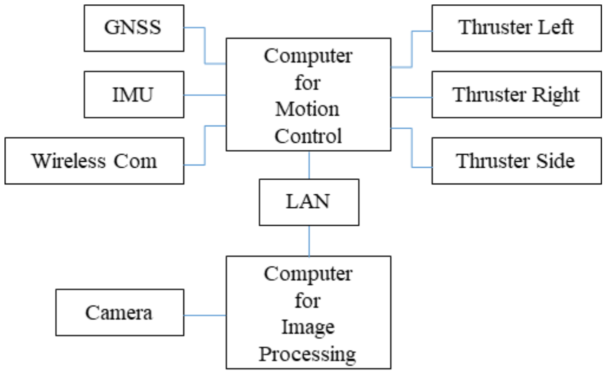


Fig.2 Control System Architecture

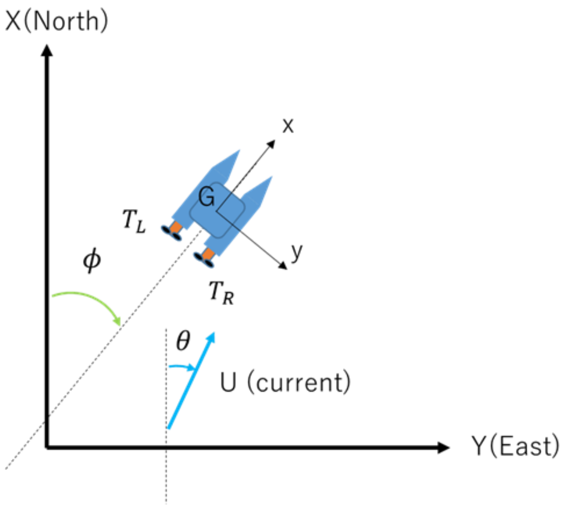


Fig.3 Coordinate System for Controller Modeling

3. Control Algorithm

Though the real dynamics of the USV is complicated due to its coupled motions of 6 DOF and environmental forces as wave, wind and current, we need to abstract the essential components of the dynamics. The main control purpose in this paper is the trajectory tracking or position keeping on the surface, so only 3 DOF of surge, sway and yaw are hired for the dynamics modeling of the USV testbed[3]. Figure 3 shows the coordinate system for the modeling of equation of motion. The USV position determined by GNSS is based on the fixed global coordinate whose origin is set from the start point. The X axis is parallel to the longitude, and the Y axis is parallel to the latitude. On the other hand, its thruster forces work based on the local coordinate whose origin is USV's COG. The nonlinear equations of motion are as follows.

Surge:

$$(m+m_a)(\dot{u}-v\omega)+\rho C_d A_x|u-u_C|(u-u_C) = T_R+T_L \tag{1}$$

Sway:

$$(m+m_a)(\dot{v}+u\omega)+\rho C_d A_y|v-v_C|(v-v_C) = T_S \tag{2}$$

Yaw:

$$(I+I_a)\dot{\omega}+\frac{1}{2}\rho C_d\frac{A_y}{2}\left|\omega\frac{L}{4}\right|\left(\omega\frac{L}{4}\right)\cdot L = -l\cdot T_R+l\cdot T_L \tag{3}$$

Here, m is mass, m_a is added mass, u is surge velocity, v is sway velocity, ω is angular velocity of yaw, ρ is water density, C_d is drag coefficient, A_x, A_y are typical area of the hull in each direction, I is the moment of inertia of the hull, I_a is the added inertia by the surrounding fluid, u_C is x component of current velocity, v_C is y component of current velocity, L is the length of the hull, T_R, T_L, T_S are thruster forces. The suffixes are as R means right, L means left, S means side. l means the length from the centerline of the hull to the thruster attached point. The drag force by the fluid is based on the modified Morrison's equation which includes the current effects.

The relation between the USV fixed coordinate and the global coordinate is as follows.

$$\phi=\phi_0+\int_0^t\omega\,dt \tag{4}$$

$$X=X_0+\int_0^t(u\cos\phi-v\sin\phi)\,dt \tag{5}$$

$$Y=Y_0+\int_0^t(u\sin\phi+v\cos\phi)\,dt \tag{6}$$

The following nominal model can be obtained by linearly approximating the nonlinear term of the fluid force using typical cruising velocity.

$$\begin{bmatrix} m+m_a & 0 & 0 \\ 0 & m+m_a & 0 \\ 0 & 0 & I+I_a \end{bmatrix}\begin{bmatrix}\dot{u}\\ \dot{v}\\ \dot{\omega}\end{bmatrix}+\begin{bmatrix}\rho C_d A_x\bar{u} & 0 & 0 \\ 0 & \rho C_d A_y\bar{v} & 0 \\ 0 & 0 & \frac{1}{64}\rho C_{dI} A_y L^3\bar{\omega}\end{bmatrix}\begin{bmatrix}u\\ v\\ \omega\end{bmatrix}= \begin{bmatrix}1 & 1 & 0\\ -l & l & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}T_R\\ T_L\\ T_S\end{bmatrix} \tag{7}$$

$$\begin{bmatrix}u\\ v\\ \omega\end{bmatrix}=\begin{bmatrix}\cos\phi & \sin\phi & 0\\ -\sin\phi & \cos\phi & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}\dot{X}\\ \dot{Y}\\ \omega\end{bmatrix} \tag{8}$$

Equation (7) is written as follows.

$$\boldsymbol{M}\ddot{\boldsymbol{\xi}}+\boldsymbol{C}\dot{\boldsymbol{\xi}}=\boldsymbol{Lf} \tag{9}$$

$$\dot{\boldsymbol{\xi}}=\boldsymbol{R}\dot{\boldsymbol{x}} \tag{10}$$

Here, $\dot{\boldsymbol{\xi}}$ is the velocity vector in USV fixed coordinate, $\dot{\boldsymbol{x}}$ is the velocity vector in the global coordinate. $\boldsymbol{R}$ is the transformation matrix depending on the yaw angle.

Since the control target is set at a point on the global coordinates, the control system is designed with the global coordinate system.

$$\boldsymbol{MR}\ddot{\boldsymbol{x}}+\boldsymbol{M}\dot{\boldsymbol{R}}\dot{\boldsymbol{x}}+\boldsymbol{CR}\dot{\boldsymbol{x}}=\boldsymbol{Lf} \tag{11}$$

Let $\boldsymbol{e}$ be the error vector between the current position and the target point.

$$\boldsymbol{e}=\boldsymbol{x}-\boldsymbol{x_t} \tag{12}$$

To implement 1-type servo system, introducing variable z as follow,

$$\dot{\boldsymbol{z}}=\boldsymbol{e} \tag{13}$$

Then, let the control state vector as $\dot{\boldsymbol{X}}=[\boldsymbol{z}\quad \boldsymbol{e}\quad \dot{\boldsymbol{e}}]^{\boldsymbol{T}}$ and the state equation as follows.

$$\dot{\boldsymbol{X}}=\boldsymbol{AX}+\boldsymbol{Bf} \tag{14}$$

Setting the optimization function J as,

$$J=\int_0^\infty(\boldsymbol{X^TQX}+\boldsymbol{f^TRf})\,dt \tag{15}$$

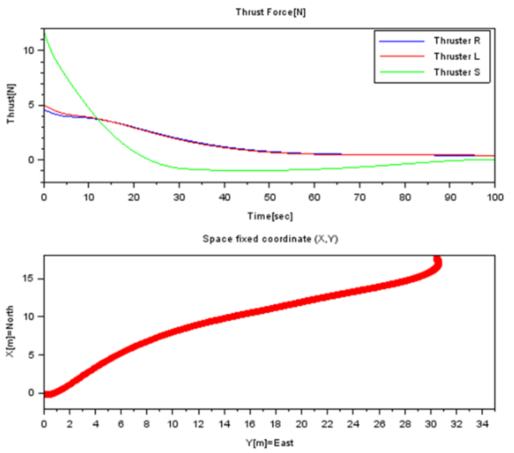


Fig.4 Simulation Result from (0,0) to (30,20) current direction was 180 deg.

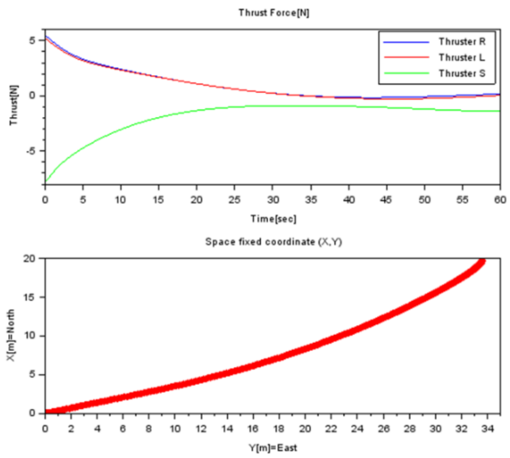


Fig.5 Simulation Result from (0,0) to (30,20) current direction was 90 deg.

The control force vector is obtained as follows by solving the LQR gain in each control step.

$$\boldsymbol{f} = -\boldsymbol{GX} \tag{16}$$

4. Simulation Results

To verify the control algorithm above, we made a simulation program and carried out simulations in several different conditions Figure 4 and 5 show the trajectory and control forces in the case of the target way point was set to (30m, 20m). The current speed was 0.3m/s and its direction was 180 deg (against to the direction of the USV) in Fig.4, and 90 deg (from the side of the USV) in Fig.5. Nonlinear drag force is considered in this simulation. The USV successfully reached the target point and the control algorithm worked as expected.

5. Conclusion and Future Work

The optimum control algorithm of the USV experimental system was designed, and some simulations were carried out by changing the direction and strength of the current. As a result, it was confirmed that a control system to reach the target way point can be constructed by using the optimum control gain based on the linearized nominal model. The verification in the sea experiment is our next step of this research.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 18K04585.

References

1. K. Harada, K. Watanabe, M. Shimpo, “Experiment System Development for Study on Control Algorithm of Unmanned Surface Vehicle Group”, The Journal of Japan Institute of Navigation, vol.140, pp.77-87, 2019(in Japanese)
2. Z. Dong, L. Wan, Y. Li, T. Liu, G. Zhang, “Trajectory Tracking Control of Underactuated USV Based on Modified Backstepping Approach”, Int. J. Nav. Archit. Ocean Eng., pp.817-832, 2015
3. T.I.Fossen,”Marine Control Systems”, Marine Cybernetics AS., ISBN82-92356-00-2, 2002

Authors Introduction

Dr. Keisuke Watanabe

He is a Professor of School of Marine Science and Technology, Tokai University in Japan. He received his B.E., M.E. and ph.D. degrees from the Dept. of Naval Architecture and Ocean Engineering, The University of Tokyo, Japan in 1991,1993 and 1996. His research interest is autonomous marine vehicles application.

Dr. Masatoshi Shimpo

He is a Professor of School of Marine Science and Technology, Tokai University in Japan. He received his B.E, M.E. from Kobe University of Mercantile Marine, in 1979,1987 and his ph.D. from Tokyo University of Marine Science and Technology, in 2004. His research interest is image processing application.

Development of a Seabed Walking Platform for Ore Sample Drilling in Deep Sea Mining

Keisuke Watanabe
Department of Navigation and Ocean Engineering, Tokai University, 3-20-1 Orido Shimizu-ku
Shizuoka, 424-8610, Japan

E-mail: keisuke@tsc.u-tokai.ac.jp

Abstract

In deep-sea mineral resource development, exploratory drilling is indispensable for estimating the amount of resources. In order to reduce the cost of exploration, a system for drilling on the seabed is needed instead of a support vessel at sea. This exploration platform must be able to move along with the undulations of the seafloor and have a structure that supports the reaction force of the drilling. The authors have been studying an eight-legged drilling platform. This paper introduces the system configuration of an eight-legged walking robot.

Keywords: Deep sea mining, Exploration phase, In-situ drilling, Sea bed walking robot

1. Introduction

Acquiring mineral resources such as rare metals from the seabed has long been debated. However, a commercial seafloor mineral resource development project has not yet been realized. The reason for this is that although there is a technical problem that the mining system in the deep sea has not been established, there is also a big problem that the method of deposit exploration necessary for judging the economic feasibility of the project has not been established.

Currently, there is a track record of drilling to a depth of around 3000 m, and in the development of technologically mature offshore oil fields, the ratio of exploration costs to the entire project is becoming relatively small. For example, according to a report[1] by Norwegian Petroleum, "In 2020, the overall costs were around NOK 245 billion. Investments made up about 60 per cent of this, operating costs 25 per cent, and exploration costs about 10 per cent.". Nevertheless, a very large amount of money is spent, and it is expected that exploration costs will need to be kept lower in the development of seafloor mineral resources than in the development of offshore oil fields.

In offshore oil field development, the size of subsea reservoirs can be estimated by seismological exploration. Then, an exploration well will be drilled for the promising reservoir. There are two types of exploration wells. One is called wildcat wells, and the other is called appraisal wells. Wildcat wells are drilled to find out if hydrocarbons are really there beneath the seabed of the target area. Once the discovery has been approved, appraisal wells are drilled to obtain more data about the size and extent of the reservoir to estimate the feasibility.

On the other hand, in the development of seafloor mineral resources, it is impossible to see veins by seismic exploration. Therefore, it is necessary to carry out a large number of wildcat-like exploration to estimate the size and range of veins by obtaining actual ore samples. For exploration drilling, it is possible to use MODU (Mobile Offshore Drilling Unit), which is commonly used in offshore oil/gas field development. However, this method is not a good idea not only because of its high operation cost, but also because of the transportation of ore samples from the seabed to the ocean. This is because, in the case of crude oil, the underground pressure causes the fluid to blow up inside the riser, but the ore sample requires energy to be sucked up inside the riser pipe. Moreover, in this case, the ore sample needs to be made into fine particles and mixed with seawater to have fluidity. Therefore, in addition to the drilling device, it is necessary to install a crusher that finely crushes the drilled ore and a mixer that mixes the ore with seawater to enhance fluidity on the seabed or at the riser entrance.

Based on the above background, the author is studying a method of deploying a large number of in-situ type subsea drilling equipment and conducting exploration drilling at low cost without using MODUs. In this paper, the 8-legged walking robot currently under consideration is introduced[3].

2. The Concept of Eight-legged Walking Robot

As a method of moving the device for in-situ drilling on the seabed, a method using a thruster such as a navigation type underwater robot can be considered. However, in the drilling robot, it is necessary to increase its own weight enough to support the reaction force generated when drilling the seabed. When moving by a thruster, it is necessary to generate a buoyancy that cancels this own weight, and a buoyancy variable mechanism will be required. It is not easy to realize a variable buoyancy mechanism at deep water. Therefore, a walking platform that can move while supporting its own weight is suitable. Figure 1 shows the prototype design of the eight-legged robot. The robot consists of two decks that can slide in orthogonal XY directions. These are called Upper Deck (UD) and Lower Deck (LD), respectively. Each deck has four legs, each of which can slide up and down. As shown in Fig. 1, each leg is L1 to L4 and U1 to U4. Walking movement is realized as follows. While the UD legs U1 to U4 land and support their own weight, the LD legs L1 to L4 are lifted and the LD is slid. Also, by landing these legs, it supports the reaction force of drilling the seabed of the drill unit. The crawler-type drilling machine being developed by JOGMEC[2] is difficult to climb up and down the slope of the seabed. On the other hand, since this 8-legged robot can change the length of each leg, it can move even in a complicated seabed topography. That is, it is possible to move along the unevenness of the seafloor topography and to go up and down the slope. If the seabed is soft ground, it is necessary to devise a mud mat so that each leg is not trapped in the ground.

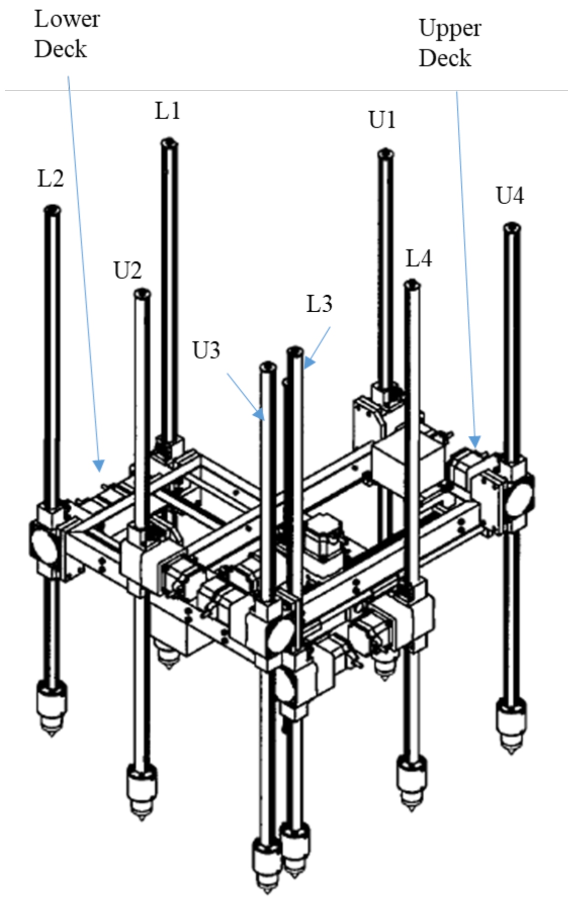


Fig. 1. Prototype design of the eight-legged robot.

The drilling unit is attached to the Upper Deck. The details of the drilling unit need to be further examined in the future, but the important functions required of this robot, including the drill unit, are as follows.

(1) Being able to collect the drilled ores so that they are not scattered in the sea.

(2) Automatically connect the drill pipes vertically to dig deeper.

(3) Equipped with drilling and moving power

(4) Being able to accurately detect the position on the seabed

In order to realize these functions, it is necessary to build experimental system and accumulate knowledge through experiments.

3. Experimental System for Walking Procedure

In order to study the walking algorithm of the 8-leg drilling platform, an experimental device as shown in Fig.

2 was manufactured. The motor was stored in a watertight container and an underwater walking experiment was conducted. An attitude (roll, pitch) sensor is attached to each deck, and the length of each leg is adjusted so that the deck is horizontal. In addition, a touch sensor is attached to the tip of each leg to know that it has touched the seabed.

As shown in Fig. 3, a slope was provided in the water tank, and obstacles were installed at the same time to realize unevenness. If the touch sensor detects the bottom of the sea, the leg descent is automatically stopped. The control algorithm is as follows. First, move from the start point toward the target point by a given distance in the X direction. The robot then moves a given Y-direction distance. In this experiment, we could not obtain a Yaw sensor that satisfies sufficient accuracy, so we did not consider the rotational motion around the Z axis. It was confirmed that it is possible to go up and down the slope very easily by using the touch sensor and the attitude sensor of the deck. Figure 4 shows the control menus of the developed interface. Each leg can be handled manually as well as automatically controlled to follow the given target point in the tank.

Fig. 2. Schematic View of Experimental Setup.

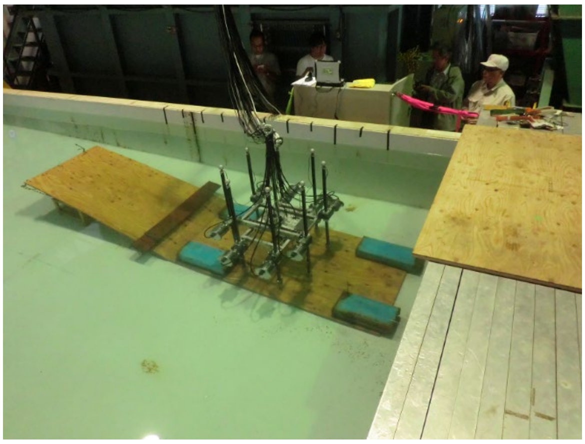


Fig. 3. Basin Experimental Setup with a Slope

4. Some Experimental Results

Figure 5 and 6 show experimental results of walking experiment in the water tank. Figure 5 shows the zig zag walking result in the horizontal plane. As shown in the graph, the robot is controlled to move X-direction from -2000 mm to -300mm, while Y-direction from 500 mm to 750 mm. The step in the graph shows the slid distance of the deck. Figure 6 shows the slope ascending and descending experiment result. As shown in the graph, it moved along the plate placed on the tank floor in Fig.3. The grey points show the Z directional distance, which means the height. It goes up from around -1400 mm to -1000 mm and goes down from a-1000 mm to -1400 mm, which means its slope walking succeeded. In this experiment, as shown in Fig. 3, four blue steps and a step by a bar in the middle of the slope were prepared. Despite these steps, we were able to confirm that our robot could walk with a very simple algorithm.

As can be seen from Fig. 2, there are many cables connecting the motors and they are relatively heavy, and in a navigation type underwater robot, the cables are at a level that hinders movement. However, since this robot is a mechanism in which the legs land on the bottom and support its own weight, no adverse effect of the cable on the walking was observed. I think this is a big advantage of this concept.

5. Discussions

As a result of this tank experiment, some future issues were found. It goes without saying that the length of the foot is closely related to the ability to climb a slope. The difference in height between the front and rear legs vs the length of the deck determine the angle of inclination of the slope that can be climbed.

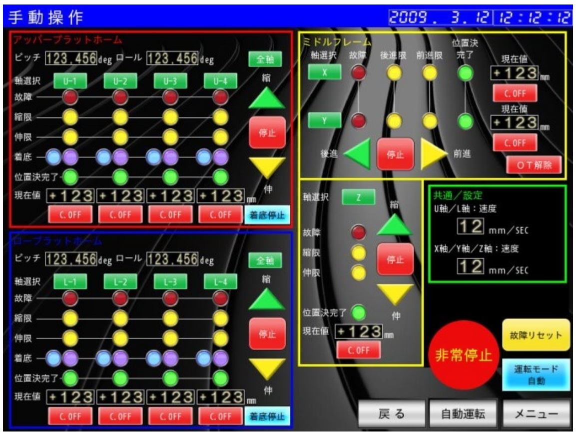


Fig. 4. Control Interface on the PC

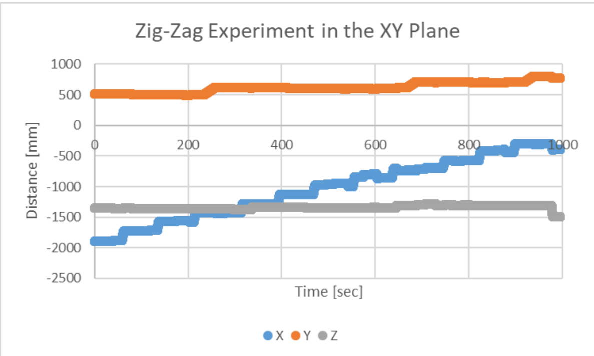


Fig. 5. Zig Zag Walking Experiment Result

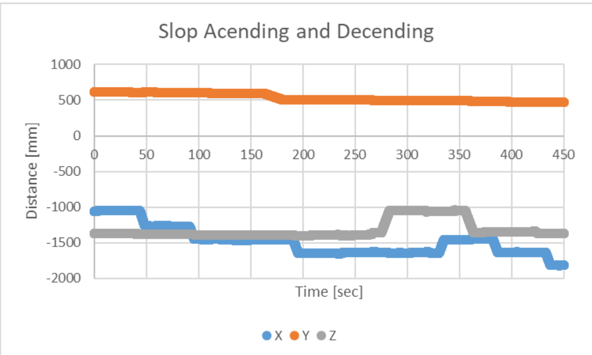


Fig. 6. Slope Ascending and Descending Result

Therefore, it is important for designing the platform size to survey the topography of the seafloor of the area that is planned to be explored in advance.

In addition, the platform cannot rotate around the z-axis. In order to change the direction of travel on the seafloor, it is necessary to consider a mechanism that allows rotation. Furthermore, it is necessary to consider how to install the robot in the deep sea.

6. Conclusion

In this paper, we introduced the in-situ type subsea drilling walking robot that we are currently studying, and introduced the underwater walking experiment in the tank. The robot has two sliding decks and eight legs, and while the four legs support their own weight, the other four legs slide and move. We constructed a water tank experiment system, conducted an experiment to walk zigzag in a plane, and an experiment to go up and down a slope, and confirmed the basic performance. The experiment was successful, and it was confirmed that a simple algorithm can make the platform move to the target point. In addition, some new research challenges have been discovered that should be investigated further. For example, it is difficult to increase the walking speed because it takes time to raise and lower the legs, and it is also difficult to rotate around the vertical axis. However, through experience in the experiments, it was confirmed that this concept is suitable for in-situ drilling because it can support the reaction force of drilling as well as various devices can be mounted on the deck.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 18K04585.

References

1. https://www.norskpetroleum.no/en/economy/investments-operating-costs/.
2. N. Okamoto, S. Shiokawa, S. Kawano, H. Sakurai, N. Yamaji and M. Kurihara, "Current Status of Japan's Activities for Deep-Sea Commercial Mining Campaign," 2018 OCEANS - MTS/IEEE Kobe Techno-Oceans (OTO), 2018, pp. 1-7.
3. Y. Watanabe, K. Watanabe, H. Suzuki, T. Ooshima, Y. Tsukamoto, "Research and Development of a Self-Walking Vertical Mining System Using DTH Drilling Unit and the Scale Model Test", OMAE2019-95394, 2019

Authors Introduction

Dr. Keisuke Watanabe

He is a Professor of School of Marine Science and Technology, Tokai University in Japan. He received his B.E., M.E. and ph.D. degrees from the Dept. of Naval Architecture and Ocean Engineering, The University of Tokyo, Japan in 1991,1993 and 1996. His research interest is autonomous marine vehicles application.

Development of Remotely Operated Vehicle for Small-size Jellyfish Extermination and its Evaluation of Extermination Motion Control

Hiroyuki Yokota
Dept. of Intelligent Mechanical Engineering, Hiroshima Institute of Technology,
2-1-1 Miyake, Saeki, Hiroshima,731-5193, Japan
E-mail: ad18107@cc.it-hiroshima.ac.jp

Shinsuke Yasukawa
Dept. of Human Intelligence Systems, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka, 808-0196, Japan
E-mail: s-yasukawa@brain.kyutech.ac.jp

Jonghyun Ahn
Dept. of Intelligent Mechanical Engineering, Hiroshima Institute of Technology,
2-1-1 Miyake, Saeki, Hiroshima,731-5193, Japan,
E-mail: j.ahn.h2@it-hiroshima.ac.jp

Abstract

In recent years, increase in the number of jellyfish has caused damage in the fishery and tourism industries. Therefore, the extermination work of jellyfish is being carried out by human hands. However, conventional methods for extermination are required a lot of time and manpower. In this paper, we propose a method for extermination work of jellyfish using underwater robot. Also, we introduce developed ROV type underwater robot, which is called J.E.N.O.S.(Jellyfish Extermination Nifty-robot for Ocean Sustentation), and its extermination motion control. The ROV is developed in consideration of the attitude control during the extermination operation. Because, the attitude, such as surge and pitch angle, of ROV becomes unstable when performing jellyfish extermination. Therefore, we equipped 8 thrusters to improve attitude stability during the jellyfish extermination. As a result, surge acceleration is reduced to about 30.0%, and pitch angle velocity is reduced to about 25.8%.

Keywords: robot design, ocean engineering, remotely operated vehicle, jellyfish extermination

1. Introduction

In recent years, the massive increase in the number of jellyfish is damaging to ocean environments such as decrease in biodiversity. Also, the number of jellyfish is increasing across Asia, damaging the fisheries, power generation and tourism industries [1][2]. It is supposed that these causes are overfishing of fishery resources, changes in coastal topography due to global warming, and eutrophication [3]. In Japan, the aurelia, which is small and poisonous jellyfish, is damaging to human activities such as fisheries and tourism industries. furthermore, they gather at the water intakes of power plants, and clog the intakes [4][5]. Therefore, jellyfish extermination work is carried out to reduce such damage.

Conventionally, jellyfish extermination work is used fishing nets. The workers catch jellyfish using fishing nets, and then caught jellyfish is cut up by the workers. However, this method is required a lot of cost such as financial cost, time, and manpower. Furthermore, the number of fishery workers in Japan has been declining in recent years due to the decreasing birthrate and aging population [6]. Underwater robot can be mentioned as a solution to this situation. The advantage of underwater robot is that, unlike humans, they do not have to bear any physical burden and can work for a long time. In this paper, we design and develop a ROV (Remotely Operated Vehicle) for jellyfish extermination work, which is called J.E.N.O.S.(Jellyfish Extermination Nifty-robot for Ocean Sustentation) and evaluate ROV motion to jellyfish extermination.

2. Development of Jellyfish Extermination ROV

2.1. *Operation of developed to ROV*

The jellyfish extermination ROV is operated on the small size fishing boat. The detail of proposed jellyfish extermination ROV operation is illustrated in Fig.1.

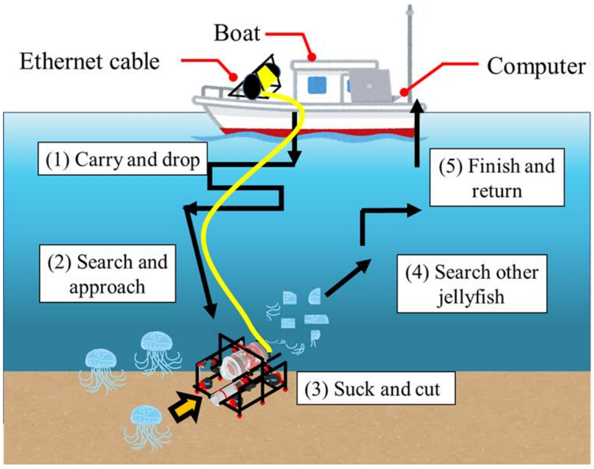


Fig. 1 Overview of the proposed jellyfish extermination ROV operation method

In the proposed operation, the ROV has 5 steps to exterminate jellyfish.

(1) The ROV and operator are carried using a boat to the field where is number of jellyfish is increasing. Then, the ROV is dropped to the sea area.
(2) The ROV is controlled by operator to search jellyfish. If a jellyfish is found, the ROV approaches to the jellyfish.
(3) A jellyfish extermination device, which is mounted to ROV, sucks the jellyfish with water. Then the jellyfish is cut by the blade.
(4) The ROV is controlled to search other jellyfish in the same sea area.
(5) After the work in the sea area is finished, the ROV is returned on the boat. Then, operator try to same works in the other sea area.

We develop a ROV which can operate these steps.

2.2. *Design concept and specifications of ROV*

The ROV is necessary several condition to operate the proposed jellyfish extermination work on the boat. The design concept of ROV, which is considered to develop, is described below.

(i) Small size and light weight
(ii) High maintainability
(iii) Wide viewing angle and high recognition rate
(iv) High positional stability

We proposed a ROV operation method, which can be done with a few workers compared to conventional method, such Fig.1. Therefore, the design concept of ROV, which satisfies (1) ~ (5), is necessary. For the design concept (i), size and weight must be designed easy enough to be drop and return on the boat by 2 people. Because the minimum number of people required to work on a boat is 2 for the safety.

The purpose of the design concept (ii) is to ensure that if a problem occurs during the operation, such as a poor connection or dead battery related to the internal system of ROV, it can be recovered and repaired quickly.

Underwater environment is dark and makes it difficult to recognize objects. Therefore, like the design concept (iii), the ROV is necessary to light source and wide viewing angle.

In the Fig.1, we proposed an operation method, which is sucked the jellyfish and cutting them. However, in this way, the posture of the ROV tends to be unstable because of the suction device applies the force to ROV. Therefore, the ROV is necessary to high positional stability for jellyfish extermination in the design concept (iv).

The designed ROV, which is considered the design concept (i) ~ (iv), is shown in Fig.2 and the ROV specifications is shown in Table 1.

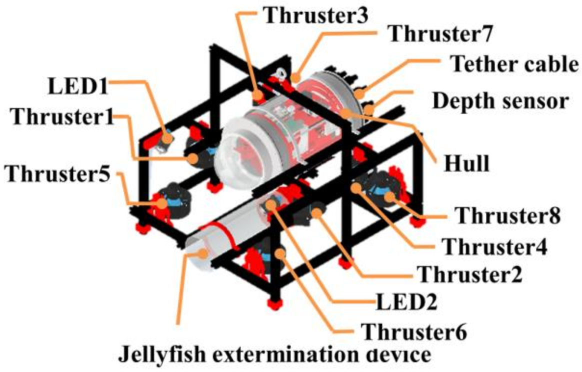


Fig. 2 Overhead view of designed the ROV for jellyfish extermination

Table 1 Specifications of developed to ROV

Dimensions	0.65 x 0.52 x 0.47 [m] (L x W x H)
Dry weight	17[kg]
Maximum depth	10[m]
Actuator	Thruster x 8, Micro servo motor x 2
Battery	Lithium-ion battery x 1
Sensors	USB camera x 1, IMU x 1, Depth sensor x 1
Umbilical Cable	100[m] length, 7.6 [mm] diameter
Equipment	LED x 2, Jellyfish extermination device x 1 Voltmeter x 1
Angle of view	Pan110[deg], Tilt120[deg]

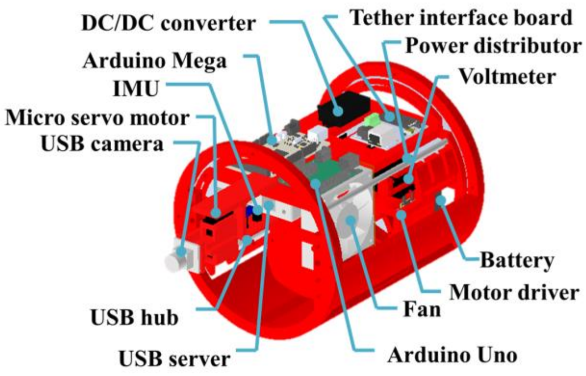


Fig. 3 Inside design of the hull of developed ROV

As shown in Fig.2, the ROV includes a hull, which is composed camera dome and acrylic cylinder, thrusters, LEDs and jellyfish extermination device. Each device is equipped using aluminum frames to ROV.

The ROV is designed 0.65 x 0.52 x 0.47[m] and 17[kg] because of design concept (i) for the drop and return work. This size and weight are made possible the extermination operation using the ROV by 2 grown mans on the boat.

The ROV hull inside is designed as a module for design concept (ii). The inside design of hull is shown Fig.3.

As shown in Fig.3, the control units such as Arduino, battery, motor drivers, sensors are placed one hull inside.

This is a single module, designed to be easily inserted into and removed from the hull by slide mechanism.

The camera system is composed by a camera, 2 LEDs and 2 micro servo motors to satisfy design concept (iii). The camera system can move to pan and tilt angle. Also, the LEDs, which has a maximum brightness of 1,500 lumens [7], are equipped.

The ROV has low positional stability during jellyfish extermination work because the center of rotation of ROV and jellyfish extermination device are mounted at a distance as shown in Fig.2. Therefore, the force from jellyfish extermination device occurs to surge direction and pitch angle. This force makes the ROV posture unstable, and it is made difficult to jellyfish extermination work. The ROV is mounted 8 thrusters which are maximum output of 5.25 [kgf] at 16 [V] [8] to satisfy design concept (iv). We placed 4 thrusters in the surge and sway directions, and other 4 thrusters are placed in the heave direction as shown in Fig.2 for high positional stability of the ROV.

2.3. *Design of power and communication system*

Power source of the ROV is employed a lithium-ion battery which is 14.8[V], 18[Ah]. The designed power system is shown in Fig.4.

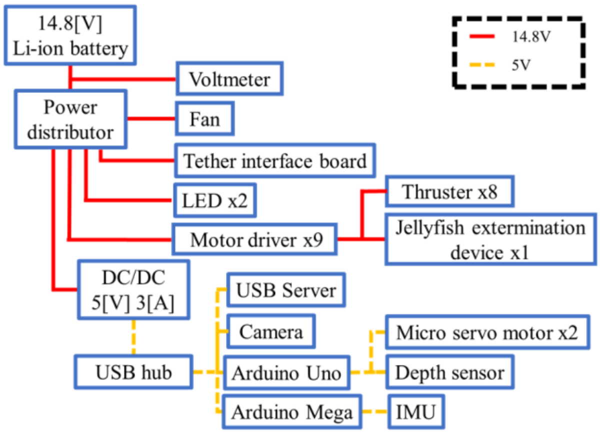


Fig. 4 Overview of power supply system

The battery is supplied DC 14.8 [V] to electronic parts as shown in Fig.4. DC 14.8[V] are employed to supply power to fan, voltmeter, LEDs, motor drivers and tether interface board. Also, DC 5[V], which are employed to sensors and processing units, are converted by DC/DC converter, and those are employed to supply power to USB server, camera, Arduino Uno and Mega.

The communication system is illustrated in Fig.5.

As shown in Fig.4, the ROV communication system is employed various communication formats. The computer on the boat is connected to ROV by ethernet cable via tether interface and then the ethernet signal is converted to USB serial by USB server. The USB hub, which connected to Arduino Mega and Uno, is connected to USB server. Arduino Mega controls thrusters, jellyfish extermination device, LEDs and get the IMU data. The PWM signal is employed for the thruster and LED control. The PWM signal, which outputted to Arduino Mega, has 1100 to 1900 [micro sec.] range. Arduino Uno controls micro servo motors for camera movement and get the depth sensor data.

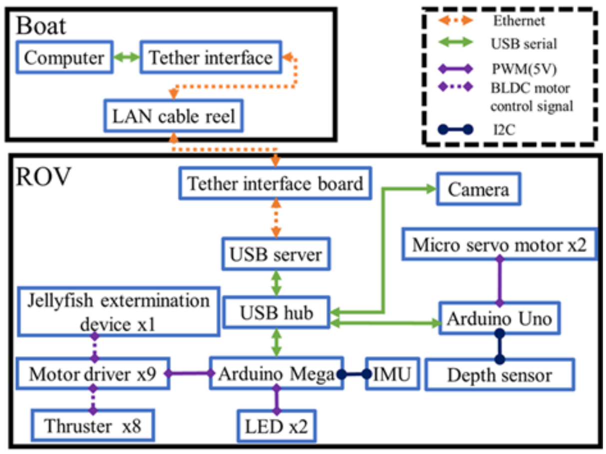


Fig. 5 Overview of communication system

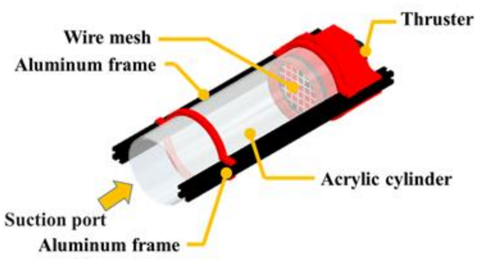


Fig. 6 Overview of jellyfish extermination device

Table 2 Specification of jellyfish extermination device

Dimensions	0.46 x 0.15 [m] (L x W)
Dry weight	2[kg]
Actuator	Thruster(T200)
Suction port diameter	100[mm]
Filter size	15 x 15[mm]
Thrust	Up to 4.07 [kgf] at 16 [V]

2.4. *Jellyfish extermination device*

We proposed jellyfish extermination method using suction device in Fig. 1. The jellyfish extermination device, which can be possible the proposed jellyfish extermination ROV operation method, is illustrated in Fig.6 and the Specifications is shown in Table2.

The jellyfish extermination device is a thruster, wire mesh, acrylic cylinder, and frames as shown in Fig.5. When the thruster rotates, the jellyfish is sucked in and cut off by the wire mesh and blade of thruster. In the proposed jellyfish extermination ROV operation method, we do not collect cut jellyfish because the cut jellyfish are used as food for marine organisms in marine ecosystem.

The wire mesh is made of stainless steel and the mesh size is 15 [mm] x 15 [mm]. It also serves to crush jellyfish and to protect hard objects other than jellyfish, such as large rocks and shells, from the thruster.

3. Evaluation experiment of developed ROV

We evaluate stability of the ROV motion and jellyfish extermination device is performance by 2 experiments.

For the motion stability, the jellyfish extermination motion of ROV is evaluated with acceleration and angular velocity. For jellyfish extermination device, the device is evaluated with jellyfish sample extermination time.

3.1. *Evaluation of extermination motion control*

The extermination motion of ROV is performed using jellyfish extermination device which mounted downside of the ROV. Therefore, when using the jellyfish extermination device, the pitch angle rotates as the robot moves forward. To solve this problem, we stabilized the posture by using other thrusters mounted on the ROV. In this experiment, the thrust of the extermination device was set to 1.44 [kgf]. The control results are shown in Fig.7, and the scene of experiment illustrated in Fig.8.

The Fig.7 (a) ~ (b) represents acceleration of surge, sway, and heave direction. And the Fig.7 (d) ~ (f) are shown angular velocity of roll, pitch, yaw angle. The dashed lines represent acceleration and angular velocity during extermination motion without control of other

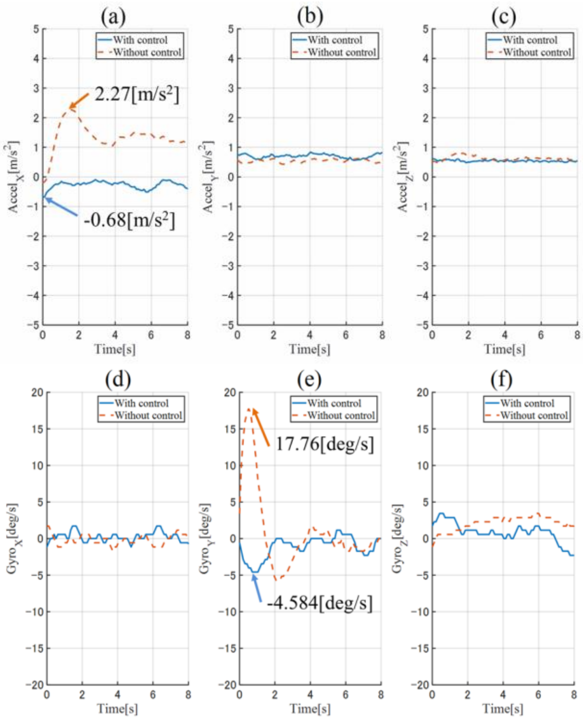


Fig. 7 Experimental results

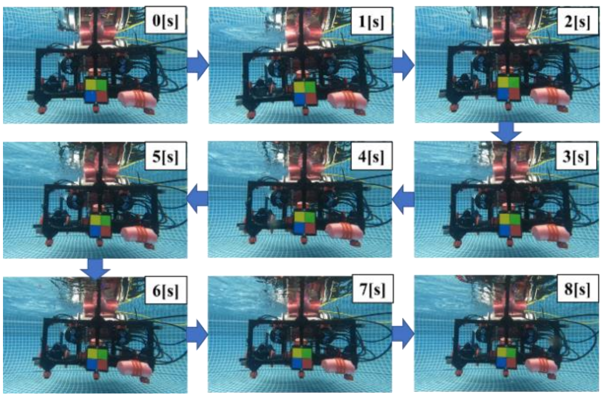


Fig. 8 Scene of the extermination motion control

thrusters, and the solid lines represent acceleration and angular velocity during extermination motion with control of other thrusters. Also, each data is measured by IMU which is mounted the ROV.

Fig.7 (a) in the dashed line, the ROV can confirm that the acceleration is applied in the surge direction during extermination motion and its maximum acceleration of the surge direction was 2.27[m/s^2]. Fig.7 (a) in the solid line represents the surge acceleration which is reduced by mounted thrusters on ROV. As a control result, the acceleration of surge direction is about 30.0% reduced. Fig.7 (b) and (c) are shown acceleration of sway and heave direction, and there was little acceleration in those direction.

Fig.7 (d) and (f) are shown angular velocity of roll, yaw angle, and there was little angular velocity in those direction. However, Fig.7 (e) in the dashed line is shown high angular velocity and its maximum angular velocity of the pitch angle was 17.76[deg/s]. Fig.7 (e) in the solid line represents the angular velocity which is reduced by mounted thrusters on ROV. As a control result, the angular velocity of pitch angle is about 25.8% reduced.

We can see in Fig. 8 that the acceleration and angular velocity are highly reduced. In Fig. 8, a marker is attached to the lower center of the ROV, and the movement of the marker was observed during the extermination motion. As a result of observation, it was confirmed that the robot hardly moved.

3.2. *Evaluation of jellyfish extermination device*

We experiment to evaluate the flow velocity of jellyfish extermination device. In this experiment, the 3 types of cylindrical jellyfish samples, which composed of water and gelatin, was employed. Each jellyfish sample was made using 200, 300, 400 [ml] of water, 5, 10, 15[g] of gelatin, 1, 2, 3[g] of red food coloring, and each size (radius x height) was about 28[mm] x 80[mm], 35[mm] x 105[mm] and 40[mm] x 130[mm].

In the experiment, the sample is dropped in front of the jellyfish extermination device and then sucked to measure the flow velocity. Using an underwater camera, the jellyfish samples were recorded at 120 frames and the flow velocity was measured. The force output by the thruster which is mounted jellyfish extermination device is 1.44 [kgf]. The scene of device evaluation is shown in Fig.9.

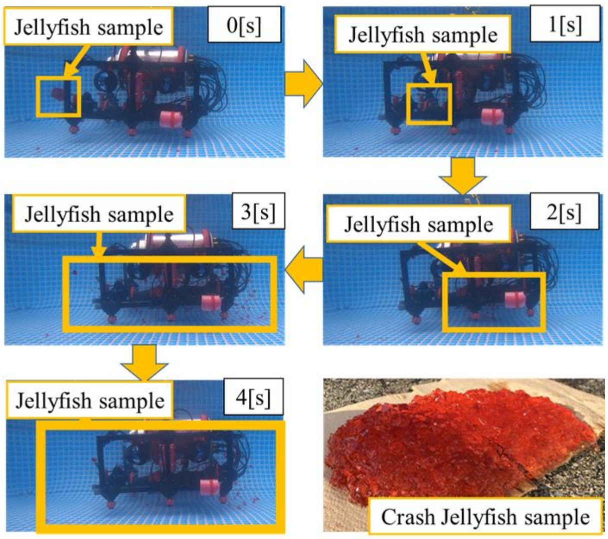


Fig. 9 Scene of jellyfish extermination device experiment

As shown in Fig.9, each jellyfish sample was chopped by jellyfish extermination device. The measured flow velocity employed the samples, which is made using 200, 300, 400 [ml] of water, was 0.60[m/s], 0.56[m/s] and 0.53[m/s]. As a result, it was confirmed that the larger the volume, the lower the flow rate.

4. Conclusion

In this paper, we design and develop a ROV for jellyfish extermination work and evaluate ROV motion to jellyfish extermination. The design concept of ROV was shown in (i) ~ (iv), and a ROV was developed that satisfies it.

In the evaluation of extermination motion control, maximum acceleration of surge direction was 2.27[m/s^2], and maximum angular velocity of pitch angle was 17.76[deg/s] by thruster which is mounted jellyfish extermination device without other thruster control. The acceleration of surge direction was reduced about 30.0[%], and angular velocity of pitch angle was reduced about 25.8[%].

There are many issues to be solved for the operation in the oceans, such as automatic recognition of jellyfish and improvement of attitude control capability such as hovering control of the ROV.

In the future, we will solve these problems and try to make practical use of the ROV for small jellyfish extermination operations.

References

1. Jennifer E. Purcell, Shin-ichi Uye, Wen-Tseng Lo, "Anthropogenic causes of jellyfish blooms and their direct consequences for humans: a review", Marine Ecology Progress Series 350, pp.153-174, November, 2007.
2. J. Gunasingh Masilamoni, K. S. Jesudoss, K. Nandakumar, K. K. Satpathy, K. V. K. Nair and J. Azariah, "Jellyfish ingress: A threat to the smooth operation of coastal power plants", Current Science Vol. 79, No. 5, pp. 567-569, 10 September 2000.
3. Rachel Chudnow, "Are Jellyfish Populations Increasing Worldwide (and Why?)", pp.1-76, April, 2008.
4. Cathy H. Lucas, Stefan Gelcich, and Shin-Ichi Uye, "Living with Jellyfish: Management and Adaptation Strategies", Jellyfish Blooms chapter 6, pp 129-150, September 2014.
5. L. Segura-Puertas, G. Avila-Soria, J. Sánchez-Rodríguez, M. E. Ramos-Aguilar, J. W. Burnett, "SOME TOXINOLOGICAL ASPECTS OF Aurelia aurita (LINNÉ) FROM THE MEXICAN CARIBBEAN", J. Venom. Anim. Toxins vol.8 no.2 Botucatu, pp.1–8, January 2002.
6. Daisuke FUJITA, "Current status and problems of isoyake in Japan", BULLETIN OF FISHERIES RESEARCH AGENCY, pp.33-42, 2010.
7. Blue Robotics, "Lumen Subsea Light for ROV/AUV".
   Blue Robotics, "T200 Thruster".

Authors Introduction

Mr. Hiroyuki Yokota

He is acquiring the B.S. Department of Intelligent Mecha-nical Engineering, in Hiroshima Institute of Technology. He is interested in underwater robotics and the preparation of a guide for writing paper manuscripts. His dream is to be a researcher on underwater robotics.

Prof. ShinsukeYasukawa

He received the Ph.D. degree in electrical, electronic, and information engineering from Osaka University, Japan, in January 2017. Currently, he is an Associate Professor of Department of Human Intelligence Systems, Kyushu Institute of Technology, Japan. His research interests include information processing in biological sensory systems and their applications in robotics.

Prof. Jonghyun Ahn

He is an Assistant Professor of Department of Intelligent Mechanical Engineering, in Hiroshima Institute of Technology, Japan. He graduated from the degree in underwater robotics Kyushu Institute of Technology, in 2017. His research interests include underwater imaging system, intelligent sensing, and underwater communication system.

A Distributed Optimal Formation Control for Multi-Agent System of UAVs

Jichao Zhao[1], Fengzhi Dai[1, 3 *], Yunzhong Song[2]

[1] College of Electronic Information and Automation, Tianjin University of Science and Technology, 300222， China;

[2] School of Electrical Engineering and Automation, Henan Polytechnic University, 454003, China;

[3] Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China

*E-mail: * daifz@tust.edu.cn*

www.tust.edu.cn

Abstract

In this paper, the distributed optimization problem for multi-agent system (MAS) formation control of unmanned aerial vehicles (UAVs) is suggested here. The situation that the internal states of a single UAV can be made fully available was aimed at, the internal optimal control law of a single UAV is designed using the optimal control theory. To cope with the obstacle that each agent in the system can only communicate with its neighbor agents, the distributed formation control law of the system is introduced based on the communication topology of the system, and further the stability of the system is analyzed in helps of graph theory. The validity of the suggested scheme is verified by both the numerical simulation and UAV platform.

Keywords: multi-agent system, UAV, optimal control, distributed formation control

1. Introduction

As the capabilities of robots continue to be improved, the application fields of robots are also expanding at the same time. However, just like humans, a single robot will show lower abilities in many cases, and in this way, the coordination of multiple agents is required to play a more significant role. The distributed coordination and cooperation ability of the agents is the foundation of the multi-agent system (MAS), that is the key to exert excess advantages to the MAS, and it is also the embodiment of the intelligence of the entire MAS.

The problem of multi-agent coordinated control includes consensus control [1], rendezvous control [2], formation control [3], and such et al. Besides, there is optimization control [4] based on the above-mentioned coordinated control. However, the optimization of a MAS has a strong dependence on the system communication network. Because the optimal control law needs to be able to obtain all the states of the collective individuals in the system in real-time, otherwise the conditions of optimal control can not be met. To deal with this problem, we designed a distributed optimal formation control, which is invariant to the structure of the system network. And this result is applied to the formation control of multiple unmanned aerial vehicles (UAVs).

The main contribution of this paper mainly includes three aspects: 1) A formation control protocol based on static consensus protocol is designed. 2) The optimal control law when the performance function of a single agent is optimal is studied. 3) Combining the formation control protocol and the optimal control law, a distributed optimization formation control protocol is designed. This protocol cannot be affected by the internal communication topology of the MAS. Even if some agents cannot communicate with each other, the system can complete formation task.

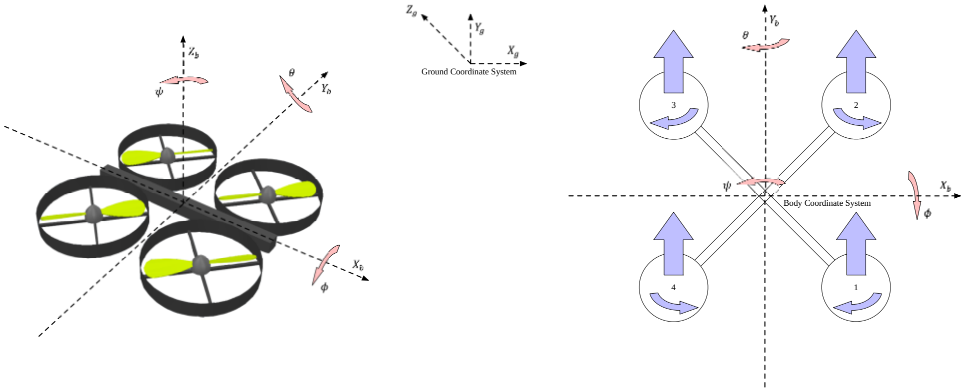


Fig.1. The illustration of UAV kinematics

2. Preliminaries

This section introduces the preliminary knowledge for studying UAV systems, including the use of graph theory to describe the communication relationships within the system, the dynamic model of a single UAV, and the state-space equation of the UAV system.

2.1. Graph theory

In order to describe the relationship of UAV systems, graph theory is used [1]. $A = [a_{ij} \in \{0,1\}]$ is the adjacency matrix of the graph G, indicating whether there is information exchange from node j to i or not. The matrix D is the out-degree matrix. The neighbors of node i are N_i. The Laplacian matrix is defined as

$$L = D - A \tag{1}$$

2.2. UAV model description

As shown in Fig.1, the UAV model can be simplified as equation (2) [5].

$$\begin{aligned} \ddot{x} &= g\theta \\ \ddot{y} &= -g\phi \\ \ddot{z} &= u_h/m - g \\ \ddot{\phi} &= u_\phi/I_x \\ \ddot{\theta} &= u_\theta/I_y \\ \ddot{\psi} &= u_\phi/I_z \end{aligned} \tag{2}$$

where x, y, z are positions in the ground coordinate system along X_g, Y_g, Z_g, respectively, ϕ, θ, ψ are roll angle, pitch angle, yaw angle of the UAV, respectively. I_x, I_y, I_z are inertia moments along X_b, Y_b, Z_b in the body coordinate system, respectively. And u_h is the lift from four propellers.

For the sake of easy computation, system (2) is transformed into the state-space format, listed in equation (3).

$$\dot{X} = AX + BU \tag{3}$$

where $X = [x \quad y \quad z \quad \dot{x} \quad \dot{y} \quad \dot{z} \quad g\theta \quad -g\phi \quad 0 \quad g\dot{\theta} \quad -g\dot{\phi} \quad 0]^{\mathrm{T}}$, $U = [u_\phi \quad u_\theta \quad 0]^{\mathrm{T}}$,

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \otimes I_3, B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \otimes I_3,$$

$\otimes$ is the Kronecker product.

When there have N UAVs, the state-space equation (3) can transform into the following form.

$$\dot{\tilde{X}} = I_N \otimes A\tilde{X} + I_N \otimes B\tilde{U} \tag{4}$$

where $\tilde{X} = [X_1^{\mathrm{T}}, X_2^{\mathrm{T}}, \dots, X_N^{\mathrm{T}}]^{\mathrm{T}}, \tilde{U} = [U_1^{\mathrm{T}}, U_2^{\mathrm{T}}, \dots, U_N^{\mathrm{T}}]^{\mathrm{T}}$, I_N represent an N-dimensional identity matrix.

The desired formation state h is defined as follows

$$h = \begin{bmatrix} h^x \\ h^y \\ h^z \end{bmatrix} \in \mathbb{R}^3 \tag{5}$$

Take the formation states (5) into the position states, three new error position states come naturally as (6).

$$\begin{bmatrix} \delta^x \\ \delta^y \\ \delta^z \end{bmatrix} = \begin{bmatrix} x - h^x \\ y - h^y \\ z - h^z \end{bmatrix} \tag{6}$$

Then the formation control problem is turned into finding a protocol $\tilde{U}$ to drive the error vector δ to zero, which indicates that

$$\lim_{t\to\infty} \|\delta_i - \delta_j\| = 0, \quad \lim_{t\to\infty} \|v_i\| = 0,$$
$$\lim_{t\to\infty} \|\Omega_i\| = 0, \quad \lim_{t\to\infty} \|\dot{\Omega}_i\| = 0, \quad i = 1,2,\cdots,N. \tag{7}$$

3. Main results

This section starts with the design of the formation control protocol. After that, the optimal control law of a single UAV is provided in detail. Finally, the optimal control law is extended to the formation control protocol in collection.

Lemma 1 [6]**.** For a $N \times N$ Laplacian matrix L, $Ne^{-Lt}, t > 0$ is a random matrix with positive diagonal elements. If L has a unique zero eigenvalue, $\mathrm{Rank}(N) = N - 1$, then its left eigenvalue has $v = [v_1 \quad v_2 \quad \cdots \quad v_N]^{\mathrm{T}} \geq 0$ and $1_N^{\mathrm{T}} v = 1, L^{\mathrm{T}} v = 0$, where $t \to \infty, e^{-Lt} \to 1_N v^{\mathrm{T}}$.

3.1. Formation control

Referring to the previous work, the consensus protocols can be divided into two types: One is static and the other one is dynamic. Based on the static consensus protocol, the following formation control protocol is employed here as (8)

$$u_i = \alpha \sum_{j \in N_i} (\delta_j - \delta_i) - \beta v_i - \gamma_1 \Omega_i - \gamma_2 \dot{\Omega}_i \quad (8)$$

where α, β, γ_1 and γ_2 all of them are positive gains. $\delta_i = [\delta_i^x, \delta_i^y, \delta_i^z]^{\mathrm{T}}, v_i = [\dot{x}_i, \dot{y}_i, \dot{z}_i]^{\mathrm{T}}$, $\Omega_i = [g\theta_i, -g\phi_i, 0]^{\mathrm{T}}, \dot{\Omega}_i = [g\dot{\theta}_i, -g\dot{\phi}_i, 0]^{\mathrm{T}}$.

Theorem 1. Assume G is the connected undirected graph. The system (4) can realize the formation as defined in (7) if protocol (8) satisfies the following conditions:

$$\alpha > 0, \beta > 0, \gamma_1 > 0, \gamma_2 > 0, \beta \gg \alpha,$$
$$\gamma_1, \gamma_2 > \beta \gg \alpha, \beta\gamma_1\gamma_2 > \beta^2 + \gamma_2^2 \alpha$$

Proof. Please referring to Reference item [5]. □

3.2. Optimal control

The solution of optimal control needs to be able to obtain all the state information in the system, which is often not satisfied in the multi-agent system. However, the optimal control law in a single agent can be studied to determine the optimal control law of the system.

Before giving the performance function, let $\bar{X} = [\delta^x \quad \delta^y \quad \delta^z \quad \dot{x} \quad \dot{y} \quad \dot{z} \quad g\theta \quad -g\phi \quad 0 \quad g\dot{\theta} \quad -g\dot{\phi} \quad 0]^{\mathrm{T}}$, then the performance function is defined as

$$J_i = \int_0^\infty \left[\bar{X}_i^{\mathrm{T}}(t) Q \bar{X}_i(t) + u_i^{\mathrm{T}}(t) R u_i(t)\right] dt \quad (9)$$

Since the states of UAV on different coordinate axes are independent, we need to set weight matrices $Q = q * I_{12}, R = r * I_3$, where $q > 0, r > 0$.

Through the optimal control theory, the optimal control law of single agent is

$$u_i^* = -R^{-1} B^T P \bar{X} \quad (10)$$

where P is the solution of the Riccati equation (11).

$$A^{\mathrm{T}} P + PA - PBR^{-1} B^{\mathrm{T}} P + Q = 0 \quad (11)$$

3.3. Distributed optimization formation control

Through the above calculation, the optimal control law (10) can be obtained. Let $K = R^{-1} B^T P$, the dimension of K must be 3×12, and it also has the following form $K = [k_1 \quad k_2 \quad k_3 \quad k_4] \otimes I_3 =$

$$\begin{bmatrix} k_1 & 0 & 0 & k_2 & 0 & 0 & k_3 & 0 & 0 & k_4 & 0 & 0 \\ 0 & k_1 & 0 & 0 & k_2 & 0 & 0 & k_3 & 0 & 0 & k_4 & 0 \\ 0 & 0 & k_1 & 0 & 0 & k_2 & 0 & 0 & k_3 & 0 & 0 & k_4 \end{bmatrix}^{\mathrm{T}}$$

The multi-agent system of UAV is isomorphic, that is, the dynamic performance of all agents is the same, so the optimal control law can be directly applied to the UAV system. Therefore, the optimal formation control law is obtained as follows

$$u_i = k_1 \sum_{j \in N_i} (\delta_j - \delta_i) - k_2 v_i - k_3 \Omega_i - k_4 \dot{\Omega}_i \quad (12)$$

where k_1, k_2, k_3, k_4 come from matrix K.

Theorem 2. Assume G is connected undirected graph. The UAV system (4) can complete the formation while making the performance function (9) optimal if it uses the protocol (12).

Proof. Substitute (1) to (2), we can get

$$U = \Gamma \otimes I_3 \bar{X}$$
$$\Gamma = L \otimes [k_1 \quad 0 \quad 0 \quad 0] + I_3 \otimes [0 \quad k_2 \quad k_3 \quad k_4]$$

Then

$$\dot{\bar{X}} = A \bar{X} + B \Gamma \otimes I_3 \bar{X} = \bar{\Gamma} \bar{X}$$

Combining the proof of Theorem 1, $\bar{\Gamma}$ can be turned into a Jordan standard:

$$\bar{\Gamma} = P J P^{-1}$$

Let v_1^T is the first row of P^{-1}, the left eigenvalue of the zero eigenvalue, and ω_1 is the right eigenvector of the zero eigenvalue of the first column of P. So $v_1^{\mathrm{T}} \omega_1$ is equal to 1. By taking time to infinity, the state of the system becomes:

$$\lim_{t \to \infty} \bar{X} = \lim_{t \to \infty} e^{\bar{\Gamma} t} \bar{X}(0)$$
$$e^{\bar{\Gamma} t} \bar{X}(0) \to (\omega_1 v_1^{\mathrm{T}}) \bar{X}(0), \qquad t \to \infty$$

According to Lemma 1, as time tends to infinity, the system (4) is asymptotically consensus and the error vector is 0, that is, the system completes the formation control task. □

4. Simulations

In this section, numerical simulation and virtual platform simulation were carried out to verify the effectiveness of the designed formation control protocol. In this section, numerical simulation and virtual platform simulation were carried out to verify the effectiveness of the designed formation control protocol.

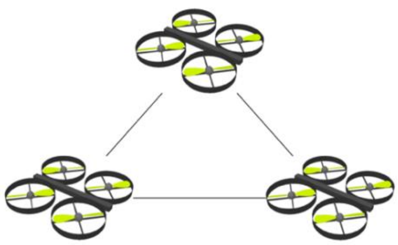


Fig.2. System communication topology diagram

4.1. Experiment 1

The numerical experiments have verified the formation control and distributed optimization formation control respectively, as shown in Fig.3 and Fig.4.

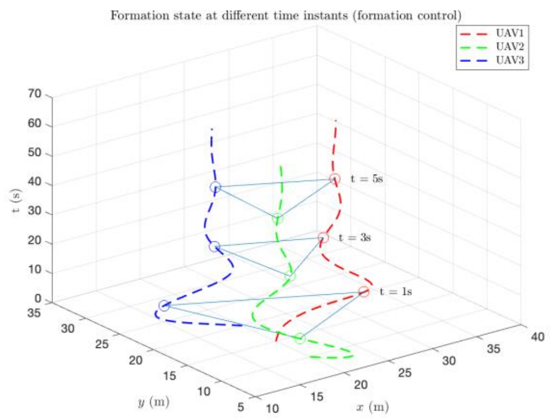


Fig.3. System states formation based on formation control

In the formation control, it can be observed that the system can complete the triangular formation, but there is a large error from the set position during the completion of the formation, and there is still an error at the 5th second. When using distributed optimal formation control, it can be observed that the system can quickly complete the triangular formation, and the error between the actual position and the set value is small.

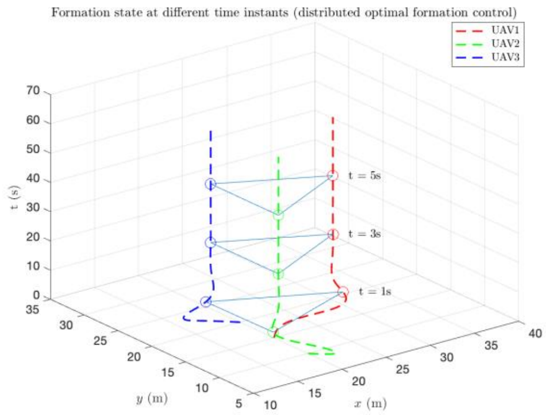


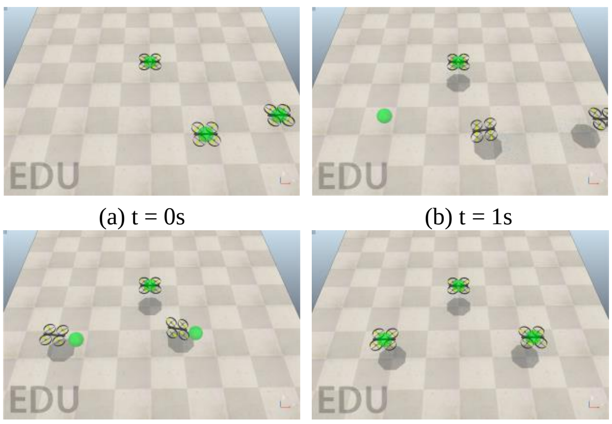
Fig.4. System states based on distributed optimization formation control

In addition, it is surprising that the optimal control not only reduces the loss of the system but also accelerates the convergence speed of the system, which is of great help in reducing the time it takes for the system to form a formation.

4.2. Experiment 2

The simulation software CoppeliaSim/Vrep was also used to verify the flight of the UAV formation. To keep the system in view during the operation, we assume that a UAV remains stationary.

Observing the position of the system at different time instants, as shown in Fig.5, it can be seen that the system can finally complete the triangular formation. The full version of the experiment video can be seen at https://www.bilibili.com/video/BV1Br4y1D7VJ/.



(a) t = 0s (b) t = 1s (c) t = 3s (d) t = 5s

Fig.5. The position of three UAVs at different time instants

5. Conclusions

In this paper, a multi-UAV system is established by analyzing the dynamic model of UAV. Based on the static consensus protocol, the formation control protocol is designed. Aiming at the problem that some UAVs in the system cannot obtain the status information of other UAVs, an optimal control law within a single UAV is designed. Finally, combining the formation control protocol and the optimal control law, a distributed optimization formation control protocol for the multi-UAV system is designed. This protocol is not interfered with by the communication topology of the multi-agent system and can optimize the performance function while the system completes the formation task.

In the next step, we plan to combine engineering applications and apply theoretical research results to practice.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China (No. 61340041, 61374079 and 61903126) and the Natural Science Foundation of Henan Province (182300410112).

References

1. J. Zhao, F. Dai, Y. Song, Analysis of the consensus protocol of heterogeneous agents with time-delays, in: *2021 Proceedings of International Conference on Artificial Life and Robotics (ICAROB)*, 2021, pp. 592-595.
2. Song Y, Zhao W. Multi-agent system rendezvous via refined social system and individual roles. *Wseas transactions on systems and control*, 2014, 9(1): 526-532.
3. Wang J, Han L, Dong X, et al. Distributed sliding mode control for time-varying formation tracking of multi-UAV system with a dynamic leader. *Aerospace Science and Technology*, 2021, 111: 106549.
4. Ren Y, Wang Q, Duan Z. Optimal Distributed Leader-following Consensus of Linear Multi-agent Systems: A Dynamic Average Consensus-based Approach. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2021.
5. J. Zhao, F. Dai, Y. Song, Consensus of Heterogeneous Mixed-Order Multi-agent Systems Including UGV and UAV, *Proceedings of 2021 Chinese Intelligent Systems Conference* 805 (2022) 202–210.
6. Zhu J, Tian Y P, Kuang J. On the general consensus protocol of multi-agent systems with double-integrator dynamics. *Linear Algebra and its Applications*, 2009, 431(5-7): 701-715.

Authors Introduction

Mr. Jichao Zhao

He is a third-year master candidate in Tianjin University of Science and Technology. His research area is about multi-agent system, control engineering.

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.

Dr. Yunzhong Song

He is affiliated with the School of Electrical Engineering and Automation, Henan Polytechnic University as a full professor. He received a PhD degree from the Department of Systems Science, Zhejiang University, P. R. China in 2006. His research interest covers multi-agent systems analysis and control, complex networks as well as nonlinear systems analysis and control. Up to now, he has published about 100 papers in peer referred journals and conference proceedings. To be specific, the work initiated by Professor SONG such as consensus of multi-agent systems with hybrid order dynamics, and social role based multi-agent systems flocking as well as hybrid-mode multi-agent systems regulation have been recognized.

A Design of Multi-Agent System Simulation Platform Based on Unmanned Ground Vehicles and A Research on Formation Control Protocol

Chuang Zhang [1], Jichao Zhao [1], Fengzhi Dai [1, 2 *]

[1] *College of Electronic Information and Automation, Tianjin University of Science and Technology, 300222, China;*

[2] *Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China*

*E-mail: * daifz@tust.edu.cn*

www.tust.edu.cn

Abstract

Inspired by ants and fishes in nature, the multi-agent system (MAS) plays a huge role in production and manufacturing in modern society. This paper takes unmanned ground vehicles (UGVs) as the object and designs a multi-vehicle test platform, which is composed of UGVs, ultra-wideband (UWB), and Bluetooth Mesh. At the same time, the UGV can achieve different control objectives by changing the main controller. The formation control protocol of UGVs is designed and its stability is analyzed. The effectiveness of the protocol is verified by numerical simulation. Finally, the designed control protocol is applied to the developed hardware test platform to verify the effectiveness of the experimental platform.

Keywords: multi-agent system, UGV, simulation platform, consensus, formation

1. Introduction

Multi-agent system (MAS) technology plays a huge role in industry and manufacturing. The technology of the multi-agent system is mainly inspired by ants, fishes, birds in nature, etc., to study the application of clustered individuals. Theoretically speaking, the agent can be anything, such as UGV, UAV (unmanned aerial vehicle), sensor, robotic arm, etc.

At present, the theoretical research on multi-agent systems is gradually becoming mature, such as the consensus control [1], rendezvous control [2], and formation control [3] of multi-agent systems. But at the application level, there are still many areas that need to be improved.

This paper designs a physical simulation platform of a multi-agent system to verify its application in actual scenarios. This platform adopts a layered structure design, which mainly includes a communication layer, an agent layer, and a positioning layer. The communication layer is mainly to establish communication channels between agents so that any agent can communicate with all other agents. The agent layer contains all the individuals that make up the system. The positioning layer plays a role in positioning each agent.

A UGV is designed as an agent of the platform, which can replace the main controller to match different programming languages, with great flexibility. Then, explains the principle of Bluetooth mesh communication technology and the principle of UWB positioning.

The dynamic consensus and static consensus of the multi-agent system are analyzed, and the static formation control protocol of the system is designed based on the static consensus. Finally, a numerical simulation is performed to verify the validity of the static formation control protocol designed. After that, the static formation control protocol is applied to the physical simulation platform to verify the effectiveness of the physical simulation platform.

2. Multi-agent System Simulation Platform

The multi-agent system simulation platform consists of three parts: 1) Agents. Agents could be anything, such as UGVs, UAVs, sensors, manipulators, etc. This paper takes the UGV as an example. 2) Communication network. The agent needs to know the status information of other agents when it is working. The purpose of Bluetooth mesh is to establish the communication channel between agents. 3) UWB positioning system. The position states of all agents come from this system.

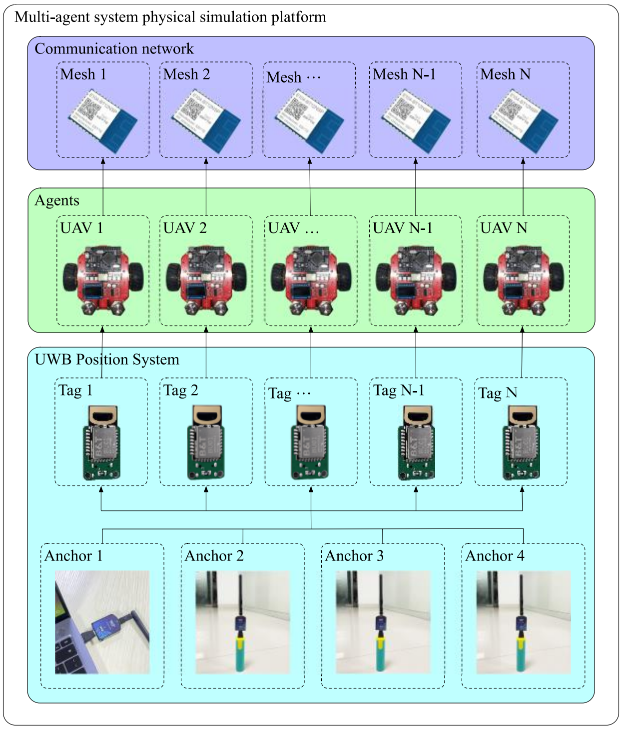


Fig.1. Multi-agent system physical simulation platform

2.1. UGV architecture

This article uses UGV as the research object as shown in Fig.2. The main controller of UGV can use Arduino, STM32, and Raspberry Pi. To achieve precise positioning, the motor pulse must be in a precise and controllable state, so the pulse value returned by the encoder combined with the PID algorithm forms a closed-loop control PWM output. To adjust the PID parameters more conveniently, the LCD color screen is used to display the return value of the encoder and the battery voltage collected by the ADC. The main purpose of real-time voltage acquisition is to keep the motor regulation within a fixed voltage range and avoid the influence of current differences.

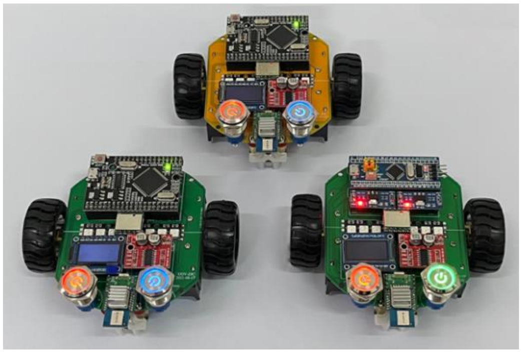


Fig.2. UGV agents

2.2. Communication network

The communication network uses Bluetooth mesh technology, and the chip model is E104-BT12NSP. The module supports the Sig Mesh and 2.4Ghz Bluetooth band. In the networking, each node can act as a proxy node or a common node, and the communication distance between each node can reach up to 50 meters. If a node is used as a relay, the received signal can be repaired and amplified, and the orientation of the network can be larger.

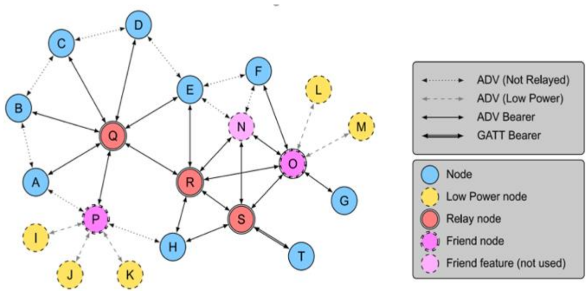


Fig.3. Bluetooth mesh networking diagram

2.3. UWB Positioning system

UWB positioning technology can be applied to the positioning, tracking, and navigation of indoor static or moving objects and people, and can provide very accurate positioning information.

UWB positioning technology uses TOF (time of flight) ranging. The TOF ranging method is a two-way ranging technology. It mainly uses the flight time of the signal between two transceivers to measure the distance between nodes. The module will generate an independent timestamp from the start. The transmitter of module A transmits a requested pulse signal at A1 of its time stamp, and module B transmits a responsive signal at time B2, which is received by module A at time A2 of its timestamp. The flight time of the pulse signal between the two modules can be calculated to determine the flight distance. Because in the line-of-sight environment, the TOF-based ranging method has a linear relationship with the distance, so the measurement results will be more accurate.

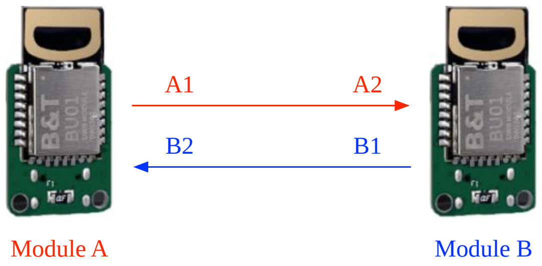


Fig.4. UWB ranging principle diagram

3. Formation Algorithm

In this part, we first elaborated the dynamic model of UGV, then introduced the consensus protocol of common second-order systems, and finally constructed the formation control protocol based on the consensus protocol.

Since the UGV above mentioned is a two-wheel differential car, we assume that the UGV first adjusts its direction when it moves, and then starts to move. In this way, we can use the common second-order system model as the dynamics model of the agent. The system dynamics model is expressed as

$$\begin{aligned} \dot{p}_i &= v_i \\ \dot{v}_i &= u_i \end{aligned} \qquad (1)$$

where $p_i = [p^x \quad p^y]^{\mathrm{T}}, v_i = [v^x \quad v^y]^{\mathrm{T}}$ represent the position and velocity of UGV, respectively. $u_i = [u^x \quad u^y]^{\mathrm{T}}$ is the control input.

3.1. Consensus protocol

The dynamic consensus protocol and static consensus protocol of the second-order system as formula (2) and formula (3) respectively.

$$u_i = \alpha \sum_{j\in N_i}(p_j - p_i) + \beta \sum_{j\in N_i}(v_j - v_i) \qquad (2)$$

$$u_i = \alpha \sum_{j\in N_i}(p_j - p_i) - \beta v_i \qquad (3)$$

where α, β are positive gain, N_i represent the neighbors of the node i [4].

The difference between dynamic consensus protocol and static consensus protocol is whether the final state of the system is static, that is, the final velocity state of dynamic consensus is equal but not necessarily 0, the position is in a constantly changing state, but the final velocity state of static consensus equal 0, the position remains stationary.

3.2. Formation protocol

Based on the static consensus protocol (3) of the system, this paper designs the following static formation control protocol (4) [5,6].

$$u_i = \alpha \sum_{j\in N_i}\big((p_j - h_j) - (p_i - h_i)\big) - \beta v_i \qquad (4)$$

where $h_i = [h^x \quad h^y]^{\mathrm{T}}$ is the formation information of agent i.

This agreement (4) can ensure that the system completes the formation within a limited time and remains stationary. Formula (2) can also be changed to a dynamic formation protocol, but it becomes another type of time-varying formation problem, and we will not discuss it here for the time being.

4. Simulations

In this part, we carried out the numerical simulation experiment and the experiment of the simulation platform respectively. The simulation experiment is set as the system has three agents and the communication relationship is shown in Fig.5. The ultimate goal of simulations is to complete the triangular array.

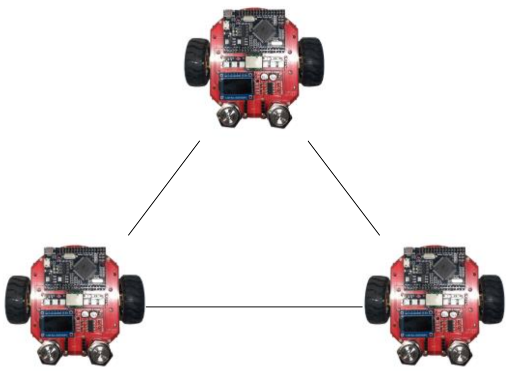


Fig.5. Multi-UGV system communication topology

4.1. Experiment 1

Set the initial states of the three UGVs as follows

$$p_1(0) = [20 \quad 15]^{\mathrm{T}}, v_1(0) = [1 \quad 1]^{\mathrm{T}},$$
$$p_2(0) = [20 \quad 10]^{\mathrm{T}}, v_2(0) = [1 \quad -2]^{\mathrm{T}},$$
$$p_3(0) = [20 \quad 20]^{\mathrm{T}}, v_3(0) = [-2 \quad 1]^{\mathrm{T}}.$$

The positive gains are $\alpha = 0.3, \beta = 1.5$. The formation control protocol is equation (4). Fig.6 and Fig.7 show the simulation results on X-axis and Y-axis respectively.

Through observing the experimental results, each agent can reach the specified position in a limited time, that is to say, the system has completed the set formation requirements.

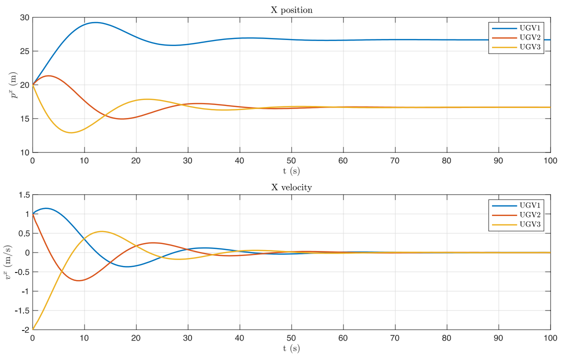


Fig.6. The states of UGVs along X-axis

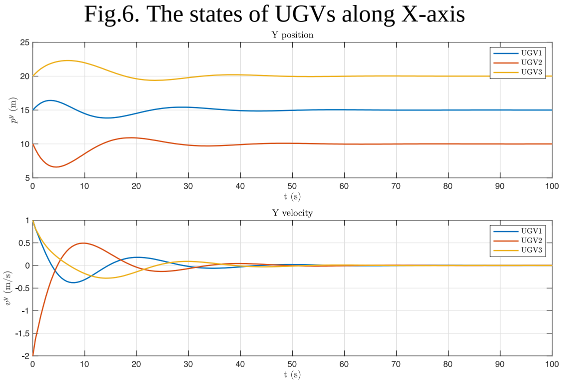


Fig.7. The states of UGVs along Y-axis

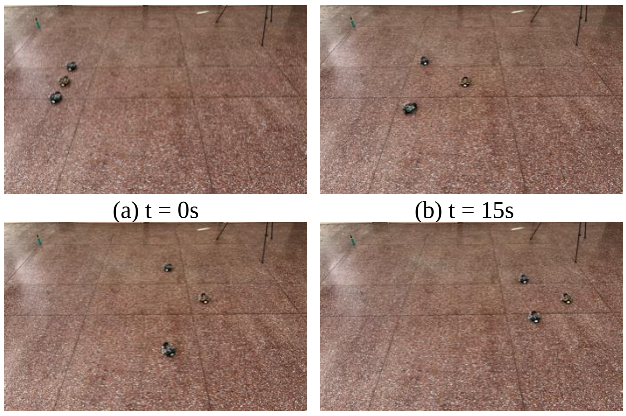
4.2. Experiment 2

After the correctness of the formation protocol is verified by numerical simulation, the formation control protocol (4) is applied to the designed physical simulation platform.

The experimental platform is shown in Fig.8. Four UWB as anchors is distributed at four corners of the experimental site to provide auxiliary location information. There are also three UWB modules mounted on the UGV to provide position coordinates. The Bluetooth module on the UGV assists communication, enabling an agent to obtain information such as position, speed, and control input of other agents.

Fig.8. Experimental scene diagram

The running effect of the simulation platform is shown in Fig.9, showing the different positions of the three UGVs at different time instants. At the beginning of the experiment, the three UGVs were lined as the line (t=0s) and then began to move away (t=15s) gradually due to the effect of the control protocol. Due to the influence of the speed, the distance would be large (t=30s), and then returned to the predetermined formation position (t=45s) to complete the triangular formation.



(a) t = 0s (b) t = 15s

(c) t = 30s (d) t = 45s

Fig.9. The operating positions of the three UGVs at different time instants

5. Conclusions

This paper designs a multi-agent system simulation platform and introduces in detail the three layers that make up the simulation platform, including the agent layer, the communication layer, and the positioning layer. A type of UGV is designed, and the UGV is used as an agent to verify the effectiveness of the simulation platform. The consensus control protocol is analyzed, and the formation control protocol is designed based on the consensus protocol and applied to the simulation platform.

In the next step, we will design different types of agents and integrate them into our simulation platform to expand the use of the platform.

References

1. Olfati-Saber R, Murray R M. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on automatic control*, 2004, 49(9): 1520-1533.
2. Song Y, Zhao W. Multi-agent system rendezvous via refined social system and individual roles. *Wseas transactions on systems and control*, 2014, 9(1): 526-532.
3. Song Y Z. Consensus of agents with mixed linear discrete dynamics. *International Journal of Control, Automation and Systems*, 2016, 14(4): 1139-1143.
4. J. Zhao, F. Dai, Y. Song, Analysis of the consensus protocol of heterogeneous agents with time-delays, in: *2021 Proceedings of International Conference on Artificial Life and Robotics (ICAROB)*, 2021, pp. 592-595.
5. J. Zhao, F. Dai, Y. Song, Consensus of Heterogeneous Mixed-Order Multi-agent Systems Including UGV and UAV, *Proceedings of 2021 Chinese Intelligent Systems Conference* 805 (2022) 202–210.
6. Song Y. Consensus of heterogeneous agents with linear discrete dynamics. *2015 34th Chinese Control Conference (CCC)*. IEEE, 2015: 7416-7422.

Authors Introduction

Mr. Chuang Zhang

He is a junior undergraduate of Tianjin University of Science and Technology. His research area is embedded development.

Mr. Jichao Zhao

He is a third-year master candidate in Tianjin University of Science and Technology. His research area is about multi-agent system, control engineering.

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.

A Study of Weighted Average Method for Multi-sensor Data Fusion

Peng Lu [1], Fengzhi Dai [1, 2 *]

[1] *Tianjin University of Science and Technology, China;*

[2]*Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China*

*E-mail: * daifz@tust.edu.cn*

www.tust.edu.cn

Abstract

With the development of sensor technology, multi-sensor data fusion has become an important research direction in the field of sensors. Among them, parametric classification algorithms have become the most intensively studied class of algorithms in the field of multi-sensor data fusion, and the weighted average method is the most important one among the parametric classification algorithms. This paper describes the composition and development of parameter classification algorithms, focusing on the process, steps and recent developments of the weighted average method, and uses the algorithm to fuse data from ultrasonic and infrared sensors. The simulation results prove that the weighted average method has a better fusion effect.

Keywords: multi-sensor, data fusion, parametric classification, weighted average method

1. Introduction

In order to meet the needs of data acquisition, a large number and various types of sensors are widely used in various fields such as military equipment, household appliances, automotive industry and medical and health care [1]. If the data that are collected by various sensors are processed separately and in isolation, it will not only increase the workload of data processing, but also cut off the intrinsic connection between the information of each sensor, which will eventually lead to the loss of key information after the combination of each sensor information. Therefore, it has become a research focus in the field of sensors to combine the data collected by multiple sensors organically by utilizing the complementary data characteristics of multiple sensors [2].

In nature, information fusion is an inherent characteristic of living organisms and is a fundamental function prevalent in humans and other animals. In the field of science and technology, the concept of "data fusion" was introduced in the 1960s, initially to address the need for multi-source correlation in military systems. In 1984, the U.S. Department of Defense established the Data Fusion Expert Group (DFS) to direct systematic research on multi-sensor information fusion technology. The International Society for Information Fusion was founded in the United States in 1998. In the 21st century, multi-sensor information fusion technology has gradually been widely used in non-military fields.

As one of the research hotspots of multi-sensor data fusion technology, fusion algorithms have received a lot of attention. Since multi-sensor data fusion involves many theories and technologies, there is no completely unified algorithm that can be adapted to all scenarios, so the appropriate algorithm needs to be selected according to different application contexts. Currently, multi-sensor data fusion algorithms are divided into the following three main categories: physical model-based methods, parameter-based methods, and cognitive theory-based methods. In this paper, we focus on the parameter-based methods.

2. Parameter Classification Algorithm

Parametric classification algorithms are the most common, applied and studied class of algorithms in the history of multi-sensor fusion technology development. Such algorithms can be further divided into those based on statistics and those based on information theoretic techniques. Commonly used algorithms in this category are weighted average, Bayesian estimation, D-S evidence theory, entropy method, neural network, and cluster analysis.

2.1. Bayesian estimation

Bayesian estimation is a method of representing various uncertain information provided by multiple sensors as probabilities and processing them using the Bayesian conditional probability formula in probability theory.

Assume that the individual decisions contained in a sample space are independent of each other $A_1, A_2, A_3 \cdots A_n$, and assume that the observations for the system are B. By taking the prior probability $\mathrm{P}(A_i)$ and the conditional probability $\mathrm{P}(\mathrm{B}/A_i)$ through the nature of the sensor itself and a priori knowledge, the probability equation can be obtained.

$$\mathrm{P}(A_i/\mathrm{B}) = \frac{\mathrm{P}(A_i\mathrm{B})}{\mathrm{P(B)}} = \frac{\mathrm{P}(\mathrm{B}/A_i)\mathrm{P}(A_i)}{\sum_{j=1}^{m}\mathrm{P}(\mathrm{B}/A_j)\mathrm{P}(A_j)} \quad (1)$$

where $\mathrm{P}(A_i/\mathrm{B})$ is the posterior probability and $\mathrm{i} = 1,2 \cdots \mathrm{m}$.

This result can be generalized to the case of multiple sensors. When there are n sensors and the observation results are $B_1, B_2, B_3 \cdots B_n$ respectively, the posterior probability of each decision at n sensors can be obtained as equation (2), assuming that they are independent of each other and independent of the observed object condition.

$$\mathrm{P}(A_i/B_1 \wedge B_2 \wedge \cdots \wedge B_n) = \frac{\prod_{k=1}^{n}\mathrm{P}(B_k/A_i)\mathrm{P}(A_i)}{\sum_{j=1}^{m}\prod_{k=1}^{n}\mathrm{P}(B_k/A_j)\mathrm{P}(A_j)} \quad (2)$$

The final decision result can be obtained according to the corresponding specific rules [3].

Bayesian estimation is a common method for the multi-sensor low-level redundant data fusion, and can be better for incomplete information with the added Gaussian noise. By this method, the information collected from each data source is represented by a probability density function, and various constraints are assumed in advance to complete the fusion of uncertainty information.

2.2. D-S evidence theory

The D-S theory of evidence was originally proposed by Dempester in 1967 and was later expanded and developed by Shafer. It is eventually developed into one of the mathematical methods that can better handle uncertainty inference problems [4].

Evidence theory proposes the concepts of belief function Bel(A) and plausibility function Pl(A) to represent the degree of support and the degree of non-doubt for A, respectively, and the interval $[\mathrm{Bel(A)}, \mathrm{Pl(A)}]$ represents the uncertainty of premise A. For the synthesis of multiple confidence levels, let$m_1, m_2, \cdots m_n$denote the confidence allocation of n data respectively, and if they are obtained from mutually independent information, the fused Bel(A) can be expressed as Equation (3).

$$\mathrm{m(A)} = \frac{\sum_{\cap A_i = A}\prod_{i=1}^{n} m_i A_i}{1-k} \quad (3)$$

where $\mathrm{k} = \sum_{\cap A_i = A}\prod_{i=1}^{n} m_i A_i$ denotes the conflict between mass functions.

The theory uses Bel(A) rather than probabilities in the metric, and uses methods that constrain the probability of certain events to construct trust functions without directly accounting for probabilities, making it very suitable for situations where probabilities are difficult to obtain.

2.3. Artificial neural network

Artificial neural network are proposed on the basis of modern neuroscience research results. Neural network is the nonlinear network system that is composed of interconnected neurons with learning, memory, computational capabilities, various processing and intelligent recognition capabilities [5]. In a multi-sensor system, there is some uncertainty in the environmental information provided by individual sensors. In contrast, artificial neural network can be expressed in the network structure in terms of network weights based on the similarity of the samples received by the current system.

The implementation of multi-sensor data fusion in artificial neural network begins with the selection of the model, topology and learning rules of the neural network according to the requirements of the intelligent system and the form of sensor data fusion. The input information of the sensor is processed into a global input function, and the function mapping is defined as the mapping function of related units and this function mapping is defined as a mapping function for the relevant units. The statistical laws of the environment are reflected into the structure of the network itself through the interaction of the artificial neural network with the environment, and then the sensor output information is learned, understood, and the assignment of weights is determined to complete knowledge acquisition and information fusion.

2.4. Weighted average method

The weighted average method is the most easily understood and most used method in the parameter classification fusion algorithm. The weighted average method is a simple and intuitive way to fuse data by weighting the data from each sensor.

The main steps are: assume that the number of sensors is n and these sensors jointly monitor the same target, the data collected by the sensors is x_i, where $i = 1,2,\cdots n$, then the weighted average of all the collected data is obtained.

$$\bar{x} = \sum_{i=1}^{n} \varphi_i x_i \quad (4)$$

Where φ_i is the weighting factor for each sensor.

The structure of the weighted average method model is shown in Fig.1.

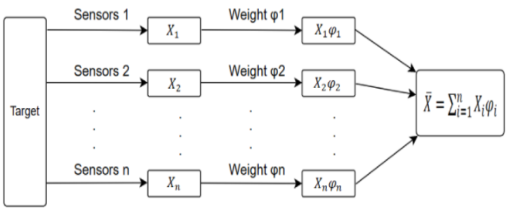


Fig.1 Weighted average method model structure

When using this method, it is necessary to reasonably determine the weighting coefficients of each sensor to ensure that the final fusion results are more accurate.

3. Test of Weighted Average Method

The main purpose of multi-sensor data fusion is to synthesize incomplete data provided by multiple sensors about a particular environmental feature to form a relatively complete and consistent sensory description for more accurate recognition and judgment functions.

In this paper, the data fusion simulation of data from two sets of sensors is performed under Windows 10 based on MATLAB 2019b software using weighted average method.

The distance data collected from ultrasonic and infrared sensor simulations were first fused. Five sets of typical data were grabbed from all fused data for comparative analysis, whcih is shown in Table 1.

Table 1. Motor parameters

Test value (cm)	Fusion results (cm)	Error
80	80.143	0.18%
80	79.856	0.18%
80	80.541	0.68%
80	79.992	0.01%
80	80.471	0.59%

From the data in the table, it can be seen that the fusion results of ultrasonic sensor and infrared sensor range data have less than 1% error, which shows the better fusion effect of the weighted average method.

At the same time, the image data fusion test was performed on the image information collected by the two vision sensors using the weighted average method.

In the first test, the values of weights for the high color vision sensor were set to larger, and the test is shown in Fig.2, where the top left image is taken by the high resolution but color distorted vision sensor, the top right image is taken by the good color but low resolution vision sensor, and the bottom image is the test result after fusion of these two vision sensor data.

Fig.2 High color image fusion test

As can be seen from the above figure, the images captured by the two vision sensors are weighted by the image data fusion, and the fused images are more vivid in color and have a certain improvement in resolution.

But we still can see that the details of the objects in the test result are rather blurred. Therefore, in the second image data weighting fusion, the weights of the high-resolution vision sensors are set larger, and the fusion result is shown in Fig.3.

Fig.3 High definition image fusion test

As can be seen from the above tests, after the second fusion, although some areas of the image are not full of color, the overall fusion effect is satisfactory because the details of the objects are clearer and the colors are brighter.

4. Conclusion

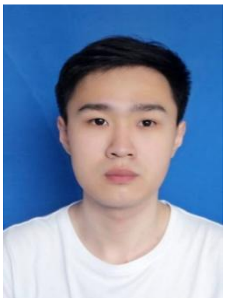
This paper presents the theory and development of parameter classification algorithms in multi-sensor data fusion techniques. In particular, the Bayesian estimation, D-S evidence theory, artificial neural networks and weighted averaging are introduced. At the end of the paper, the weighted average method is tested for data fusion simulation, and the test results show that the method has obvious fusion effect and has strong practicality.

References

1. Varshney P K, Multisensor data fusion, *Electronics & Communication Engineering Journal*, 1997, 9(6): pp.245-253.
2. Khaleghi B, Khamis A, Karray F O, et al., Multisensor data fusion: A review of the state-of-the-art, *Information Fusion*, 2013, 14(1): pp.28-44.
3. Fang H, Tian N, Wang Y, et al., Nonlinear Bayesian Estimation: From Kalman Filtering to a Broader Horizon, *IEEE/CAA Journal of Automatica Sinica*, 2018, 5(2): pp.401-417.
4. Xiao F Y, A Novel Evidence Theory and Fuzzy Preference Approach-Based Multi-Sensor Data Fusion Technique for Fault Diagnosis, *Sensors*, 2017, 17(11): pp.2-20.
5. Rosa, João & Guerra, Daniel & Horta, Nuno & Martins, Ricardo & Lourenço, *Overview of Artificial Neural Networks*, DIO: 10.1007/978-3-030-35743-6_3.

Authors Introduction

Mr. Peng Lu



He is a second year master student at Tianjin University of Science and Technology. His research is about multi-sensor fusion, artificial intelligence, and robotics.

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 the College of Electronic Information and Automation, Tianjin University of Science and Technology, China.

A Design of Wide-angle Open and Close Multifunctional Smart Windows

Zilong Liu [1], Luqi Shen [1], Jiarun Xu [1], Longyu Gao [1], Fengzhi Dai [1, 2 *]

[1] *Tianjin University of Science and Technology, China;*

[2] *Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China*

E-mail: [*] *daifz@tust.edu.cn*

www.tust.edu.cn

Abstract

Current smart windows always have low intelligence and a few functions. And almost all kinds of traditional windows have the problem of inconvenient cleaning outside. The paper proposed a new mechanical structure design of the window, so that the window can be turned by nearly 180 degrees, which is convenient for cleaning the inside and outside of the window. In terms of intelligent systems, a variety of sensors combined with Internet modules are used to realize the remote intelligent control (turning with the wind, and automatically closing windows when there is no one at home when bad weather). The added child mode uses visual inspection to identify children. When a child is detected, the window will be automatically closed to prevent the child from climbing the window and falling from the building.

Keywords: Smart window, new structure, machine vision

1. Introduction

With the development of science and technology and the improvement of people's living standards, the global smart windows are developing rapidly. Thanks to the development of construction industry, the current smart window market in my country is growing steadily year by year, becoming the future development trend of the door and window industry and occupying an important market position.

In the process of research on smart windows, it is found that existing smart windows have obvious defects. The specific manifestation is: low degree of intelligence, some products can only realize remote control.

Although the technology has been upgraded, it is only electric, not smart. Secondly, all types of smart windows have a single function. For example, smart dimming glass windows are researched on the permeability of glass; smart security windows are developed in terms of automatic alarms.

The traditional intelligentization of windows is mainly to upgrade the indoor brightness, voice control, and anti-theft. It has few functions, low practicality, and low degree of universality, so it is limited in practical applications. The traditional smart window is shown in Fig 1.

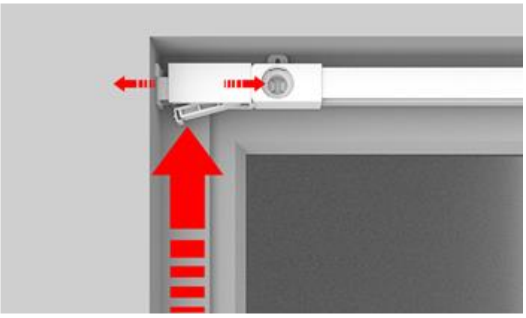


Fig.1. Traditional smart windows

At the same time, we have noticed that almost all smart windows are only functionally studied. The structure of the windows has not been upgraded [1]. Therefore, in addition to the limitations of non-intelligence for various types of smart windows, the problem of the outside of the window that is not conducive to cleaning still exists, which has become a major pain point for users of smart windows. The structure of a traditional window is shown in Fig.2.

Fig.2. Traditional window structure

In response to the existing problems of smart windows, a new structure of smart windows was designed by our project team. It combines the advantages of traditional casement windows and smart windows to improve the shortcomings of the two types of windows to the greatest extent.

It not only solves the problem that the outside of the window is difficult to clean, but also realizes the diversity of windows in terms of functions, which can cope with various scenarios, and is more intelligent and humane.

2. Hardware Design

ESP8266 is used as the core controller of the control system. The sensor part includes compass module, temperature and humidity sensor, air dust sensor, rain sensor, harmful gas sensor, combustible gas sensor [2].

The 1.3-inch display screen is used to display various data of the window. Stepper motors and stepper motor drivers are used to control the mechanical structure of the window. The main elements of the window are shown in Fig.3.

2.1. Hardware connection

The single-chip microcomputer is connected to the sensor to receive various data and conduct a comprehensive analysis, and through the motor control the mechanical structure part, complete the control of the window state. And the necessary information is displayed on the display, so that the user can grasp the status of the window in real time. The connection mode of each sensor and the one-chip computer is shown as in Fig.4.

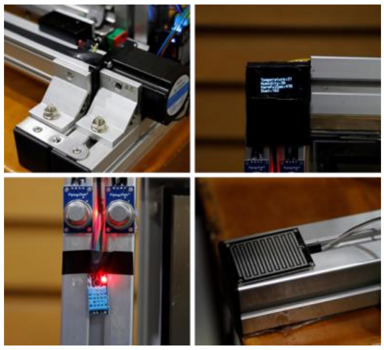


Fig.3. The main elements of the window

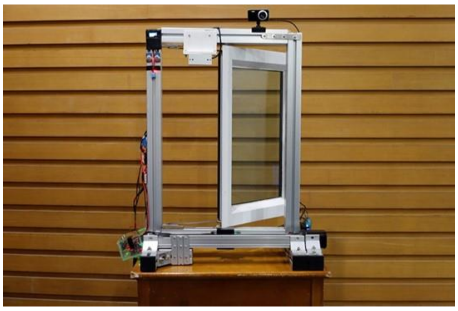


Fig.4. Nearly 180° flipped windows

A stepping motor is connected between the controller and the mechanical structure of the window, and the stepping motor is controlled by sending a pulse signal to the motor driver to open or close the window.

The Bluetooth module mounted on the window is connected to the controller through the USART bus to realize short-distance wireless transceiver function. The circuit connection diagram of the control system is shown in Figure 5.

2.2. Structural design

The clever design of the mechanical structure is the key to the large-angle opening and closing of the smart window. By adding the connecting rod design (the motor drives the lead screw, the slide moves horizontally through the movement of the lead screw, and the window is rotated by the connecting rod during the movement of the slide), only one motor is needed to realize the nearly 180° flip of the window.

The connecting rod structure and the window mechanical structure are shown in Figures 6 and 7 respectively.

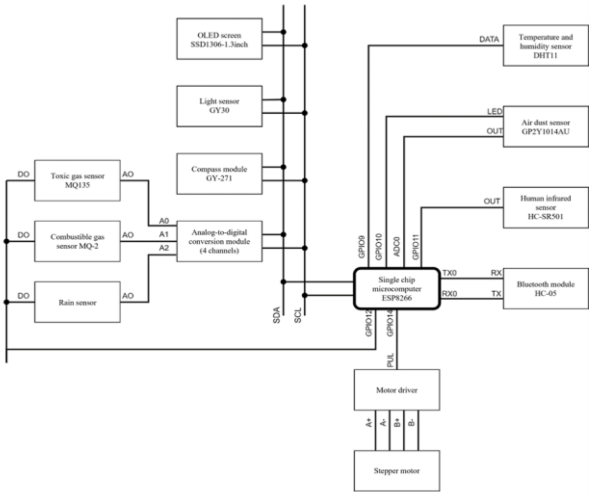


Fig.5. Control system connection diagram

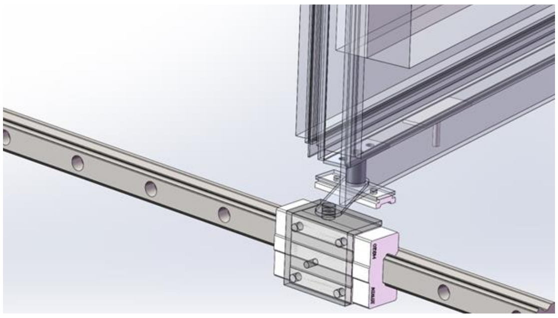


Fig.6. Connection rod structure

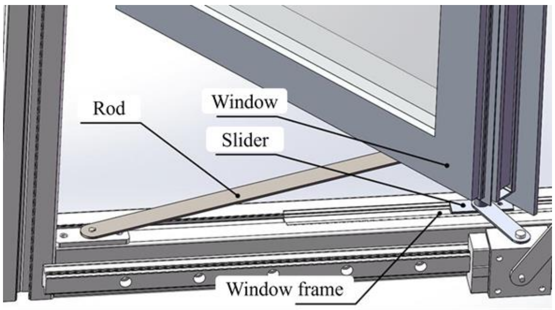


Fig.7. Mechanical structure drawing

Fig.8 shows the three typical open states of windows, namely 30°, 90° and 160°. Compared with existing windows, it effectively solves the problem that the outside of the window is difficult to clean, thereby greatly reducing the risk of falling. Clean the outside of the window is shown in Fig.9.

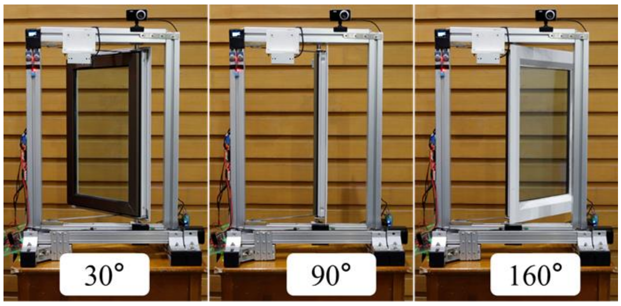


Fig.8. Mechanical structure drawing

Fig.9. Clean the outside of the window

3. Algorithm Design

3.1. Control algorithm design

In a general program cycle, the controller uses polling to obtain external information and display it on the display. However, due to sudden conditions such as rain outside the window, toxic gas, combustible gas and so on indoors, higher priority countermeasures are required. Therefore, various sensors can send interrupt signals to the controller through digital output pins, which ensures priority Deal with emergencies. The work flow chart is shown in Figure 10, which can realize the following functions:

- When the weather is bad and there is no one in the house, the windows will be closed automatically [3];
- Automatically open windows when the indoor air is bad or contains harmful gas;
- Adjust the opening and closing angle by itself according to the external wind direction.

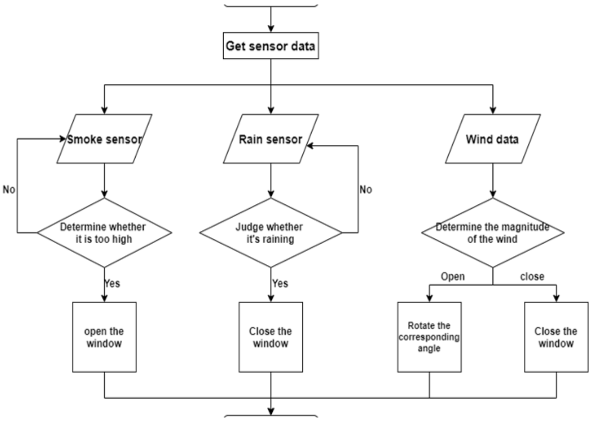


Fig.10. work flow chart

3.2. Child detection algorithm design

When the user is within the shooting range of the camera, the camera will automatically find and detect the user's face image.

A fixed-length feature vector is extracted from the convolution feature map, and then the feature vector is input into two fully connected layer networks. The age and gender classification training is performed on the extracted facial features through the random forest algorithm. When the camera detects that the child is close to the window, the window is automatically closed to prevent the child from climbing the window and falling from the building. The result of child face detection and recognition is shown in Fig.11.

Fig.11. Children's face recognition results

4. Conclusion

In view of the difficulty of cleaning the outside of traditional windows, a new type of multi-scene smart window is proposed. The structure is based on the original traditional casement window design, and a linkage mechanism is added, which enables the window to be turned over by nearly 180 degrees, effectively solving the problem of difficult cleaning of the outer side of the window.

At the same time, the window can adjust the opening and closing angle of the window according to the wind direction to ensure the maximum air intake. When the camera detects that there is a child's activity near the window, it will sound an alarm and automatically close the window to prevent accidents caused by children climbing the window. In summary, this wide-angle opening and closing multifunctional smart window will have broad market prospects.

References

1. Liqing Geng, Yuechen Li, Shanfeng Shao, et al., The Design of the Intelligent Window Control System Based on Wireless, *2018 10th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, 2018 Feb: pp. 294-297.
2. Weihai Song, Weili Liu, Yingyue Pan, Design of Intelligent Rainwater Detection Window Based on STM32 Single-Chip Microcomputer, *2019 Chinese Automation Congress (CAC)*, 2019 Nov: pp. 278-281.
3. Zhao Sun, Shuyi Wei, Xiuxia Zhang, Intelligent Window Control System Design Based on Single Chip Microcomputer, *2020 IEEE International Conference on Progress in Informatics and Computing (PIC)*, 2020 Dec: pp. 300-303.

Authors Introduction

Mr. Zilong Liu

He majored in Machinery and electronics engineering in College of Mechanical Engineering, Tianjin University of Science and Technology. He studied mechanical control engineering and other related knowledge, and achieved certain results in related professional competitions.

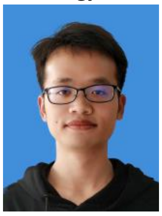
Mr. Luqi Shen

He majored in industrial design at the School of Mechanical Engineering, Tianjin University of Science and Technology. He learned some computer programming and related knowledge .

Mr. Jiarun Xu

He majored in vehicle engineering at the School of Mechanical Engineering, Tianjin University of Science and Technology. He mainly learns about vehicle stability analysis, new energy and other related knowledge.

Mr. Longyu Gao



He majored in automation in College of Electronic Information and Automation, Tianjin University of Science and Technology. He learned SCM and data crawler and other related knowledge, and achieved certain results in relevant professional competitions.

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.

A Low-intensity Laser Control System Design

Yuhui Cheng, Fengzhi Dai *

Tianjin University of Science and Technology, 300222，China;

E-mail: * *daifz@tust.edu.cn*

www.tust.edu.cn

Abstract

Laser-assisted therapy has a good therapeutic effect on specific symptoms in clinical and experimental applications. This paper proposes a low-intensity laser control system design scheme, which is divided into operation unit, control unit and work unit. The main control unit is based on the STM32F407ZGT6 chip design, the operating unit uses a 10.1-inch serial screen, and the working unit consists of 8 low-intensity laser generators. The system can control the power, frequency and working time of 8 laser generators at the same time, and has the function of automatically saving working parameters when the power is turned off.

Keywords: single chip microcomputer, laser, pulse width modulation, power failure protection

1. Introduction

In recent years, laser diagnosis and treatment have achieved a mass of successful applications in the field of medicine [1]. This treatment has many advantages, such as non-invasive, short treatment time and low risk [2]. These characteristics make laser physiotherapy technology very popular.

Low-intensity laser physiotherapy equipment has high requirements for performance, especially the accuracy of control signals. It also requires products with good safety and reliability. In terms of function, for different needs, different laser sources are needed for treatment. How to control multiple laser generators to work together at the same time is a difficult point.

Based on the above discussion, this design is optimized for the precise control of multiple light sources. The design of laser physiotherapy apparatus proposed in this article can control eight laser generators to work at the same time.

In order to ensure the reliability, the parameters manually set in this design can be saved in real time. This way can avoid the loss of the set parameters after the power is cut off due to an emergency. When the device is powered on again, the system can be set to the working parameters of the laser generator soon.

2. The Hardware Structure

The final control effect achieved by the laser controller is closely related to each functional module. According to the hardware structure, the laser physiotherapy instrument can be divided into the following three parts: operation unit, control unit and working unit, which is shown in Fig 1.

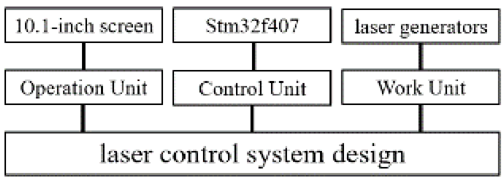


Fig.1 System structure diagram

The operating unit consists of a 10.1-inch touch screen and an emergency stop button. The control unit integrates the main control chip, signal output interface, power supply module and buzzer on a PCB board. The working unit is composed of a laser drive circuit and a laser generator.

The PCB board is designed to reduce circuit complexity and improve the practicability of the device. For a qualified product, the PCB design can directly affect its performance. The design of the PCB board is shown in Fig.2.

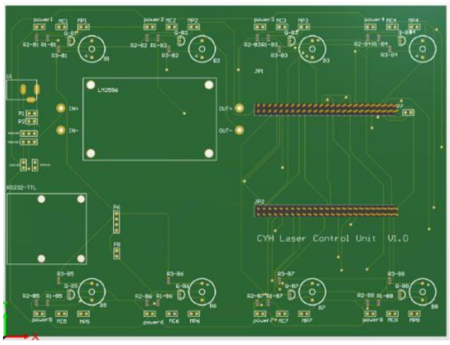


Fig.2. The designed PCB board

2.1. Capacitive touch screen

The DWIN serial capacitive touch screen is used as the data input terminal. Its creen resolution is 1024*600 pixels, and the ideal working power supply is 12V, 1A. The operating temperature is between -20°C and 70°C. The DWIN capacitive touch screen is shown in Fig.3.

Fig.3. Front and back of the DWIN screen

The screen supports RS232 and TTL two serial port modes for communication, and its corresponding working voltages are 12V and 5V respectively. In order to match the working voltage of the single-chip microcomputer, the TTL communication mechanism is adopted in this design. The communication data format is 16-bit hexadecimal characters, and the data format is shown in Fig.4.

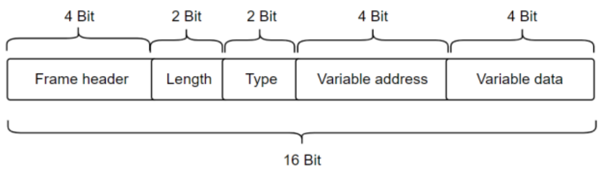


Fig.4. Serial communication data format

2.2. Eemergency module

In order to deal with possible unexpected situations, an emergency stop button is added. The emergency stop button used in this design is a double-in and double-out type, which is usually used to connect the power supply line and the grounding line.

In order to realize the feedback of the emergency stop signal, this design connects the ground terminal interface to the IO port of the single-chip microcomputer. The power supply line end interface is connected to the working unit. The emergency stop button is shown in Fig.5.

Fig.5. Emergency button

2.3. Main control unit

The main control unit includes a control chip, a power supply circuit and a signal output interface. The control chip model is STM32F407ZGT6. The chip adopts a 32-bit Corte-M4 CPU with FPU, the main frequency is up to 168MHZ, and it is equipped with 140 IO ports with terminal functions. The chip has 1MB of Flash memory and 192+4KB RAM. The equipped 17 timers can well meet the requirement of multi-channel PWM signal output [3]. Other related features are as follows:

- Up to 15 communication interfaces, including 6x USARTs running at up to 11.25 Mbit/s, 3x SPI running at up to 45 Mbit/s, 3x I²C, 2x CAN, SDIO
- Analog: two 12-bit DACs, three 12-bit ADCs reaching 2.4 MSPS or 7.2 MSPS in interleaved mode
- Operating Voltage 2.0V ~ 3.6V.
- Operating temperature range: -40 ºC ~ 105 ºC.

The design of the main control chip is shown in Fig.6.

Fig.6. STM32F407ZGT6 chip

2.4. Working unit

The working unit is composed of a laser drive module and a laser module. The laser drive module is shown in Fig.7.

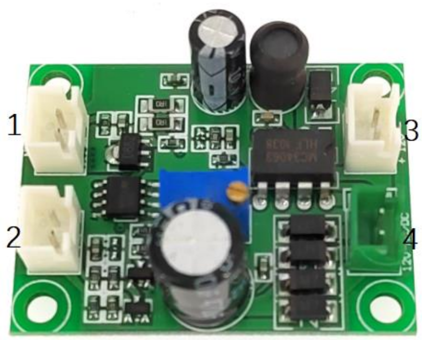


Fig.7. Laser drive module

The laser driver module has four interfaces, the interface 1 is the laser output terminal, the interface 2 is the signal input terminal, and the interfaces 3 and 4 are the power ports. By controlling the signal input of the laser drive module, the power and frequency of the laser output can be adjusted.

The wavelength of the laser source used in this design is 650±20nm. This light source has a wide range of applications in the field of biomedicine. Its working parameters are as follows:

- Operating temperature range: -10 ºC ~ 50ºC.
- Operating voltage: 10~12V
- Operating current: <200mA
- Laser Safety: Class IIIB

The laser generation module is shown in Fig.8.

Fig.8. Laser generation module

3. System Circuit Module Design

This system adopts the STM32407ZGT6 core board, which is equipped with the smallest circuit system of single-chip microcomputer and the download and debug interface.

Therefore, when designing the circuit, only need to consider the signal output interface circuit, voltage conversion circuit and buzzer drive circuit. This kind of scheme is more convenient in design, so as to realize the design of PCB quickly.

3.1. Voltage conversion circuit design

The voltage conversion circuit converts the 12V DC input from the power adapter into 5V DC through the LM2596 module.

LM2596 is a switch type step-down chip, usually is a step-down circuit with a constant voltage output. The chip samples the output voltage value through a resistor and inputs to the chip's feedback terminal. The output voltage can be changed by changing the resistance of the resistor. The schematic diagram of the power circuit design is shown in Fig 9.

Formula (1) is to calculate the Vref. The value of R1 in this circuit is fixed at 2.5KΩ, and the value of R2 is determined by the chip model. When the output voltage Vo is 5V, the resistance value of R2 is 7.6KΩ, then the reference voltage Vref is 1.24V.

$$Vref=Vo*R1/（R1+R2） \quad (1)$$

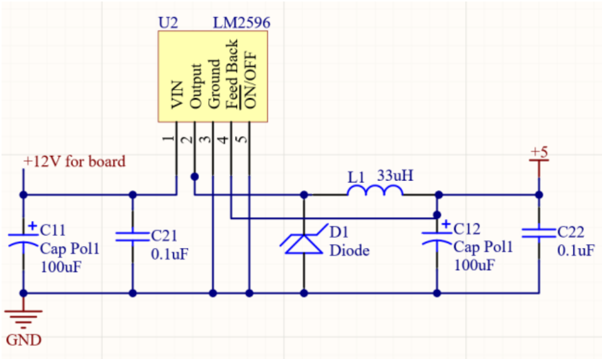


Fig.9. Schematic diagram of the power circuit

3.2. Design of buzzer drive circuit

For an active buzzer, the only thing to do is to input the drive level at the signal end and amplify the drive current through a transistor to make a sound.

Through the test, the actual working time of the buzzer is about 4.7V, and the working current is 30mA.

The schematic diagram of the buzzer drive circuit is shown in Fig.10. In the circuit, R1 and R3 play a current limiting role, and R2 is a pull-down resistor, which can improve the turn-off speed of the transistor.

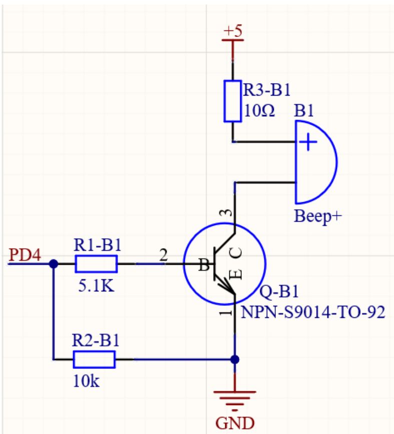


Fig.10. Schematic diagram of the buzzer drive circuit

4. The Functional Module

The various hardware modules of the system ensure the reliability of the system's implementation functions, and the control logic and interactive operations are completed by software codes.

After the system is powered on, enter the initial interface, and enter the preparation interface with a delay of three seconds after confirming to enter the system. At this time, the microcontroller will read the last set parameters and enter the ready state.

When the operator sets the parameters and confirms the work, the system will turn on the laser source and start timing. After the countdown time is over, the corresponding buzzer will give a reminder and turn off the laser. The system work flow chart is shown in Fig.11.

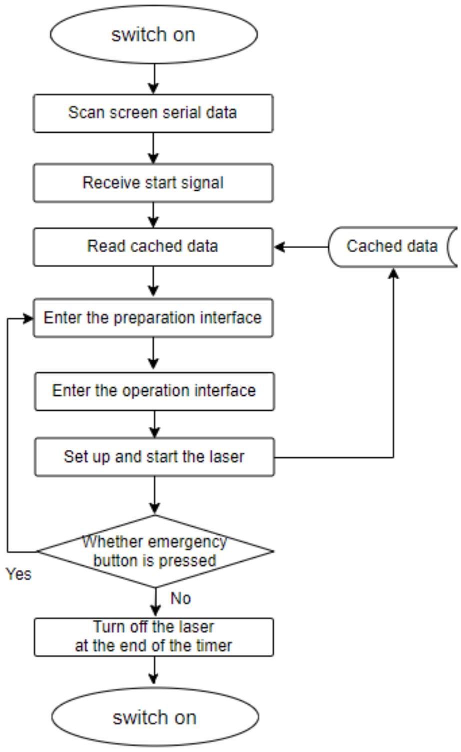


Fig.11. System work flow chart

5. Conclusion

Through testing, it can be concluded that this design can complete the work of controlling eight laser generators at the same time. In continuous mode, four-speed power adjustment can be achieved. In the pulse mode, the default working time is 30 minutes, and each laser generator has a separate timing channel, which can realize the time setting of 0-99 minutes. The default laser power is 50%, and the adjustable range of power is

0~100%. The default laser frequency is 50Hz, and the adjustable range is 1~1K Hz. When an emergency stop is needed, the emergency stop module can also respond quickly and save data.

References

1. Moskvin SV, Agasarov LG. Laser Acupuncture: 35 Years of Successful Application in Russia (Narrative Review). *Journal of lasers in medical sciences vol*. 2020,11: pp.381-389.
2. Tony Y. Chon, Molly J. Mallory, et al., Laser Acupuncture: A Concise Review, Medical Acupuncture. 2019,6: pp.164-168.
3. Cui Lin, Zhu Lei, Bai Lu. Design of PWM waveform generator with adjustable parameters based on STM32 (in Chinese). Information and Communication, 2018, 1: pp.129-131.

Authors Introduction

Mr. Yuhui Cheng

He received his Bachelor's degree from the Internet of Things Engineering, Tianjin University of Science and Technology, China in 2019. He is currently studying for a master's degree in electronic information at Tianjin University of Science and Technology.

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 Tianjin University of Science and Technology, China.

A Review of Machine Vision Based Fruit Recognition Applications

Tianyi Zhang [1], Fengzhi Dai [1, 2 *]

[1] *College of Electronic Information and Automation, Tianjin University of Science and Technology, 300222，China;*

[2] *Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China*

*E-mail: * daifz@tust.edu.cn*

www.tust.edu.cn

Abstract

Machine vision is widely used in various fields. The main applications in the field of fruit picking are fruit identification, fruit quality detection, fruit ripeness detection and grading, etc. And fruit ripeness detection technology is important to improve the quality and market competitiveness of fresh and stored fruits. This paper focuses on the application of machine vision in fruit identification, fruit ripeness detection and grading in the past three years, and the application is more mature in many fruits such as citrus, blueberry, cherry, etc. It uses a number of algorithms to enable accurate identification of fruits and processing of their images to control the robotic arm for a variety of operations such as picking.

Keywords: machine vision, image processing, edge detection, maturity inspection

1. Introduction

With the rising development of the modern economy, consumers have higher requirements for the quantity and quality of fruits. In order to reduce human losses during fruit picking and to save labor for more efficient picking, more and more machine vision technologies are being used in fruit picking and identification detection, etc.

Machine vision is the use of computers to simulate the visual function of the human eye, extracting information from images or image sequences, and identifying the form and motion of three-dimensional objects in the objective world [1]. Its ultimate goal is to use machines to fully interpret, simulate, reproduce and process human vision. In the future intelligent development of agriculture, machine vision technology will play a very important supporting role. And it will also face many technical challenges [2].

This paper is based on the application of machine vision in agriculture, especially in the field of fruit picking and identification in recent years [3, 4]. In the paper, we are writing about the key developments and research dynamics of machine vision technology in fruit recognition, and we do a detailed understanding and description of it. The main technology includes the use of image processing and other algorithms to complete operations. These operations consist of picking, defect identification and sort identification of citrus, navel oranges and other fruits.

2. Fruit (Citrus) Maturity Recognition

Citrus Research Institute of CAAS has proved that citrus peel color can be used as an indicator of citrus fruit maturity after experimental research. It provides the theoretical basis for the final realization of online automatic detection of citrus maturity.

Therefore, Zhuang Yibin and others from Zhejiang University of china have developed a machine vision system for non-destructive detection of citrus maturity. The machine vision system consists of CCD camera, image acquisition card, monitor, light box and computer [2].

According to the contrast of light source effects, the choice is to capture images of citrus against a white background. An image acquisition method was determined by building more than a dozen color models.

This image acquisition method is to obtain color pictures in RGB format through the camera for storage and display, and then convert them to HIS format for easy computer processing. The relationship between the RGB and HIS can be expressed as

$$\begin{cases} I = (R+G+B)/3 \\ S = I - \dfrac{[Min(R,G,B)]}{3} \\ W = cos^{-1}\left[\dfrac{2R-G-B}{2\sqrt{R^2+G^2+B^2-RG-GB-RB}}\right] \end{cases}$$

$$H = W, \text{ if } B \leqslant G \qquad (1)$$
$$H = 2^C - W, \text{ if } B > G$$

The study detected the green to orange color of citrus peel during the maturity process, and the color change was used to determine whether the citrus was mature or not.

3. Fruit (navel orange) Identification Processing

In the process of navel orange picking, the robot captures images in real time through the camera, and the field of view includes the sky, branches, leaves, fruits, and the earth, etc.

Moreover, there are overlapping, blocking, and uneven lighting, and the background is very complex, which puts the robustness of the recognition algorithm to a severe challenge.

To solve such problems, Guangli Chu et al. designed a machine vision-based method for sphere-like fruit recognition. The method first uses an image segmentation algorithm based on color normalization difference. It separates the fruit from the background [3]. Then uses a single-edge pixel tracking algorithm to extract the edges of the fruit, and uses a straight-line projection method to eliminate corner points. This step is done to improve the accuracy. Finally, the circles are detected by the least squares method to identify the fruit.

The algorithm can be divided into image Denoising processing, color segmentation, occlusion problem processing, and fruit recognition these four steps.

3.1. Image denoising processing

For easier post-processing, the image is first processed for noise reduction. The algorithm uses a Gaussian filtering algorithm to reduce the impact of image noise. Gaussian filtering is shown in the Fig.1.

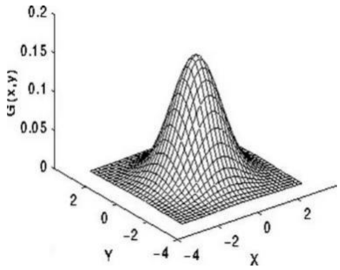


Fig.1. Gaussian filtering

3.2. Color segmentation

Considering that the mature navel orange is orange or red warm color, the fruit background is mostly green, cyan and other cool colors. The fruit and background colors are clearly distinguished, so the color segmentation can achieve good results. There uses normalized chromatic aberration for image segmentation. The relationship between the S and R、G、B can be expressed as

$$S = \frac{255(2R+B)}{2(R+G+B)} \qquad (2)$$

When $S \geqslant T$ (the fixed threshold), the pixel value of the point remains unchanged, and when $S < T$, the point becomes black. Processing comparison images is shown in Fig.2, and image(a) is unprocessed image, image(b) is processed image.

(a) unprocessed image

(b) processed image

Fig.2. Processing comparison images

3.3. Occlusion problem processing

When the navel orange fruit is on the tree, there will be and other fruits or leaves block each other. The mutually blocked fruits or leaves must have cross corner points, here use the straight line projection method to eliminate the corner points, so that the edge of the two fruits at the cross point break. his allows for better identification of the navel orange.

3.4. Fruit recognition

Based on the fact that navel oranges are mostly round in shape, this study uses least squares to achieve circle detection. After initial processing to eliminate edges that are clearly not round, further processing determines round fruits. Circle detection result is shown in the Fig.3.

Fig.3. Circle detection result

Finally, the data is transmitted to the actuator for picking. The study used a machine vision algorithm to recognize 500 similar images and it achieved a processing time of 83.548ms and an accuracy of 95%.

4. Fruit (cherry) Defect Detection

Wang Zhao from Beijing Forestry University used machine vision technology to extract image information from the cherry surface. He proposed the corresponding image processing algorithm to achieve defect detection and recognition on the cherry surface.

Achieving cherry surface defect recognition is divided into two steps: cherry image background removal processing and cherry defect feature extraction.

4.1. Cherry image background removal processing

The first choice is to extract its R component to enhance the contrast between the cherry image and the background image. To facilitate the subsequent processing, it is converted into a Binary Image, Morphological processing is applied to enhance the image, after which the edge extraction algorithm is applied to extract the edges to get the cherry and defective contour edges.

Subsequently, the cherry contour is filled to get the complete cherry contour, the contour is filtered to remove the non-conforming contour.finally the original color image is filled.

4.2. Cherry defect feature extraction

According to the cherry defects mainly have four kinds of rotten, bird peck, scratches, deformed fruit. Among them, rotten in the image acquired by machine vision is manifested as a large number of spots in the binary map; bird peck and scratches are manifested as black irregular lines inside the edge of the image and other defect characteristics.

First of all, the obtained de-background map is weakened by the brightness of the reflective region, which is converted into a binary image. By eliminating the external contours, the defective binary image can be obtained.

5. Conclusion

This paper describes the application of machine vision for recognition and detection on fruits such as citrus, navel orange and cherry, including some operations of machine vision in this aspect, such as: image enhancement, image edge detection, etc. Understanding the specific operation steps and role of machine vision in fruit recognition in recent years.

With the refinement of neural networks, fuzzy control and other intelligent algorithms and the development of artificial intelligence, the richness and speed of image processing technology is greatly enhanced, coupled with the advent of the 5G era, will also lay the technical foundation for the further integration of machine vision technology and fruit picking and agricultural production.

References

1. Hongkun Tian, Tianhai Wang,Yadong Liu, et al., Computer vision technology in agricultural automation — A review. Information Processing in Agriculture, 2020, 7(1), pp. 1-19.
2. Ying Yibin, Rao Xiuqin, Ma Junfu. Methodology for nondestructive inspection of citrus maturity with machine vision (in Chinese). Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2004, 20(2), pp. 144-147.
3. Chu Guangli, Zhang Wei, Wang Yanjie ,et al., A method of fruit picking robot target identification based on machine vision. Journal of Chinese Agricultural Mechanization, 2018, 39(2), pp. 83- 88.

Authors Introduction

Ms. Tianyi Zhang

She is a second-year master candidate in Tianjin University of Science and Technology, majoring in brain-computer interface, 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.

A Portable Electrocardiograph System Design based on STM32 Chip

Tianyi Zhang*, Huating Liu

College of Electronic Information and Automation, Tianjin University of Science and Technology, 300222，China

*E-mail: * zhangtianyi085@163.com*

www.tust.edu.cn

Abstract

ECG monitoring instruments occupy an important position in the medical field of application.It has a great reference value for the testing of basic cardiac functions and its pathological studies. For the shortcomings of conventional ECG monitoring instruments and equipment, which are large and not easy to carry, this paper designs a portable ECG monitor. In this paper, we designed a portable infrared induction heart rate measurement device, using STM32C8T6 as the core chip, to complete the extraction and processing analysis of ECG signals. The complete set thus composed can make basic diagnosis of the collected data and is successfully applied to daily life.

Keywords: signal acquisition, digital signal process, embedded system, STM32

1. Introduction

With the progress of society, economic development and the gradual improvement of people's living standards, the aging of the population is becoming more and more serious. The incidence of heart disease is also on the rise. [1]. However, the prevention, treatment and control rates are still low due to limitations in detection methods. Prevention rates are a key factor in the effective prevention and treatment of cardiovascular disease. Effective and convenient ECG monitoring instruments are a powerful tool to accomplish this task.

Conventional ECG monitoring devices are bulky, expensive and not portable, and can only record ECG activity while the patient is lying down for a short period of time, obtaining very little information. In order to enable more convenient diagnosis in more settings, a great variety of portable ECG devices have emerged.

Portable monitoring device can monitor and store data anywhere and anytime. This paper mainly designs a portable ECG device, which collects ECG signals through sensors, performs processing operations such as filtering and enhancement by STM32 chip, and finally is displayed by the display. It has the features of low price, small size, easy to carry and easy to use.

2. General Design

By combining embedded technology, digital signal processing technology and signal acquisition technology, we design an embedded ECG monitoring system that can perform signal extraction and analysis functions. [5-8]

The whole system consists of six modules: signal acquisition module, signal process module, keypad module, display circuit module, PC communication module and MCU module. The system structure schematic diagram is shown in Fig.1.

The ECG signal is acquired by electrodes and subsequently sent to the ECG acquisition circuit. After processing by preamplification, main amplification, high and low pass filtering, the system gets the ECG signal that meets the requirements and sends it to the STM32 ADC for AD conversion.

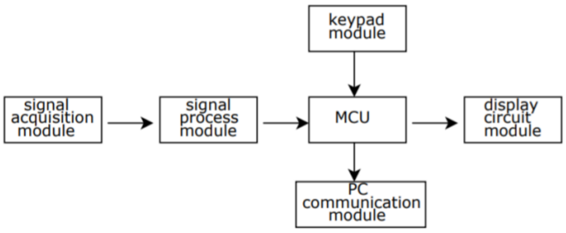


Fig.1. The system structure schematic diagram

The system control chip adopts STM32 and the display adopts TFT-LCD. Its touch function plus a small number of keys can establish a good human-computer interaction environment, which can be displayed and played back in real time by LCD. The data can be reliably transferred to a PC via USB for further analysis.

3. System Hardware Design

The hardware selected for the system includes the following two aspects: choice of processor and the human-machine interaction interface.

3.1. Choice of processor

The selection of the processor should consider five issues such as processing speed, complexity of the completed task, complexity of peripheral circuits, production cost and high power consumption.

Considering the above aspects, we finally chose STM32F103C8T6, a new 32-bit ARM core processor chip from STMicroelectronics, from the STM32 family. The main control chip is shown in Fig.2.

Fig.2. STM32C8T6 chip

In the design, the processor is responsible for signal acquisition, signal filtering and processing, display of ECG waveforms, data storage, and communication.

3.2. Human-machine interaction interface

Human-machine interface is a device for communication between human and machine, which can transmit human commands to MCU and also let the device display the information we want to know. This part of the display interface design choose color LCD screen, power consumption and size of the choice to combine the actual use.

According to the above points this design chose 3.2 inch true color TFT LCD touch screen, with 320 * 240 pixels, 260,000 colors, 16-bit parallel interface, it can be directly driven by AVR, ARM7, STM32 and other MCUs. The screen-related parameters are as follows:

- Resolution: QVGA 240 x 320
- Size: 3.2 inches
- Controller: IL9320
- Touch screen: 4-wire resistive type
- Pins: 30PIN
- Backlight: 4 LEDs in parallel

The touch screen is shown in Fig.3.

Fig.3 The touch screen

4. System Circuit Design

In the circuit design, this design adopts Pre-amplifier circuit, right leg driver circuit, filter circuit, trap circuit and power supply circuit. These modules greatly improve the function of the device.

4.1. Pre-amplifier circuit

The pre-amplifier circuit has to perform the function of differential signal amplification. this part of the circuit is crucial in the whole acquisition circuit, so it is necessary to choose a suitable differential op-amp chip. The following points are generally considered in the selection.

- Gain

- Frequency Response
- Common Mode Rejection Ratio
- Input Impedance
- Low noise, low drift

4.2. Right leg driver circuit

The right leg drive circuit is typically used in biosignal amplifiers to reduce common mode interference. Due to the large amount of external interference to humans, the capacitive coupling between the ECG electrodes and the power lines generates displacement currents[1].

We use a right-leg drive circuit, which can have a suppressive effect on 50Hz interference and does not come at the expense of losing the frequency component of the ECG signal.

4.3. Filter circuit and trap circuit

To filter out interference, a band-pass filter is designed so that ECG signals with frequencies from 0.05 Hz to l00 Hz pass through. signals outside this range will be substantially attenuated.

The active bandpass filter is used in this design. The bandpass filter is composed of a high and low pass filter with a high pass filter as of f=0.03Hz and a low pass filter as of f=110Hz.

Although the design of the right leg drive circuit, but there is still 50Hz interference into the circuit, so this design adds a 50Hz trap circuit. By this method to filter out the industrial frequency interference, the experimental results show that the signal waveform is clear and distinctive by high and low pass filtering followed by trap circuit.

4.4. Power circuit

The power supply circuit design mainly considers which type of power supply device to use, the input and output voltages, the output current, and the control state [15]. The ECG acquisition circuit requires a ±5V power supply, the STM32 operates at 3.3V and this design is powered by a 7.2V battery.

5. System Software Design

Software Flow DiagramThe system software is divided into two broad parts.

(1) Lower computer software, i.e. STM32 application.It mainly completes the acquisition of ECG signals, signal filtering, RTC module, LCD display and serial communication, etc.; (2) the upper computer management software. Its function is to complete the reception and processing of data, which mainly includes the reception, display and storage of data [2].

Software Flow Diagram is shown in Fig.4.

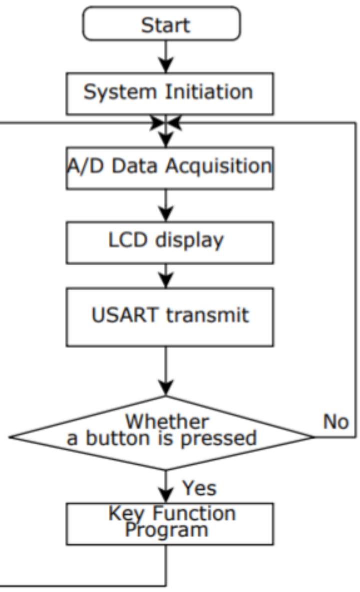


Fig.4. Software Flow Diagram

6. Conclusion

The system is a portable and miniaturized ECG monitoring system, using the cost-effective Cortex-M3-based STM32 chip as the microprocessor. Thus, we designed a portable ECG signal collector based on STM32 chip with real-time monitoring and low cost.

It is easy to carry the characteristics of storing data, can complete a long period of ECG monitoring, and real-time ECG waveform display through TFT-LCD color LCD, through the key to make the system has a good human-computer interaction interface. According to the debugging results, the system has completed the expected vision better and can collect signals from the human body and process and display them more correctly.

References

1. Zhou J . Elimination of Power-line Interference from ECG Signals. *Journal of Biomedical Engineering Research*, 2003.
2. Wang D X , Wang G J . Development of a portable ambulatory ECG monitor based on embedded microprocessor unit. *Space Medicine & Medical Engineering*, 2005, 18(3):196-200.

Authors Introduction

Ms. Tianyi Zhang

She is a second-year master candidate in Tianjin University of Science and Technology, majoring in brain-computer interface, machine vision.

Ms. Huating Liu

She is a second-year master candidate in Tianjin University of Science and Technology, majoring in flushbonading

Design and Development of The Parking Space Autonomous Management System

Yiting Gao*, Tianyi Zhang
Tianjin University of Science and Technology, Tianjin, China

*E-mail: *2801110578@qq.com*
www.tust.edu.cn

Abstract

The article purpose is to change the manual or semi-manual management model that parking Spaces now have, so as to make the parking space management more standardized and more effective. This system used the Browser/Server architecture, JSP(Java Server Page) as the front r & d tool, SQL Server as the background data warehouse for r & d. The system key functions realized in the last system include: user registration and login, message block, parking reservation, change of login password,fee settlement block and parking information query. The design manufacture of this parking space autonomous management system can increase the effect of parking space management to a large extent, all reflecting the information period of the special good.

Keywords:autonomous, JSP,B\S,tomcat

1. Introduction

With the rapid development of automobile industry, China's automobile ownership has increased sharply [1]. As a part of transportation facilities, with the busy and continuous development of transportation, people's requirements for its management are also increasing. They all hope that the management can achieve the effect of convenience, quickness and safety.

The scale of parking lots is different, and the management mode is also different. Managers need to choose and apply economic and stable management procedures according to their own conditions, so as not to choose a high-cost management system. This paper aims to design a simple, stable and practical parking lot management information system, hoping to have its own characteristics in fault tolerance, practicability and easy operation, and maintain certain scalability to meet the information management needs of different parking lots.

The whole system adopts the two-layer mode of separating the business logic layer and the user presentation layer. This developed mode can separate the database operation class from the user layer, which is convenient for code modification and system maintenance in the future. The parking management system can make the parking lot management information and gain an advantage in the increasingly competitive vehicle industry

2. The Hardware Structure and Software

In the production of parking lot reservation management system, a fully equipped computer is very useful for. The required configurations are: server:Tomcat,development software:MyEclipse,database: MySQL, R & D language: JSP, Java language.At the same time, enough memory can ensure that the code runs smoothly. In addition to the structure of the web page, in a whole parking lot system, users need to hold a card to park, so as to run the whole system. There are also access cabinets to display the current status.The design of the access cabinet is shown in Fig.1.

Fig.1 The design of the access cabinet

2.1. Server

The server we selected is Tomcat. Tomcat is a free open source web application server. It is a lightweight application server [3]. It is widely used in small and medium-sized systems and when there are not many concurrent access users. It is the first choice for developing and debugging JSP programs. Tomcat is shown in the Fig.2.

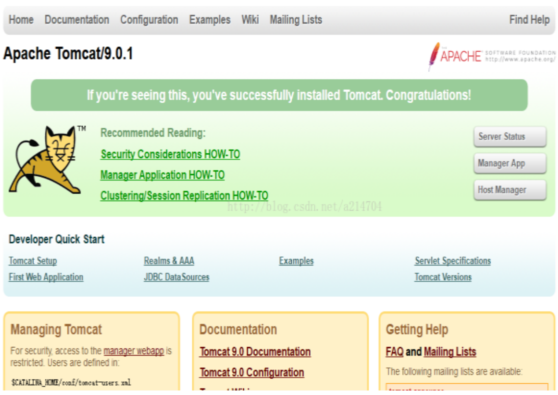


Fig.2. Tomcat server

2.2. Database

The function of database is to manage all kinds of data orderly and provide unified interfaces and services for other applications. Three database uses of MySQL:

(1) Multifunctional database: this option provides fast access to both transactional storage engine (InnoDB) and non transactional storage engine (MyISAM);

(2) Transactional database only, this option mainly optimizes the transactional storage engine (InnoDB), but the non transactional storage engine (MyISAM) can also be used;

(3) Non transactional database only. This option mainly optimizes the non transactional storage engine (MyISAM). Note that the transactional storage engine (InnoDB) cannot be used.

3. Software Design

The parking lot reservation system is designed and developed by individuals, and the development time is about one month. The hardware facilities needed to develop the software are a single computer. First, it is developed on the computer. After the development is completed, it is tested on the server. The development tool used is Eclipse and the database is Mysql.

3.1. Database design

The design of the database is the cornerstone of the project. According to the analysis of the key functions of the system, the tables of the parking space autonomous management system include administrator information table, parking space information table and parking information table.

3.2. Module detailed functional design

According to the functions of the parking reservation system, the system can be mainly divided into two modules: the front and rear platforms, which are the interaction of database information after user login and the operation of database information after administrator login. Function diagram of parking reservation system is shown in the Fig.3.

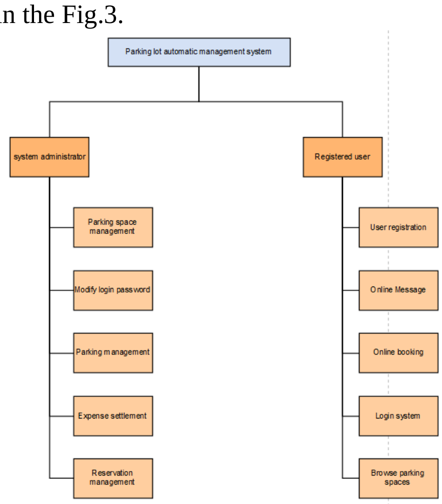


Fig.3. Function diagram of parking reservation system

3.3. Authorized design

Authorized design is shown in Table 2.

Table 2.Authorized design

Authorization number	User name	Object name	Jurisdiction
001	Administrators	Administrators	add,delete,check and change
002	User	User	add,delete,check and change

4. Software Test

The content of software testing [3,4] includes software code testing and software system function testing. The specific content refers to simulating and inputting some real values in the program, then running it, and checking some errors on the way.

4.1. General check of source code

The inspection is mainly a functional spot check of some key modules to check the modules. For example, whether the names of variables and functions in the code meet the specification requirements of the development software; You can also check whether the notes in the program are standardized, whether the amount of notes meets the requirements and whether the notes are identified accurately; Check whether the data display meets the standards, etc.

(1) Naming convention check

Check whether the naming in the source code, such as variables, functions, objects, etc., comply with the agreed specifications.

(2) Note check

Check whether the notes of the program meet the specifications, whether the amount of notes meets the specified requirements, and whether the instructions of notes are correct. For example, the amount of notes is required to be more than 20%.

(3) Interface inspection

Check whether the database interface and external interface are named properly, whether they are correctly connected with functions, and whether they clarify the functions to be completed.

(4) Data type check

Some of the codes involved have definitions of numbers. For example, whether the floating point type is used to define the price, the definition of date, the definition of number, etc. are in line with the specification.

5. Conclusion

Parking lot management is a troublesome and monotonous work, which repeatedly manages the vehicles in and out of the warehouse every day, and there is great instability in the type and entry time of vehicles, as well as great randomness in the parking spaces used by vehicles, which brings great inconvenience to the staff. So we developed this system, hoping to achieve more convenient and accurate operation under the use of this system.

This system used the Browser/Server architecture, JSP(Java Server Page) as the front r & d tool, SQL Server as the background data warehouse for r & d. The system key functions realized in the last system include: user registration and login, message block, parking reservation, change of login password, parking information management, parking information management, fee settlement block and parking information query. The design and manufacture of this parking space autonomous management system can increase the effect of parking space management to a large extent, all reflecting the information period of the special good.

References

1. Beibei Ding, Xiaona Yang, Investigation and Analysis on the development status of intelligent transportation in China, *Automobile Practical Technology*, 2021, 46: pp.199-201.
2. Yuankun Du, Yuxin Huang, Tomcat6.0 Configuration and applicationof connection pool, *Computer CD software and Application*, 2015, 18(02): pp.114-115.
3. Qile Qi, Research on Key Technologies of embedded computer software testing, *Electronic production*, 2021, 21:pp. 99-100+78.
4. Bo Weng,Yanfeng Yan, Research on software testing in big data environment, *Internet Weekly*, 2021, 22:pp. 42-44.

Authors Introduction

Ms.Yiting Gao 	She is a first-year master candidate in Tianjin University of Science and Technology, majoring in neural network, deep learning.

Ms.Tianyi Zhang 	She is a second-year master candidate in Tianjin University of Science and Technology, majoring in brain-computer interface, machine learning.

An Overview of Obstacle Avoidance Methods for Unmanned Vehicles

Peng Lu*, **Haokang Wen**

Tianjin University of Science and Technology, China

*E-mail: * lupeng970504@163.com*

www.tust.edu.cn

Abstract

Autonomous obstacle avoidance is one of the popular research elements in the field of intelligent unmanned vehicles, and it is a key technology to realize the automatic travel of intelligent unmanned vehicles. This paper introduces the traditional algorithms and intelligent algorithms related to autonomous obstacle avoidance, and analyze the advantages and drawbacks of the corresponding algorithms. The development of the corresponding control strategies is also summarized and summarized in order to provide some reference for the research on obstacle avoidance of intelligent unmanned vehicles.

Keywords: avoiding obstacles, unmanned vehicle, control strategy, algorithm

1. Introduction

Obstacle avoidance path planning for intelligent unmanned vehicles is an important and challenging task that plays a crucial role in the operation of unmanned vehicles [1]. The obstacle avoidance system of the unmanned vehicle uses the sensors installed in the vehicle and the sensory information provided by the outside world to sense the speed, acceleration, position, status and other information about itself and the outside environment during the vehicle's travel. Then, by calculating and analyzing relevant information, we make decisions for safe vehicle driving and control the actuator to realize automatic vehicle driving.

Unmanned vehicles will encounter static or dynamic obstacles in the course of travel, at which time relevant algorithms and autonomous obstacle avoidance strategies need to be introduced to achieve real-time, accurate autonomous obstacle avoidance and ensure the safe operation of the vehicle.

At present, the algorithm research in the field of autonomous obstacle avoidance has achieved certain results. In this paper, the existing algorithms are classified into traditional algorithms and intelligent algorithms according to the applicable environment, and they are elaborated and summarized respectively.

2. Traditional Obstacle Avoidance Algorithms

The more classical methods of traditional obstacle avoidance include visual graph method, grid method, artificial potential field method, virtual force field method, etc.

2.1. Visual graph method

The visual graph method, proposed by Lozano-Perez and Wesley in 1979, is the classical algorithm for global motion planning of unmanned vehicles. In the visual graph method, the unmanned vehicle is described by points and the obstacles are described by polygons, and the starting point, the target point and the vertices of the polygon obstacles are combined and connected. It is required that the line between the starting point and each vertex of the obstacle, between the target point and each vertex of the obstacle, and between each vertex of the obstacle and the vertex cannot cross the obstacle, i.e.

the line is "visible". The vehicle then uses some optimization algorithm to search for the optimal path from the starting point to the target, and then the shortest path from the starting point to the target is obtained by accumulating and comparing the distances of these lines.

The visual map method can find the shortest path, but the search time is long and inflexible, and it is troublesome to reconstruct the visual map once the starting point and target point of the unmanned vehicle are changed. The visual map method is applicable to polygonal obstacles, but not to circular obstacles. The tangent diagram method and the Voronoi diagram method improve the visual diagram method. The tangent map method uses the tangent of the obstacle to represent the arc, so the map of the shortest path from the starting point to the target point, and the unmanned vehicle must almost walk close to the obstacle. The disadvantage is that the possibility of unmanned vehicles colliding with obstacles can be high if position errors arise during control. The Voronoi diagram method represents the arc by a path that is as far away from obstacles and walls as possible. As a result, the path from the starting point to the target point will grow, but with this type of control, the unmanned vehicle will not encounter obstacles even if position errors are generated.

2.2. Grid method

The grid method was proposed by W.E. Howden in 1968. It is a physical model in the shape of a grid to represent the likelihood of an obstacle's appearance. The grid method uses the basic element as the minimum grid granularity to rasterize the map, and the basic element in the free zone takes the value of 0, in the obstacle zone or contains the obstacle zone is 1, so that a map can be used for path planning in the computer. When each grid is given a passage factor, the path planning problem becomes a problem of finding the optimal path between two grid nodes on the grid network. In a grid map, the size of the grid affects the amount of environmental information stored and the length of time. The smaller the grid division, the greater the storage of environmental information, the higher the resolution and the better the obstacle avoidance in complex environments, but the longer the computation time.

2.3. Artificial potential field

The artificial potential field was originally proposed by Khatib in 1985 and has since been widely used in path planning for mobile robots such as unmanned vehicles. Its basic idea is to consider the motion of an unmanned vehicle in its surroundings as the motion of an unmanned vehicle in an artificially created virtual force field. In the virtual force field, the target point generates a gravitational force that guides the vehicle towards the target point and the obstacle generates a repulsive force that prevents the vehicle from colliding with the obstacle. The motion of the vehicle is controlled according to the combined forces of gravity and repulsion to produce a collision-free optimal path.

The driving paths planned by applying an artificial potential field are generally smooth and safe, with simple algorithms and good real-time performance. But the algorithm also has some drawbacks. For example, when there is an obstacle near the target point, the repulsive force is much greater than the gravitational force and the vehicle will have difficulty reaching its destination. When the gravitational and repulsive forces at a point are exactly equal in magnitude, the combined force on the unmanned vehicle is zero, i.e. it is caught in a local minimum problem. A great deal of research has been carried out by scholars in various countries to optimize these issues.

The virtual force field is a real-time obstacle avoidance algorithm for unmanned vehicles that combines the grid method with an artificial potential field. The VFF algorithm is a local obstacle avoidance algorithm that uses a grid to represent the environment and a force field method to control the unmanned vehicle.

3. Intelligent Obstacle Avoidance Algorithm

Intelligent obstacle avoidance algorithms are generally stochastic search algorithms based on biological intelligence or physical phenomena, commonly known as fuzzy logic algorithm, genetic algorithm, neural network methods and ant colony algorithm [1].

3.1. Fuzzy logic method

Instead of simplifying the actual situation and thus building a mathematical model as in classical control theory, fuzzy control uses human experience and decision making to reason with the corresponding fuzzy logic and describes the entire time-varying control process in a language with fuzziness. The fuzzy logic obstacle avoidance algorithm is an intelligent obstacle avoidance algorithm based on real time sensing information and reference to human driving experience, which achieves local path planning by building a fuzzy control rule base to obtain planning information [2]. The method develops corresponding fuzzy control rules based on information such as the angle between the direction of motion of the dynamic obstacle and the unmanned vehicle, the time of collision and the position of the collision point on the obstacle.

For fuzzy control of unmanned vehicle obstacle avoidance, the key issue is to establish a suitable fuzzy controller. The fuzzy controller is mainly responsible for the fuzzification of sensor sensing information, the operation of fuzzy relations, fuzzy decision making and the non-fuzzy processing of the obstacle avoidance decision results, and thus intelligently controls the obstacle avoidance behavior of the unmanned vehicle. The fuzzy controller is shown in Fig.1.

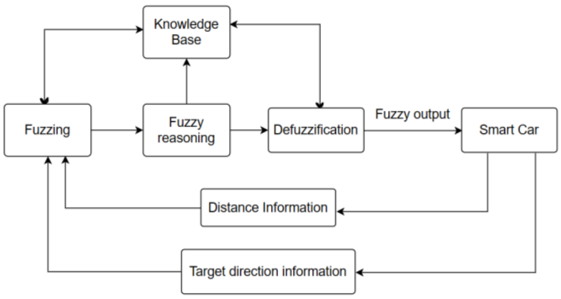


Fig.1 Fuzzy controller for obstacle avoidance

There are 3 main points in establishing a suitable fuzzy controller.

(1) Substitution of linguistic variables for mathematical variables.

(2) Describe the relationships between variables using fuzzy control condition statements.

(3) Describing complex relationships with fuzzy algorithms.

Fuzzy logic obstacle avoidance algorithms have the disadvantage of not being able to learn on their own and being inflexible. This is because fuzzy rules are pre-defined by experience and do not change with the input of unknown environmental information, and the number of fuzzy rules increases exponentially with the number of inputs [3].

3.2. Genetic algorithms

Genetic algorithms are a method for searching for optimal solutions [4]. It is an intelligent algorithm that simulates the evolutionary principles of biology and achieves species evolution through evolution and genetic variation. Genetic algorithm-based obstacle avoidance methods are an effective intelligent algorithm in the field of autonomous obstacle avoidance research for unmanned vehicles [5].

The genetic algorithm treats all path points as a population and uses binary coding to encode each path point, then selects the path points according to the fitness function and performs combinatorial crossover and mutation with the help of genetic operators, gradually evolving to produce increasingly optimized approximate solutions. At the same time, global path planning is combined with local path planning in the planning process and corresponding obstacle avoidance strategies are proposed depending on the type of collision between the unmanned vehicle and dynamic obstacles. The algorithm can effectively guide unmanned vehicles to achieve obstacle avoidance in dynamic environments and obtain collision-free optimal or sub-optimal paths. The genetic algorithm based obstacle avoidance process is shown in Fig.2.

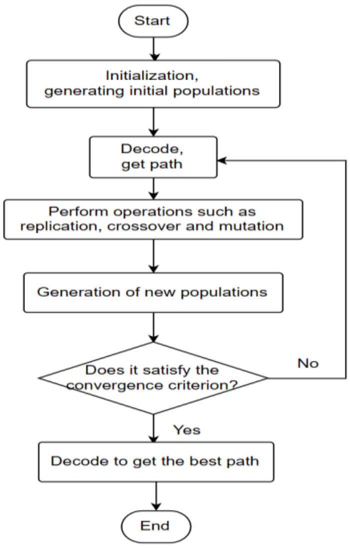


Fig.2 Genetic algorithm based obstacle avoidance process

3.3. Neural network

A neural network is a mathematical or computational model that mimics the structure and function of a biological neural network [6]. Neural networks consist of a large number of artificial neurons linked together for computation. In most cases artificial neural networks can change their internal structure based on external information and are an adaptive system. Artificial neural networks are typically optimized by a mathematical statistics-based type of learning method, a nonlinear statistical data modeling tool that can model complex relationships between inputs and outputs.

Neural network obstacle avoidance path planning methods tend to build a neural network model about the travel path of an unmanned vehicle from an initial position to a target position. The model input is the sensor information and the motion direction of the previous position or the previous position of the unmanned vehicle, and after the model is trained, the motion direction of the next position or the next position of the unmanned vehicle is output. The structure of the neural network obstacle avoidance system is shown in Fig.3.

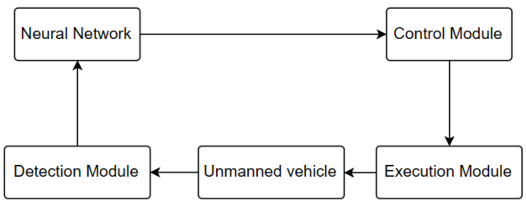


Fig.3 Neural network obstacle avoidance structure

There is now another emergence of building dynamic neural network-based obstacle avoidance algorithms for unmanned vehicles. The dynamic neural network can automatically adjust its structure according to the complexity of the state of the unmanned vehicle environment and realize the mapping relationship between the state of the unmanned vehicle and its obstacle avoidance action in real time, which can effectively reduce the computing pressure of the unmanned vehicle information processing system.

3.4. Ant colony algorithm

The ant colony algorithm is an intelligent optimization algorithm proposed by Dorigo et al. Bionomists have discovered that ants transmit information between individuals through a substance called pheromone, which the ants can sense to guide their direction. The basic principle of the ant colony algorithm is shown in Fig.4.

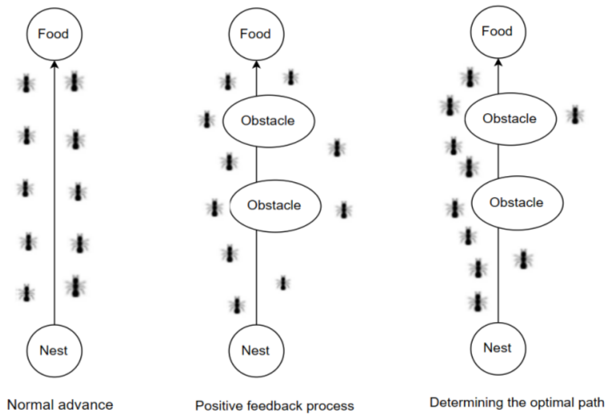


Fig.4 Principle diagram of ant colony algorithm

When an ant finds food, it will release a volatile secretion, or pheromone, into the environment to attract other ants to come, so that more and more ants will pass this path. As more ants pass on the path, the more pheromones are left behind and the more likely it is that new ants will choose this path, which is a positive feedback process. Some ants do not repeat the same path as others, they will find another way, and if the other path is shorter than the original one, then, gradually, more ants are attracted to this shorter path. Finally, after a period of running, the colony can always find the shortest path between the food source and the nest. The ant colony algorithm is to simulate the foraging behavior of ant colony through continuous iteration to complete the process of finding the shortest and optimal path.

The ant colony algorithm has the advantages of positive information feedback mechanism, strong robustness and easy parallel implementation. However, in the initial stage if the pheromone is missing or the path size is too large, it will lead to too slow path planning. Therefore, when using the ant colony algorithm for unmanned vehicle obstacle avoidance path planning, the grid method is usually used first to model the environment of obstacles, and then the global obstacle avoidance path planning is performed using the ant colony algorithm.

4. Conclusion

The article presents a comprehensive analysis and description of the main traditional and intelligent algorithms in the field of unmanned vehicle obstacle avoidance. Both traditional and intelligent algorithms can solve the problem of unmanned vehicle obstacle avoidance to some extent, and each algorithm has its advantages and limitations.

However, with the development of driverless technology, the complexity of its application scenarios is increasing. Therefore, in practical situations, to perform accurate, safe and fast obstacle avoidance relying only on a single algorithm is limited.

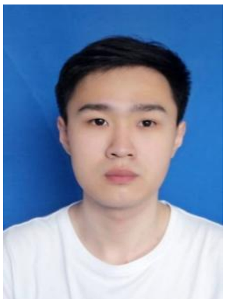
Therefore, how to overcome the limitations of existing algorithms, effectively integrate multiple algorithms, build on their strengths and avoid their weaknesses, and achieve path planning with completely unknown obstacle information is the focus of future research in the field of unmanned vehicle obstacle avoidance.

References

1. Nazemizadeh M, Trajectory planning of robots in presence of obstacles: A review study, *Universal Journal of Computational Analysis*, 2013, 1: pp.24-32.
2. C. Perez-D'Arpino, W. Medina-Melendez, J. Guzman, et al., Fuzzy logic based speed planning for autonomous navigation under Velocity Field Control, *2009 IEEE International Conference on Mechatronics*, 2009, pp.1-6.
3. Yao P, Wang H, Su Z, Real-time path planning of unmanned aerial vehicle for target tracking and obstacle avoidance in complex dynamic environment, *Aerospace Science & Technology*, 2015, 47: pp.269-279.
4. Lee J, Kim D W, An effective initialization method for genetic algorithm-based robot path planning using a directed acyclic graph, *Information Sciences*, 2016, 332: pp.1-18.
5. Ahmed F, Deb K, Multi-objective optimal path planning using elitist non-dominated sorting genetic algorithms, *Soft Computing*, 2013, 17(7): pp.1283-1299.
6. Parhi D R, Singh M K, Real-time navigational control of mobile robots using an artificial neural network, *ARCHIVE Proceedings of the Institution of Mechanical Engineers Part C Journal of Mechanical Engineering Science 1989-1996 (vols 203-210)*, 2009, 223(7): pp.1713-1725.

Authors Introduction

Mr. Peng Lu



He is a second year master student at Tianjin University of Science and Technology. His research is about multi-sensor fusion, artificial intelligence, and robotics.

Mr. Haokang Wen

He is the third-year graduate student of Tianjin University of Science andTechnology. 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.

Synchronization of Novel 5D Hyperchaotic Systems

Hong Niu
College of Electronic Information and Automation,
Tianjin University of Science & Technology, China
E-mail: spots@163.com
www.tust.edu.cn

Abstract

In this paper, synchronization of novel five-dimensional (5D) autonomous hyperchaotic systems is studied. The synchronization control law is proposed based on the center translation method. A structure compensator is formulated to make the mathematical model of the error system the same as that of the response system, and a linear feedback controller is designed via the Lyapunov stability theory to make the error system globally asymptotically stable at the origin. Thus, the two 5D hyperchaotic systems are synchronized. Some relevant numerical simulation results, such as the curves of the corresponding synchronization state variables and the errors, are given to illustrate the feasibility and effectiveness of the synchronization control law.

Keywords: novel 5D hyperchaotic system, hyperchaos synchronization, center translation method, Lyapunov stability theory, global asymptotic stability

1. Introduction

Hyperchaos was first presented in 1979 by Otto Rössler.[1] Because hyperchaos is much more complicated than chaos, hyperchaos synchronization has greater application significance and engineering value in secure communication.

In this paper, the mathematical model of the novel 5D hyperchaotic system is given as the drive system. Hyperchaos synchronization of the 5D systems is studied based on the center translation method. Corresponding numerical simulation results are presented to demonstrate the validity of the synchronization method.

2. The Novel 5D Hyperchaotic System

The dynamic equations of the novel 5D hyperchaotic system are

$$\begin{aligned} \dot{x}_1 &= a\left(y_1 - x_1\right), \\ \dot{y}_1 &= \left(c-a\right)x_1 + cy_1 + w_1 - x_1 z_1, \\ \dot{z}_1 &= -bz_1 + x_1 y_1, \\ \dot{v}_1 &= mw_1, \\ \dot{w}_1 &= -y_1 - hv_1, \end{aligned} \qquad (1)$$

where $x_1, y_1, z_1, v_1, w_1 \in R$ are state variables, and a = 23, b = 3, c = 18, m = 12 and h = 4.[2]

Let the initial values of the system (1) be (x_{10}, y_{10}, z_{10}, v_{10}, w_{10}) = (1, 1, 1, 1, 1), then the Lyapunov exponents respectively are $\lambda_{11} = 0.8732 > 0$, $\lambda_{12} = 0.1282 > 0$, $\lambda_{13} = -0.0013 \approx 0$, $\lambda_{14} = -0.5770 < 0$ and $\lambda_{15} = -8.4231 < 0$. It indicates that the system (1) is hyperchaotic. The attractors of the 5D hyperchaotic system (1) are shown in Fig. 1.

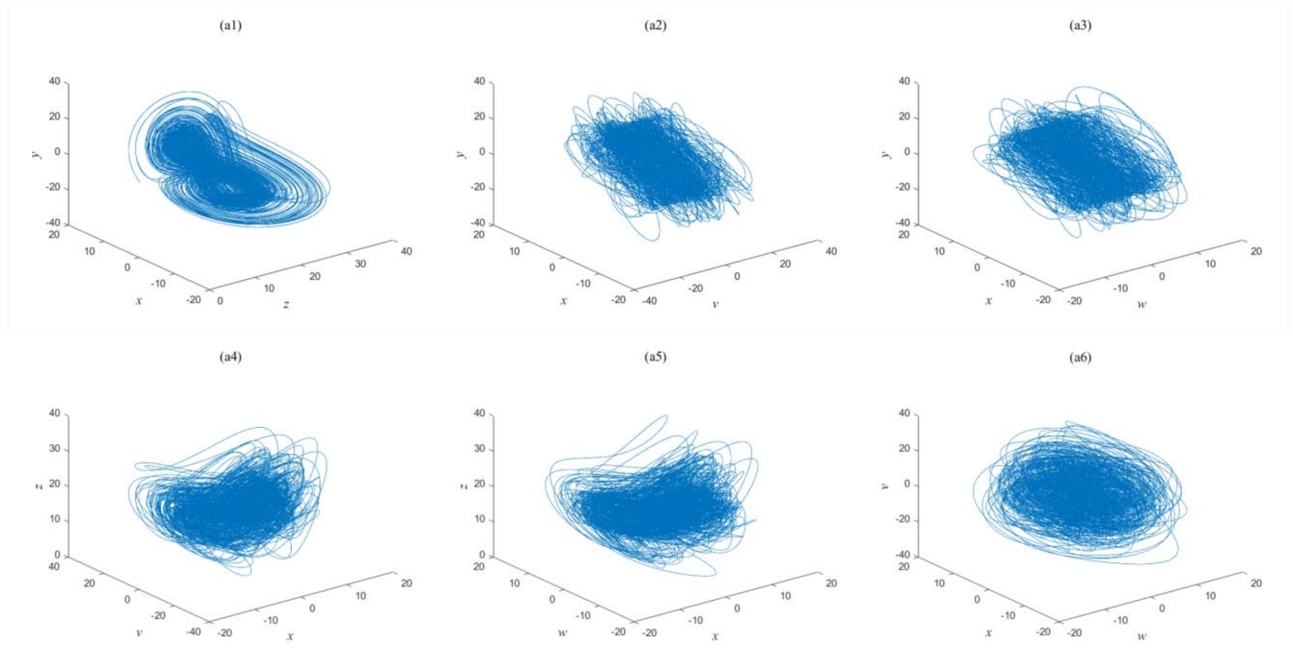


Fig. 1. Attractors of the 5D hyperchaotic system: (a1) *z-x-y*; (a2) *v-x-y*; (a3) *w-x-y*; (a4) *x-v-z*; (a5) *x-w-z*; (a6) *w-x-v*

3. Hyperchaos Synchronization Based on Center Translation Method

3.1. *Formulation of error system*

Take the system (1) as the drive system, then the response system is formulated as

$$
\begin{aligned}
\dot{x}_2 &= a\left(y_2 - x_2\right) + u_{s1} + u_{c1},\\
\dot{y}_2 &= \left(c-a\right)x_2 + cy_2 + w_2 - x_2 z_2 + u_{s2} + u_{c2},\\
\dot{z}_2 &= -bz_2 + x_2 y_2 + u_{s3} + u_{c3},\\
\dot{v}_2 &= mw_2 + u_{s4} + u_{c4},\\
\dot{w}_2 &= -y_2 - hv_2 + u_{s5} + u_{c5},
\end{aligned} \tag{2}
$$

where

$$\boldsymbol{u}_s = \left[u_{s1} \quad u_{s2} \quad u_{s3} \quad u_{s4} \quad u_{s5}\right]^{\mathrm{T}}$$

and

$$\boldsymbol{u}_c = \left[u_{c1} \quad u_{c2} \quad u_{c3} \quad u_{c4} \quad u_{c5}\right]^{\mathrm{T}}$$

are structure compensator and synchronization controller to be designed. Let $\boldsymbol{u}_s = \mathbf{0}$, $\boldsymbol{u}_c = \mathbf{0}$, and the initial values of the response system (2) be (x_{20}, y_{20}, z_{20}, v_{20}, w_{20}) = (5, 0, 4, 3, 8), then the Lyapunov exponents respectively are $\lambda_{21} = 0.9121 > 0$, $\lambda_{22} = 0.1175 > 0$, $\lambda_{23} = -0.0008 \approx 0$, $\lambda_{24} = -0.5533 < 0$ and $\lambda_{25} = -8.4755 < 0$. It shows that the response system (2) is also hyperchaotic.

Let

$$
\begin{aligned}
\boldsymbol{e} &= \left[e_1 \quad e_2 \quad e_3 \quad e_4 \quad e_5\right]^{\mathrm{T}}\\
&= \left[x_2 - x_1 \quad y_2 - y_1 \quad z_2 - z_1 \quad v_2 - v_1 \quad w_2 - w_1\right]^{\mathrm{T}}
\end{aligned}
$$

be the synchronization error and

$$
\boldsymbol{u}_s = \left[u_{s1} \quad u_{s2} \quad u_{s3} \quad u_{s4} \quad u_{s5}\right]^{\mathrm{T}}
= \begin{bmatrix} 0 \\ x_2 z_1 + x_1 z_2 - 2x_1 z_1 \\ -x_2 y_1 - x_1 y_2 + 2x_1 y_1 \\ 0 \\ 0 \end{bmatrix},
$$

then the error system is simplified as

$$
\begin{aligned}
\dot{e}_1 &= a\left(e_2 - e_1\right) + u_{c1},\\
\dot{e}_2 &= \left(c-a\right)e_1 + ce_2 + e_5 - e_1 e_3 + u_{c2},\\
\dot{e}_3 &= -be_3 + e_1 e_2 + u_{c3},\\
\dot{e}_4 &= me_5 + u_{c4},\\
\dot{e}_5 &= -e_2 - he_4 + u_{c5}.
\end{aligned} \tag{3}
$$

Comparing the mathematical model of the error system (3) with that of the controlled system (2) in Ref.

2, it can be found that the two models are similar. Hence, the synchronization controller $\boldsymbol{u}_c$ is designed as

$$\begin{aligned}\boldsymbol{u}_c &= \begin{bmatrix} u_{c1} & u_{c2} & u_{c3} & u_{c4} & u_{c5} \end{bmatrix}^{\mathrm{T}} \\ &= \begin{bmatrix} -k_1e_1 & -k_2e_2 & -k_3e_3 & -k_4e_4 & -k_5e_5 \end{bmatrix}^{\mathrm{T}},\end{aligned}$$

where $k_1, k_2, k_3, k_4, k_5 \geq 0$.

3.2. *Design of linear feedback synchronization controller*

Theorem 1. *Let* $\boldsymbol{x} = \boldsymbol{0}$ *be an equilibrium point for* $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x})$, *where* $\boldsymbol{f}: D \to R^n$ *is a locally Lipschitz map from a domain* $D \subset R^n$ *into* R^n. *Let* $V: R^n \to R$ *be a continuously differentiable function such that*

$$V(\boldsymbol{0}) = 0 \text{ and } V(\boldsymbol{x}) > 0, \quad \forall\, \boldsymbol{x} \neq \boldsymbol{0}$$

$$\|\boldsymbol{x}\| \to \infty \Rightarrow V(\boldsymbol{x}) \to \infty$$

$$\dot{V}(\boldsymbol{x}) < 0, \quad \forall \boldsymbol{x} \neq \boldsymbol{0}$$

then $\boldsymbol{x} = \boldsymbol{0}$ *is globally asymptotically stable.*[2]

Take a continuously differentiable function

$$V = \frac{1}{2}\left(e_1^2 + e_2^2 + e_3^2 + \frac{h}{m} e_4^2 + e_5^2 \right)$$

as a Lyapunov function candidate for the error system (3). Then, the derivative $\dot{V}$ is derived as

$$\begin{aligned}\dot{V} &= e_1\dot{e}_1 + e_2\dot{e}_2 + e_3\dot{e}_3 + \frac{h}{m} e_4\dot{e}_4 + e_5\dot{e}_5 \\ &= -(k_1 + a)e_1^2 + ce_1e_2 - (k_2 - c)e_2^2 \\ &\quad -(k_3 + b)e_3^2 - k_4\frac{h}{m}e_4^2 - k_5e_5^2 \\ &\leq -\left(k_1 + a - \frac{c}{2}\right)e_1^2 - \left(k_2 - \frac{3}{2}c\right)e_2^2 \\ &\quad -(k_3 + b)e_3^2 - k_4\frac{h}{m}e_4^2 - k_5e_5^2.\end{aligned}$$

For $\dot{V} < 0$, the parameters k_1, k_2, k_3, k_4 and k_5 should satisfy that

$$\begin{matrix} k_1 + a - \dfrac{c}{2} > 0, \\ k_2 - \dfrac{3}{2}c > 0, \\ k_3 + b > 0, \\ k_4\dfrac{h}{m} > 0, \\ k_5 > 0, \end{matrix} \Rightarrow \begin{matrix} k_1 > \dfrac{c}{2} - a, \\ k_2 > \dfrac{3}{2}c, \\ k_3 > -b, \\ k_4 > 0, \\ k_5 > 0, \end{matrix} \Rightarrow \begin{matrix} k_1 = 0, \\ k_2 = 30, \\ k_3 = 0, \\ k_4 = 1, \\ k_5 = 1. \end{matrix}$$

Thus, the linear feedback synchronization controller $\boldsymbol{u}_c$ is designed as

$$\begin{aligned}\boldsymbol{u}_c &= \begin{bmatrix} u_{c1} & u_{c2} & u_{c3} & u_{c4} & u_{c5} \end{bmatrix}^{\mathrm{T}} \\ &= \begin{bmatrix} 0 & -30e_2 & 0 & -e_4 & -e_5 \end{bmatrix}^{\mathrm{T}}.\end{aligned}$$

From Theorem 1, the error system (3) is globally asymptotically stable at the origin. It indicates that the response system (2) is synchronized with the drive system (1).

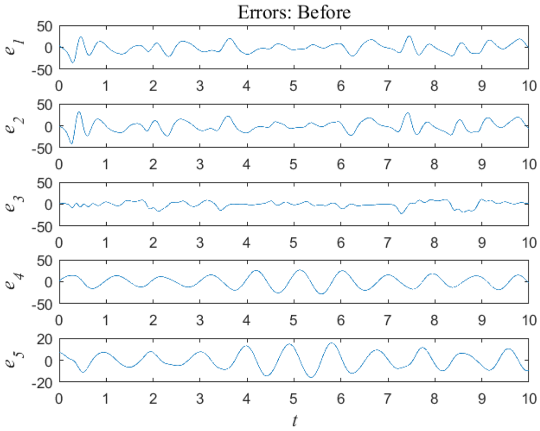
3.3. *Numerical simulation*

Remark 1. *The initial values of the drive system (1) and the response system (2) are* (x_{10}, y_{10}, z_{10}, v_{10}, w_{10}) = (1, 1, 1, 1, 1) *and* (x_{20}, y_{20}, z_{20}, v_{20}, w_{20}) = (5, 0, 4, 3, 8) *respectively in this paper.*

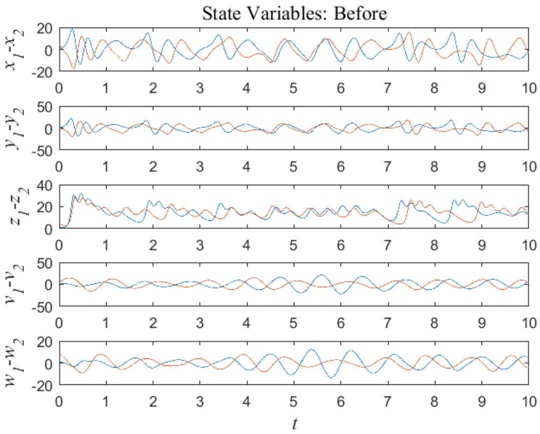
Definition 1. *After adding the structure compensator* $\boldsymbol{u}_s$ *and the linear feedback synchronization controller* $\boldsymbol{u}_c$ *to the response system (2), the Lyapunov exponents of the response system (2) are called sub-Lyapunov exponents.*[3]

Theorem 2. *The response system (2) and the drive system (1) will synchronize only if the sub-Lyapunov exponents are all negative.*[3]

The curves of the errors and the corresponding state variables before and after adding the structure compensator $\boldsymbol{u}_s$ and the linear feedback synchronization controller $\boldsymbol{u}_c$ to the response system (2) are shown in Fig. 2 and Fig. 3 respectively. Comparing Fig. 3 with Fig. 2, it can be found that the errors e_1, e_2, e_3, e_4 and e_5 converge to zero asymptotically and rapidly and the corresponding state variables are synchronized well after adding $\boldsymbol{u}_s$ and $\boldsymbol{u}_c$ to the response system (2). Moreover, the sub-Lyapunov exponents of the response system (2) are λ_{21c} = -1.0292, λ_{22c} = -1.0355, λ_{23c} = -3.0000, λ_{24c} = -17.4669 and λ_{25c} = -17.4690, which are all negative. From Theorem 2, the response system (2) and the drive system (1) have synchronized.

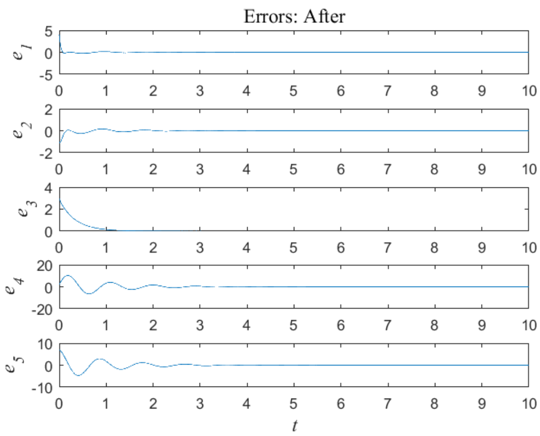


(a) Errors

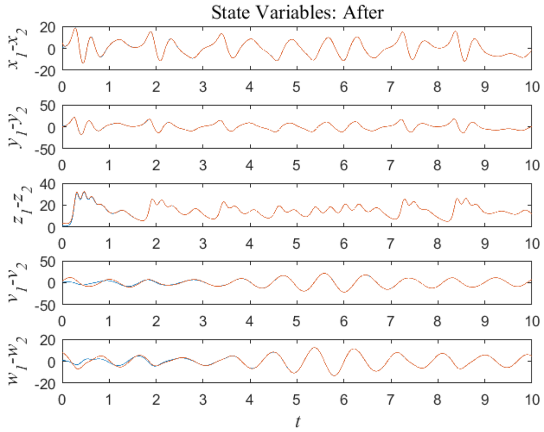


(b) State variables

Fig. 2. Before: (a) Errors; (b) State variables



(a) Errors



(b) State variables

Fig. 3. After: (a) Errors; (b) State variables

4. Conclusions

Synchronization of the novel 5D hyperchaotic systems is proposed based on the center translation method in this paper. Numerical simulation results illustrate the feasibility of the synchronization method. The study has some engineering significance. Furthermore, the circuit implementation of the synchronization system is under investigation and will be reported elsewhere.

References

1. O. E. Rössler, An equation for hyperchaos, *Physics Letters A*, 1979, 71(2–3): pp. 155–157.
2. Q. Wei, H. Niu, Control of a novel 5D hyperchaotic system, *Proceedings of the 2021 International Conference on Artificial Life and Robotics*, On line, January, 2021, pp. 655-658.
3. L. M. Pecora, T. L. Carroll, Synchronization in chaotic systems, *Physical Review Letters*, 1990, 64(8): pp. 821-824.

Author Introduction

Dr. Hong Niu
She received the Ph.D. degree in control science and engineering from Tianjin University, China, in 2014. She is currently a lecturer with the College of Electronic Information and Automation, Tianjin University of Science & Technology, China.

Data expansion method by combining unnecessary sentence deletion and most important sentence addition

Tomohito Ouchi, Masayoshi Tabuse
Graduate School of Life and Environmental Sciences, Kyoto Prefectural University,
1-5 Shimogamohangi-cho, Sakyo-ku, Kyoto 606-8522, Japan
E-mail:t_ouchi@mei.kpu.ac.jp, tabuse@kpu.ac.jp

Abstract

We are studying data expansion methods in automatic summarization systems. In our research, the method of expanding the input article with unnecessary sentences deleted is the most effective of the extended methods. In the previous research, we have tried a method of adding most important sentences. In this research, we propose a method that combines the deletion of unnecessary sentences and the addition of most important sentences. We propose a hybrid method with two methods, one is to add important sentences first and the other is to add important sentences last.

Keywords: automatic summarization, data augmentation, Pointer-Generator Model, Extractive summarization.

1. Introduction

Currently, the amount of information on the Internet is expected to increase at an average annual rate of 29% from 2010 to 2024, reaching 143 ZB by 2024 [1]. In terms of text data, the number of websites worldwide in 2018 was about 1.6 billion, however, it was about 1.8 billion in 2021 [2]. Since it has increased by about 200 million in three years, it is expected that text data will increase steadily in the future. Under such circumstances, the issue of selecting information is an urgent problem. Automatic summarization struggles that issue. However, it can be said that extractive summarization that only made up with sentences is not sufficient. Since the sentence-to-sentence connection is not taken into consideration, readability is lacking. Therefore, it is needed generative summarization as a technology that looks ahead. A generative summarization basically uses the Encoder-Decoder model, which learns the relationship between input and output and generates one word at a time in the output when a new input comes in during the test. Various models have been proposed [3,4]. In this study, the Pointer-Generator model [3] uses as the baseline model. One of the issues with the generative summarization model is that data maintenance is costly. We have to attach a manual summary to each article in order to make the generative summarization model. Therefore, we focused on data augmentation as a method that can be applied to any model. This is to create extended data from existing data. As a result, it was confirmed that the accuracy of the evaluation metric ROUGE [5] of Pointer-Generator model applied by the data augmentation method is improved by about 1% compared to baseline model.

Next, we explain the method of data augmentation simply. We decide the importance of each sentence in each article. To decide the importance of sentences, we used the topic model. In the existing research [6], the sentence with the lowest importance is removed to obtain extended data. Furthermore, a previous study [7] examined the effect of expansion by adding the most important sentences. This method is further divided into two methods depending on the position of addition. The method of adding to the beginning of existing data is called “add-s”, and the method of adding to the end of existing data is called “add-e”. Also, the method by removing the lowest important sentence is called “remove”. This method was proposed method in [7].

And in this research, we propose a new method. It's a combination of “remove” and “add”. It is called a “hybrid”. The method that combines “remove” and “add-e” is called “hybrid-e”. The method that combines “remove” and “add-s” is called “hybrid-s”. These five techniques (“remove”, “add-e”, “add-s”, “hybrid-e”, and “hybrid-s”) are described in Section 2. Experiments and results are described in Section 3. And discussions are given in Section 4.

2. Data Augmentation Method

This section describes the five models used in the data augmentation method. In each method, each sentence is scored in an article, and create the extended data using this score.

2.1. Topic Model

The topic model is used in the existing method and the new method. For how to determine the importance of sentences using the topic model, we referred to existing research [8]. The topic model is one of the language models that assumes that one document consists of multiple topics. In addition, each topic has an appearance word distribution. The method of determining the importance of a sentence is as follows.

1. Calculate the frequency of occurrence in a topic with words that make up a sentence
2. Sum of all the words that make up the sentence
3. Divide by the square root of the sentence length
4. Sum on all topics

2.2. Proposed method

The five methods ("remove", "add-e", "add-s", "hybrid-e", "hybrid-s") use the topic model. First of all, we calculate the score of sentences importance in input one article using topic model. In the “remove” method, the lowest important sentence is removed to existing data. In the “add-e” method, the highest important sentence is added to end of existing data. In the “add-s” method, the highest important sentence is added to beginning of existing data. In the "hybrid" method, important sentences are added and then unnecessary sentences are deleted. Figure 1 shows the remove method processing procedure.

Proposed Method
1. make topic model
2. $r \leftarrow$ None
3. $i \leftarrow 0$
4. For $s \in S$
   $i_s \leftarrow$calculate score(s, topic model)
   If $i < i_s$
      $r \leftarrow s$
      $i \leftarrow i_s$
5. $E \leftarrow S$ remove r

出力 E

Figure 1 Proposed method(“remove”) processing procedure

3. Experiment and Results

3.1. Parameter Setting

The CNN / DailyMail dataset is used as the dataset for training, evaluating, and testing. The training data, evaluating data, and test data are 287,226 articles, 13,768 articles, and 11,490 articles, respectively. The model used for the experiment is the Pointer-Generator model, which is composed of a copy mechanism and a coverage mechanism when learning. In the Copy mechanism we calculate the error of the evaluating data each time the epoch ends and we use the model of the epoch with the lowest error in Early Stopping. Early Stopping what we mean here, uses a model that waits twice as many epochs as the error seems to be the minimum, unless the minimum value is updated. Next, in the coverage mechanism, the same processing is performed in the coverage loss. We use ROUGE as using for evaluation on existing research.

The program used in this research uses PyTorch. It has been confirmed that this program can achieve the same result as [3]. The hidden layer vector size was set to 256 and the embedded vector size was set to 128. The batch size was set to 8. In the original paper, the batch size is 16, so double learning is required to learn the same number of articles. The beam size was set to 4. The beam search will be described later. The number of vocabulary was set to 50,000. The learning rate was set to 0.15.

In this program, the number of words used to encode an input article is limited to 400. This setting has no effect on learning an extended data. Specifically, an extended data is the same as an original data. This is because the extracted sentence may not be within 400 words from the beginning. We must confirm that the extracted sentence is present in the input article. Therefore, I found the article with the most number of words among the articles used in the training data. The number of words with the most words was 2,380. And the upper limit of the number of words used in encoding the input article was set to 2,380. Table 1 shows the values of ROUGE when the maximum number of words is 400 and 2,380. In the Table 1, f, r, and p represent the F value, recall, and precision, respectively.

Table1 the values of ROUGE when the maximum number of words is 400 and 2,380

	ROUGE-1-f	ROUGE-1-r	ROUGE-1-p	ROUGE-2-f	ROUGE-2-r	ROUGE-2-p
400	0.3935	0.4372	0.3800	0.1709	0.1891	0.1662
2380	0.3958	0.4181	0.3994	0.1741	0.1832	0.1770

	ROUGE-L-f	ROUGE-L-r	ROUGE-L-p
400	0.3616	0.4014	0.3493
2380	0.3644	0.3846	0.3679

Table 1 shows when the upper limit of the number of words is increased from 400 to 2,380, the value of ROUGE increases slightly. In the following, the experiment is performed with the upper limit of the number of words set to 2,380.

In this experiment, we calculated the average ROUGE value of three experiments in order to eliminate the randomness of the parameters as much as possible.

3.2. Beam search

Greedy method contrasts with beam search. Specifically, in greedy method, when generating a word, one word with the highest generation probability is selected, while in beam search, processing is performed while holding the top K words. Then, we make the final summarizations by multiplying the probabilities of each word generation, and make the highest one the final summarization. In this experiment, this K value is set to 4. The following Table 2 summarizes the parameter settings.

Table 2 Parameter settings

hidden vector size	256
embbed vector size	128
batch size	8
beam size	4
vocabulary size	50,000
lerning rate	0.15
input word size	2,380

3.3. Experimental Results

The results are shown in Table 3, 4 and 5.

Table 3 Results of learning 115,000 articles using 6 methods

	normal	remove	add-e	add-s	hybrid-e	hybrid-s
ROUGE-1	0.3441	0.3519	0.3470	0.3436	0.3429	0.3462
ROUGE-2	0.1389	0.1426	0.1400	0.1382	0.1374	0.1394
ROUGE-L	0.2968	0.3025	0.2977	0.2963	0.2945	0.2970

Table 4 Results of learning 57,000 articles using 6 methods

	normal	remove	add-e	add-s	hybrid-e	hybrid-s
ROUGE-1	0.3137	0.3302	0.3231	0.3210	0.3256	0.3302
ROUGE-2	0.1145	0.1272	0.1239	0.1183	0.1264	0.1269
ROUGE-L	0.2676	0.2828	0.2773	0.2745	0.2806	0.2813

Table 5 Results of learning 28,000 articles using 6 methods

	normal	remove	add-e	add-s	hybrid-e	hybrid-s
ROUGE-1	0.3056	0.3188	0.3043	0.3265	0.3187	0.3052
ROUGE-2	0.1079	0.1201	0.1048	0.1252	0.1212	0.1050
ROUGE-L	0.2595	0.2711	0.2601	0.2790	0.2730	0.2573

Table 3 shows the results of the 6 methods (normal, remove, add-e, add-s, hybrid-e, hybrid-s) when 115,000 articles were trained, Table 4 when 57,000 articles were trained, and Table 5 when 28,000 articles were trained. “normal” method is baseline model.

3.4. Disccusion

Among all the number of articles, “remove” showed the best effect of expansion. In training data 115,000 articles, “remove” showed the best effect. Next, “add-e” had good effect. In training data 57,000 articles, “hybrid-s” and “remove” showed the best effect. In training data 28,000 articles, “add-s” showed the best effect. In training data 28,000 articles, “add-s” showed the best effect. Next, “remove” and “hybrid-e” had almost the same good effect. “add-e”, “add-s”, “hybrid-e”, and “hybrid-s” were not found to be superior or inferior in this experiment. For “remove”, too few or too many articles seemed to reduce the effect. However, the number of articles in the middle seemed to be the most effective. In order to verify this, the following additional experiment was conducted.

3.5. Additional Experiment

The purpose of this experiment is to see at which number of articles the effect is most visible. For this purpose, we tried experiments with training data of 10,000, 28,000, 45,000, 57,000, 90,000, 180,000, and 287,226 articles. The methods used were “normal” and “remove”. The results are shown in Table 6 and 7.

Table 6 Results of each the number of articles by "normal" and "remove"

	10,000	28,000	45,000	57,000	90,000	180,000	287,226
normal	0.3226	0.3358	0.3398	0.3340	0.3538	0.3758	0.3841
remove	0.3335	0.3473	0.3506	0.3591	0.3627	0.3827	0.3916
difference	0.0109	0.0115	0.0108	0.0251	0.0089	0.0069	0.0075

Thus, with 57,000 articles at the top, the expansion effect of remove is getting worse for both fewer and more articles than that. And we can also see that it gets better for fewer articles than for more. It is expected that the reason why it got worse with more articles is that it is not possible to generate a more varied input article vector. The reason why it was worse for fewer articles is that the input article vector was biased. Since the expansion is only a slight modification of the original article, if the vector is too biased, it will remain biased even after the expansion.

4. Conclusion

In this study, we proposed "hybrid-e" and "hybrid-s" in addition to the existing work [7]. Both methods showed the effect of expansion, but not better than "remove"." The reason why "add-e", "add-s", "hybrid-e", and "hybrid-s" had a certain effect was because the extended data was readable and understandable to humans. In the existing study [6], the EDA method was not so effective because the readability of the extended data was reduced.

In this study, we also examined at what number of articles the effect of expansion is most apparent. As a result of the verification, it was found that the effect of the expansion was diminished for both too few and too many articles.

As a future task, we would like to show the high applicability of the proposed method by verifying whether the effect of the expansion can be seen in a better model that is currently being devised [9], although in this study we only tested the Pointer-Generator model.

References

1. Worldwide Global DataSphere Forecast, 2020–2024: The COVID-19 Data Bump and the Future of Data Growth IDC ANALYZE THE FUTURE. https://www.idc.com/getdoc.jsp?containerId=US44797920
2. Total number of Websites Internet Live Stats. https://www.internetlivestats.com/total-number-of-websites/
3. Abigail. See, Peter J. Liu, et al., "Get To The Point: Summarization with Pointer-Generator Networks", arXiv:1704.04368 (2017)
4. Yang Liu, Mirella Lapata, "Text Summarization with Pretrained Encoders", arXiv:1908.08345v2 (2019)
5. Chin-Yew Lin, "ROUGE: A Package for Automatic Evaluation of Summaries", Workshop on Text Summarization Branches Out, Post-Conference Workshop of ACL 2004, Barcelona, Spain (2004)
6. T. Ouchi, M. Tabuse, "Effectiveness of Data Augmentation in Pointer-Generator Model", Journal of Advances in Artificial Life Robotics, Vol. 1, No. 2, pp.95-99 (2020)
7. T. Ouchi, M. Tabuse, "Comparison of Data Augmentation Methods in Pointer–Generator Model", Journal of Robotics Networking and Artificial Life, Vol 8, No. 2, pp.85-89(2021)
8. H.Sigematsu, I.Kobayashi, "Generation of abstracts considering importance of potential topics", The Association for Natural Language Processing (2012) (In Japanese).
9. Liu, Y. and Lapata, M. "Text Summarization with Pretrained Encoders." In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP) pp. 3730-3740 (2019)

Authors introduction

Mr. Tomohito Ouchi

He received his Master's degree from Department of Environmental Science, Graduate School of Life and Environmental Sciences, Kyoto Prefectural University, Japan in 2019. He is currently a Doctoral course student in Kyoto Prefectural University, Japan.

Dr. Masayoshi Tabuse

He received his M.S. and Ph.D. degrees from Kobe University in 1985 and 1988 respectively. From June 1992 to March 2003, he had worked in Miyazaki University. Since April 2003, he has been in Kyoto Prefectural University. His current research interests are machine learning, computer vision and natural language processing. IPSJ and IEICE member.

Evaluation of a system that the reading of sentences by a voice synthesizer and the highlighting sentences

Sota Kobayashi, Masayoshi Tabuse
Graduate School of Life and Environmental Sciences, Kyoto Prefectural University,
1-5 Shimogamohangi-cho, Sakyo-ku, Kyoto 606-8522, Japan
E-mail: s-kobayashi@mei.kpu.ac.jp, tabuse@kpu.ac.jp
https://www.kpu.ac.jp/

Abstract

It is generally known that there are individual differences in human cognitive function. We made a hypothesis; giving both visual and audio stimuli may make it easier for people to catch information. In this study, we set three indicators consists of memory, understanding and concentration for an experiment. The difference in learning effect due to the reading situation was measured. We concluded voice synthesizer and highlighting system can help reading in some cases.

Keywords: psychology of leaning, cognitive science, verbatim memory, voice synthesizer

1. Introduction

It is generally known that there are individual differences in human cognitive function. For example, some people have an advantage over reading the text, while others have an advantage over hearing the words. We made a hypothesis; giving both visual and audio stimuli may make it easier for people to catch information. To test this hypothesis, we conducted an experiment.

2. Development Environment

Development environment is as follows.

- Language: C# (DotNet Framework 4.5)
- Integrated Development Environment: Microsoft Visual Studio Community 2019 (Version 16.8.2)
- Voice Synthesizer: CeVIO Creative Studio 7 (Version 7.0.23.1) [3]
- Morphological Analysis Engine: MeCab (MeCab.DotNet 0.0.26) [4]
- Heart Rate Monitor: Polar H10

3. Method

To test the hypothesis, we made a system that can highlight sentences and read by the voice synthesis. The process is as below.

1. Extract a target sentence based on the symbols like commas and periods that separate sentences.
2. The system gives the target sentence to CeVIO CS 7 to take pronunciation data. The system stores data into list structures whose members are mora and time-to-read. Mora is a unit in phonology. In Japanese, each character of kana corresponds to a mora. Time-to-read is the time to read a mora (msec).
3. The system gives the target sentence to MeCab to take Morphological data. The system stores data into new list structures named Manuscript whose members are morpheme, lexical category and pronunciation. Morpheme is the smallest unit of meaningful lexical item in language.
4. The system gives each morpheme to CeVIO CS 7 to check if the pronunciation is the same as that from MeCab. True or False will be followed by False join the next node other for the later processes.

5.The system stores reading-time into each structure of Manuscript. The reading-times are calculated by referring to moraDataList. If the structure of Manuscript has a True member, reading-time can be directly calculated. If False, the system finds out the next True and sum up the times before that.
After this process ends, each node of Manuscript has the following members: morpheme, lexical category, pronunciation, coordination-of-reading (boolean) and time-to-read (msec).
6.To make the Manuscript more comfortable to read, the software combines the structures. First, combining by referring to the lexical category; if the word order was as determined, combine the structures (e.g., a postpositional particle is next after a noun). Second, combining by referring to the reading time; if the reading time is less than 1000 msec, combine the structure with next to.
7.Executing highlight; give the Target Sentence to CeVIO CS 7 to let it start reading out and then start highlighting.

4. Experiment

To search for the effect of the differences in reading, we set 3 items of learning efficiency: memory, understanding and concentration. We referred to [1] on memory and understanding, and [2] on concentration.

4.1 Subjects

In this experiment, the subjects consist of 9 college students (male: A~F, female: G~I).

4.2 *Problem*

Experiment was conducted by using the problem set of [1]. Problem set consists of two elements: verbatim memory problems and understanding problems. The problems are excerpted from a novel, essay or fairy tale. They consist of around 270 characters.

In verbatim memory problems, the system shows a sentence to a subject and asks whether the same sentence was included in the problems or not. In the understanding problems the system shows a sentence to a subject and asks whether the sentences said the same meaning as the problem or not.

As [1] said so, the difficulties of the problem set have been adjusted to be the same level.

4.3 Experiment

We set 3 ways to read:

・read to oneself
(hereinafter called SIL)
・read assisted by voice synthesis
(hereinafter called OFF)
・read assisted by voice synthesis + highlight
(hereinafter called ON)

Subjects read the sentences with 3 ways and answered the questions; verbatim memory and understanding. Subjects had been equipped with a Heart Rate Monitor 5 minutes before start to solve the problem.

5. Results

Table 1 shows the result of verbatim memory and understanding. Each sentence has 4 verbatim memory and 4 understanding problems. One correct answer gives one point, so max 4 points. Table 2 shows both the average and standard deviation of verbatim memory and understanding scores.

Table 1. Score and average for each subject

Subject		A	B	C	D	E	F	G	H	I
memory	SIL	3	1	1	1	2	3	3	3	3
	OFF	4	2	2	2	2	3	2	2	3
	ON	2	3	2	1	3	3	2	2	1
understand	SIL	0	3	2	3	4	3	3	3	3
	OFF	2	1	1	2	2	2	4	2	4
	ON	3	1	2	2	2	4	3	3	3
average		2.33	1.83	1.67	1.83	2.50	3.00	2.83	2.50	2.83

Table 2. Ave and stdev of verbatim memory and understanding.

verbatim memory			understanding		
	ave	stdev		ave	stdev
SIL	2.22	0.916	SIL	2.67	1.05
OFF	2.44	0.685	OFF	2.22	1.03
ON	2.11	0.737	ON	2.56	0.831

As can be seen from Table 2, OFF got the highest average and lowest stdev scores of verbatim memory. It implies that OFF facilitates memorizing and prevents the influence of subjects' cognitive function. It means that OFF can be an effective way to memorize.

SIL got the highest average of understanding which implies that SIL facilitates understanding. ON got the lowest stdev of understanding which implies that ON prevents the influence of subjects' cognitive function. Incidentally, ON got a relatively high score. It means that ON can be an effective way to understand.

Table 3 shows the score difference between the way of reading for each subject.

Table 3. Score difference between the way of reading for each subject

subject		A	B	C	D	E	F	G	H	I
memory	SIL - OFF	-1	-1	-1	-1	0	0	1	1	0
	OFF - ON	2	-1	0	1	-1	0	0	0	2
	ON - SIL	-1	2	1	0	1	0	-1	-1	-2
understand	SIL - OFF	-2	2	1	1	2	1	-1	1	-1
	OFF - ON	-1	0	-1	0	0	-2	1	-1	1
	ON - SIL	3	-2	0	-1	-2	1	0	0	0

Six correlation coefficients of rows were calculated; both memory and understand have 3 combinations then 6 combinations exist.

Figure 1 and 2 show scatter plot of OFF-ON vs ON-SIL (memory) and SIL-OFF vs ON-SIL (understand), whose correlation coefficient absolute values are biggest in 6 correlations.

As can be seen from the Figure 1 and 2, each of them certainly has a correlation. Here, we consider what the correlations mean.

(1) Correlation coefficient r = -0.763 for *OFF-ON vs ON-SIL (memory)*

① Subject has high OFF-ON: disturbed by highlight

② Subject has low OFF-ON: enhanced by highlight

③ Subject has high ON-SIL: not so good at read to oneself but enhanced by highlight

④ Subject has low ON-SIL: good at read to oneself but disturbed by highlight

The negative correlation coefficient implies that ①,④ and ②,③ have high positive correlations.

For ①,④: who good at read to oneself could be disturbed by highlight.

For ②,③: who not good at read to oneself could be enhanced by highlight.

According to the above discuss, who good at memorize by rearing to oneself should not use highlight. On the other hand, highlight can be helpful to memorize for who not good at read to oneself.

(2) Correlation coefficient r = -0.774 for *SIL-OFF vs ON-SIL (understand)*

① Subject has high SIL-OFF: better at reading to oneself than reading by using reading system

② Subject has low SIL-OFF: better at reading by using reading system than reading to oneself

③ Subject has high ON-SIL: not so good at read to oneself but enhanced by highlight

④ Subject has low ON-SIL: good at read to oneself but disturbed by highlight

The negative correlation coefficient implies that ①,④ and ②,③ have high positive correlations.

For ①,④: who good at read to oneself could be not good at using reading system and be disturbed by highlight.

For ②,③: who not good at read to oneself could be good at using reading system and be enhanced by highlight.

According to the above discuss, who good at understand by reading to oneself should not use highlight or reading system. On the other hand, reading system and highlight could be useful for who not good at reading.

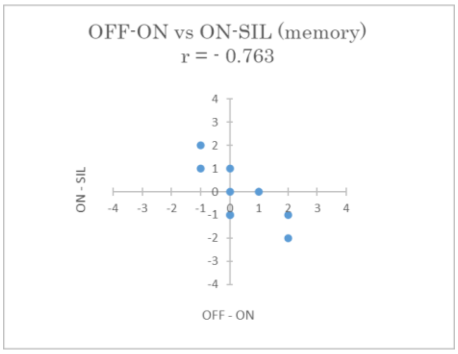


Figure 1. Scatter plot of OFF-ON vs ON-SIL (memory)

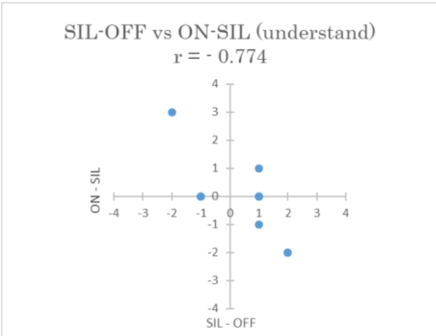


Figure 2. Scatter plot of SIL-OFF vs ON-SIL (understand)

Table 4 shows the personal average heart rate for each behavior. Rest 1 is the lowest one minute's average while being rest before solving the problem. Rest 2 is as same to that but after solved the problem. REST in the bottom row is the smaller value of rest 1 and rest 2, which is the representative value. Whose reason is that some subject can be nervous before or after experiment. While in nervous, the subjects are not in rest, so we took lower one as representative heart rate value of rest.

Table 4. Heart Rate (HR) for each subject (bpm)

subject	A	B	C	D	E	F	G	H	I
rest 1	86.8	97.9	80	101.9	71.3	62.3	87.4	78	71.9
SIL	86.5	106.1	81.1	103.5	77.2	70.1	99.3	81.4	72.3
OFF	90.3	97.9	82.5	102.3	79.6	69.3	97.1	83.5	72.4
ON	91.1	105.1	81.2	102.1	79.8	68.3	92.2	84.4	75.6
rest 2	87.6	97.6	81.7	97.3	78.2	67.2	94	80.5	70.1
REST	86.8	97.6	80	97.3	71.3	62.3	87.4	78	70.1

Table 5 shows the average and standard deviation of each subject. Table 6 is the average and standard deviation of difference between reading and rest. As can be seen from Table 6. ON-REST has highest average and lowest standard deviation. It implies that ON could be an effective way for who doesn't have much concentration power.

Table 5. Ave and stdev of heart rate (bpm)

heart rate	ave	stdev
rest 1	81.9	12.1
rest 2	83.8	10.5
REST	81.2	11.5
SIL	86.4	12.7
OFF	86.1	10.9
ON	86.6	11.4

Table 6. Ave and stdev of difference between reading and rest

heart rate	ave	stdev
SIL - REST	5.19	3.70
OFF - REST	4.89	2.90
ON - REST	5.43	1.98

6. Conclusion

We investigated the effects of the system that helps reading by a voice synthesis and highlighting. There are cases that memorizing, understanding or concentration are increased. This result implies that properly using voice synthesis or highlight may help learners.

7. Acknowledgement

I wish to acknowledge professor Yasunari Yoshitomi and specially-appointed professor Taro Asada for useful discussions.

8. References

1. T. Terao, M. Takahashi, S. Kiyokawa "Roles of articulatory movements and speech feedback in Japanese text comprehension during oral reading", The Japanese journal of psychology 89(6), 618-624, 2018. Shinrigaku-kenkyu. (In Japanese)
2. H. Takatsu, O. Osamu "Assessment of Concentration Using Heart Rate Variability during taking class in college students", Journal of National Institute of Technology, Toyota College 39(0), 149-152, 2006 (In Japanese)
3. CeVIO Project, https://cevio.jp/, Accessed 28 Jan. 2021
4. MeCab.DotNet, https://github.com/kekyo/MeCab.DotNet/blob/master/README.ja.md, Accessed 28 Jan. 2021

Authors Introduction

Mr. Sota Kobayashi

He received his B.S. degree from Informatics and Environmental Science, Kyoto Prefectural University, Japan in 2020. He is currently a master course student of Kyoto Prefectural University.

Dr. Masayoshi Tabuse

He received his M.S. and Ph.D. degrees from Kobe University in 1985 and 1988 respectively. From June 1992 to March 2003, he had worked in Miyazaki University. Since April 2003, he has been in Kyoto Prefectural University. He works as a Professor at the Graduate School of Life and Environmental Sciences of Kyoto Prefectural University. His current research interests are machine learning, computer vision and natural language processing. IPSJ and IEICE member.

Mouse Cursor Control System using Eye Gaze and Detection of Eye Opening and Closing

Masayoshi Tabuse[1], Mana Fukumoto[2], Yasunari Yoshitomi[1], Taro Asada[1]
1 Graduate School of Life and Environmental Science, Kyoto Prefectural University
1-5 Shimogamohangi-cho, Sakyo-ku, Kyoto 606-8522, Japan
E-mail: tabuse@kpu.ac.jp
2 Department of Informatics and Environment Science, Kyoto Prefectural University
1-5 Shimogamohangi-cho, Sakyo-ku, Kyoto 606-8522, Japan

Abstract

It is necessary to support of computer operation for a physically disabled person. One of the possible physical movements of the physically disabled person is Eye movement. Eye gaze and detection of eye opening and closing of a person makes it possible to operate a computer. In our system, a web camera, Gaze pointer, dlib C++ library and OpenCV library are used to recognize eye gaze and detect eye opening and closing. Changing the eye gaze, we can move a mouse cursor. Recognizing closed eye, we can carry out an operation of mouse click. In this paper, we evaluated the effect on operability due to the eye gaze and detection of eye opening and closing.

Keywords: Mouse Cursor Control system, Eye gaze, eye opening and closing.

1. Introduction

In recent years, a computer input system without the mouse and keyboard has been developed as a support system of computer operation for a physically disabled person However, many of them are contact type system that must be attached to the body which imposes a physical burden on the user. In addition, many non-contact type system are special sensors and expensive.

Therefore, in the previous study [1], we proposed a non-contact system which operates the mouse cursor using an inexpensive Web camera. The system recognizes the direction of the face and the opening and closing of the mouth and eyes so that the mouse pointer moves to the face direction and the system clicks the mouth in the case of the opening mouth and the closing eye. However in the previous study we assume that the user can turn his face and open and close his mouth and eyes so that the system cannot be used by severe amyotrophic lateral sclerosis (ALS) patients.

So, it is necessary to develop a mouse cursor control system that can be used by patients with severe ALS. In this research, we developed a mouse cursor control system only by eye movements, and evaluated its performance.

2. The communication means

In ALS, it is important to consider communication means as the symptoms progress. If a patient can move his arms and fingers, he can talk in writing, and if he can move his mouth, he can communicate using shape recognition of mouth. If he can move his eyes, he can communicate using a transparent letter board. However, communication means using shape recognition of mouth and a transparent letter board always require the help of a helper. On the other hand, if he can move his body or eyes, it is possible to communicate by selecting an appropriate computer input system. This study targets people with advanced symptoms who have residual eye movements.

3. Comparison with the previous study

In the previous study, we assumed that the user could turn his face and open and close his mouth and eyes. In this study, it is assumed that the user cannot change the direction of his face, cannot move his mouth but he can move his eyes. The movement of the mouse cursor depends on the direction of his eyes and click process is performed by gazing. In addition, we make it possible to change click to double-click by changing the input mode by closing his eye.

4. The proposed system

The movement of the mouse cursor depends on his eye gaze and click process is performed by gazing for 2 seconds or longer. To prevent unintended click processing, the click function will be stopped if both eyes are closed for 4 seconds or longer. The proposed system can be divided into three elements: "movement of mouse cursor", "mouse click processing", and "controlling mouse click function". Figure 1 shows a flowchart of the outline of the system.

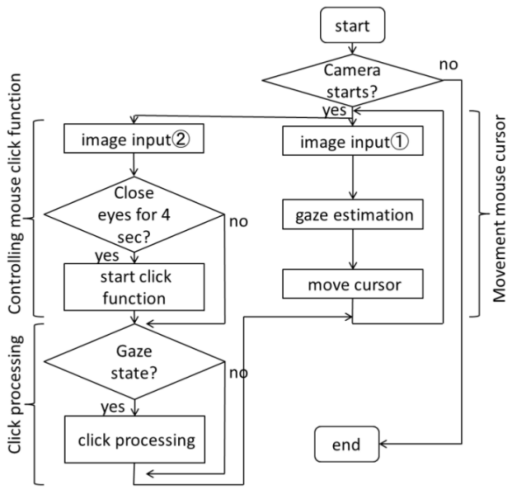


Fig 1. Flowchart of the outline of the system.

5. Movement of the mouse cursor

In this research, we use a Web camera and GazePointer software [3] to estimate the eye gaze and move the mouse cursor based on the eye gaze. Before using the system, the user calibrates eye gaze.

6. Mouse click

6.1. *Judgment of gaze state*

The system gets the coordinate M_0 of the mouse cursor at time t = 0 and sets this coordinate to the point G_0. If the coordinates M_1 of the mouse cursor got by the system at the next time t = 1 are within the rectangular area R_0 of 114 × 106 centered on G_0, the system sets R_1 with the midpoint of the coordinates M_1 and G_0 as the point G_1. When the coordinate Mt of the mouse cursor at time t is in the rectangular area R_{t-1} (Fig. 2) for about 2 seconds, it is judged as "gaze state" and the coordinates of G_t are determined to be the gaze coordinates. If M_t is not in the rectangular area R_{t-1}, initialize the time and continue the calculation (Fig. 3).

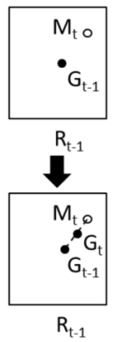


Fig.2 The coordinate M_t is in the region R_{t-1}

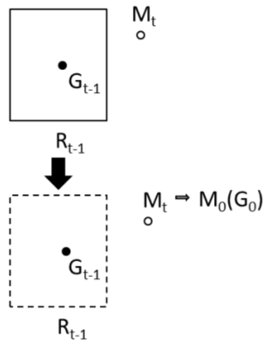


Fig.3 The coordinate M_t is out the region R_{t-1}

6.2. *Process of Mouse click*

If the state is the "gaze state", the coordinate of the point Gt is the coordinate of the gaze and the system Moves the mouse cursor to the point Gt to perform click processing.

6.3. *Judgment of eye opening and closing*

The judgment of opening and closing eyes is used by the method described in [1]. First, an open source C ++ library Dlib [4] is used to detect facial feature points (68 points in total).

The eye area is determined by the coordinates of the feature points of both eyes. Next, a luminance value histogram of the eye area is created, a threshold value is obtained by the discriminant analysis method [4], and binarization is performed. From the binarized image, the vertical width h_e and the horizontal width w_e of the eye are determined. The ratio R_e is obtained from he and we by equation, if R_e is above the threshold value, it is determined that the eyes are open, and if R_e is below the threshold value, it is determined that the eyes are closed. From the previous study [1], the threshold is set to 0.3.

$$R_e = \frac{h_e}{w_e}$$

6.4. *Controlling the mouse click*

When the mouse click function is activated and one eye is closed for about 1 second, the double click and single click functions are switched. Also, if both eyes are closed for about 4 seconds, the mouse click function will be stopped. On the contrary, when the mouse click function is stopped and both eyes are closed for about 4 seconds, the mouse click function is activated. Figure 4 shows a flowchart for controlling the mouse click function.

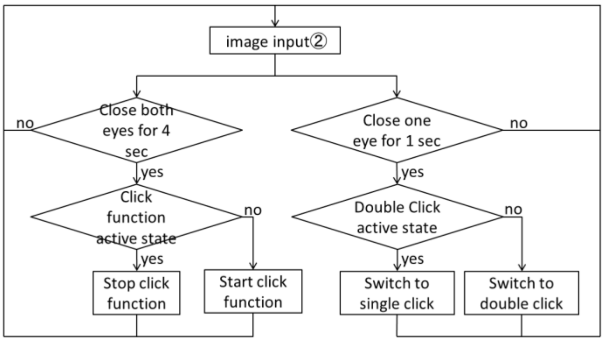


Fig.4 Flowchart of controlling the mouse click

7. Experiments

7.1. *Operation screen*

Figure 5 shows the operation screen when using the system. The user's face is displayed in the upper left and upper right of the screen. The upper left screen displays the eyes gaze detection, and the upper right displays the eye opening and closing judgment.

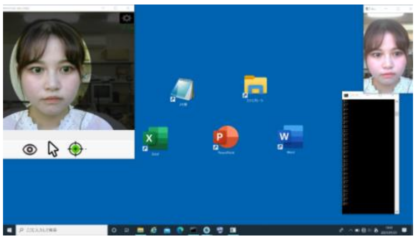


Fig.5 Operation screen

7.2. Experiment contents

We conducted the following two evaluation experiments comparing the proposed system with the previously reported system [1] for 10 subjects in their 20s (male: A to F, female: G to J).

・Evaluation experiment on mouse cursor movement and mouse click

・Evaluation experiment on character input

Before the experiment, the subjects explained the operation method of each system and the contents of the experiment, and then practiced for 2 minutes. When using the proposed system, calibration was performed for each experiment. After the experiment, a questionnaire was conducted.

7.3. Evaluation of mouse cursor movement and mouse click

The subjects use each system to double-click the icon (Notepad) installed on the desktop, and measured the required time (Experiment 1). Also, for the operation of double-clicking five icons installed on the desktop, the required time was measured for each system (Experiment 2), and the difference in the required time between single operation and continuous operation was investigated.

7.4. Evaluation of character input

Using the Windows screen keyboard (Fig. 6), we measured the time it took to enter "kyoto" in Notepad, and evaluated the difference in operability in character input of each system (Experiment 3).

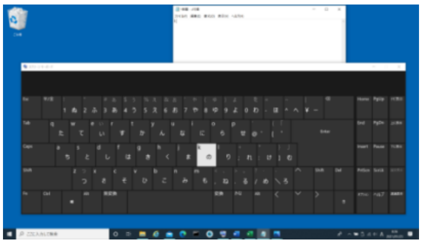


Fig.6 Character input

8. Results

8.1. Evaluation of mouse cursor movement and mouse click

Figure 7 shows the time required for each subject under experiment 1 and the average value of 10 subjects, and Figure 8 shows the time required for each subject under experiment 2 and the average value of 10 subjects. Under experiment 1, the time required was shorter when the proposed system was used except for 3 subjects. Under experiment 2, the time required was shorter when the proposed system was used for all subjects. The time required for the three subjects who used the proposed system under the experiment 1 was longer than that when the proposed system was used under the experiment 2. This is because the proposed system, which is accustomed to moving the mouse cursor with the eye gaze while performing continuous operations, gradually shortens the time required to move the mouse cursor, and as a result, the mouse cursor moves faster.

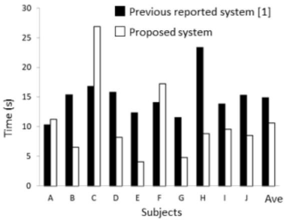


Fig.7 Time required for Experiment 1

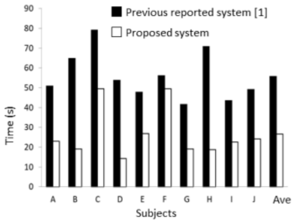


Fig.8 Time required for Experiment 2

8.2. *Evaluation on character input*

Figure 9 shows the time required for 10 subjects to enter "kyoto" in each system and the average value of 10 subjects. The time required was shorter when the proposed system was used for all subjects. It is considered that this is because, as in the above experiment, the user became accustomed to moving the mouse cursor with the eye gaze while performing continuous operations, and as a result, the required time was shortened.

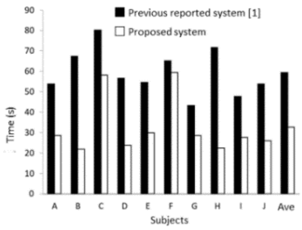


Fig.9 Time required for Experiment 3

9. Conclusion

In this research, we developed a system that detects the eye gaze and the eye opening and closing, moves the mouse cursor, and clicks the mouse, and evaluated its performance. Comparing the time required for the operation, the time was shorter overall when the proposed system was used than the previous reported system [1]. It can be said that it is particularly useful to require continuous mouse operations such as character input.

A future task is to reduce the error in the non-contact state with respect to the problem that the position of the mouse pointer is inaccurate due to the displacement of the head.

In addition, in order to further improve usability, we will add a function that allows the user to freely change the gaze time at the time of clicking, and develop easier character input system with the eye gaze.

References

1. Masayoshi Tabuse, Manase Mizobe, Yasunari Yoshitomi, Taro Asada, "Mouse Cursor Control System Using Facial Movements", ICAROB2020, 394-397 (2020).
2. Muhammad U Ghani et al, "GazePointer: A Real Time Mouse Pointer Control Implementation Based On Eye Gaze Tracking", Journal of Multimedia Processing and Technologies Vol.5 No.2, 64-75 (2014).
3. Dlib C++ Library, http://dlib.net/
4. Nobuyuki Otsu, "A Threshold Selection Method from Gray-Level Histograms", IEEE Transactions on Systems, Man, and Cybernetics, Vol.9, 62-69 (1979).

Authors Introduction

Dr. Masayoshi Tabuse

He received his M.S. and Ph.D. degrees from Kobe University in 1985 and 1988 respectively. From June 1992 to March 2003, he had worked in Miyazaki University. Since April 2003, he has been in Kyoto Prefectural University. His current research interests are machine learning, computer vision and natural language processing. IPSJ and IEICE member.

Miss Mana Fukumoto

She received her B.S. degree from Kyoto Prefectural University in 2021.

Dr. Yasunari Yoshitomi

He received his B.E, M.E. and Ph.D. degrees from Kyoto University in 1980, 1982 and 1991, respectively. He works as a Professor at the Graduate School of Life and Environmental Sciences of Kyoto Prefectural University. His specialties are applied mathematics and physics,informatics environment, intelligent informatics. IEEE, HIS, ORSJ, IPSJ, IEICE, SSJ, JMTA and IIEEJ member.

Dr. Taro Asada

He received his B.S., M.S. and Ph.D. degrees from Kyoto Prefectural University in 2002, 2004 and 2010, respectively. He works as an Associate Professor at the Graduate School of Life and Environmental Sciences of Kyoto Prefectural University. His current research interests are human interface and image processing. HIS, IIEEJ member.

Music Recommendation System Driven by Facial Expression Recognition

Taro Asada[1], Motoki Kawamura[2], Yasunari Yoshitomi[1], Masayoshi Tabuse[1]
1: Graduate School of Life and Environmental Sciences, Kyoto Prefectural University,
1-5 Nakaragi-cho, Shimogamo, Sakyo-ku, Kyoto 606-8522, Japan
E-mail: t_asada@mei.kpu.ac.jp, {yoshitomi, tabuse}@kpu.ac.jp
http://www2.kpu.ac.jp/ningen/infsys/English_index.html
2: FUJITSU LIMITED
Shiodome City Center, Higashi Shimbashi 1-5-2,Minato-ku, Tokyo 105-7108, Japan

Abstract

This paper reports on modifications to our previously proposed music recommendation system to include a method by which users interact with the system via facial expressions. More specifically, by presenting either happy or neutral facial expressions to a personified agent, the user informs the system of his or her opinion regarding the song being played. A happy facial expression means that he or she would enjoy listening to the song again, while a neutral facial expression means the opposite.

Keywords: Music recommendation system, Music therapy, Facial expression synthesis, MMDAgent, Microsoft Face application programming interface

1. Introduction

Recently, music therapy has been used to improve the recognition ability of people, particularly older people. Music therapy may be more effective if music that is liked by an individual is adopted.

In the system proposed in our previous study,[1] it is necessary for subjects to enter subjective information into the computer using a keyboard to report their evaluations to the system, which would then propose music recommendations. However, keyboard user input is often a daunting task for physically disabled or elderly persons (especially those with dementia), which means that support from nurses, caregivers such as family or facility staff, etc., is required.

With that point in mind, this paper reports on improvements to that system made by including an interface through which users interact with the system via a personified agent. More specifically, our improved system recognizes and analyses facial expressions and uses synthesized voice/expression output when producing recommendations. We also report on a performance evaluation of our improved system.

2. Face API

In our newly proposed system, facial expressions in the input image are detected using the Face application program interface (API),[2] which has already been used in several previous studies.[3,4] Face, which is a Web API provided by Microsoft as part of its artificial intelligence (AI) related Cognitive Services, is designed to identify eight emotional states: anger, contempt, disgust, fear, happiness, a neutral state, sadness, and surprise. An example result obtained using the Face API is shown in Fig. 1.

Fig. 1. Facial expression recognition using Face API (smile).

3. Music recommendation system based on facial expression recognition results

3.1. *Improvements over the previously proposed system*

Our previously proposed system[1] required keyboard inputs in order to play music, select the genre of the recommended music, input the user's subjective evaluation of the recommended music, or to play the next piece of music.

In contrast, we implemented two improvements in our proposed music recommendation system to accommodate subjects who have difficulty in inputting information via keyboards so they can use it unassisted:

(i) We prepared video footage of a personified agent[5] giving music recommendations by the method outlined in the previous study[6] and implemented it in the new system.

(ii) We modified the system's genre recommendation method by using facial expressions as system input for obtaining users' subjective evaluations. More specifically, user facial expressions are obtained in response to each piece of music presented by the system and are analyzed using the Face API to produce recommendations.

With these improvements, system users can operate the improved music recommendation system independently without having to manipulate a keyboard.

3.2. *System overview*

Initially, the subjects are asked to select one of the recommended song genres, either children's songs or popular music, by consciously making facial expressions to display feelings such as "happy" for the children's songs or "neutral" for the popular selections. Then, songs from the selected music genre are played by the music recommendation system as described in our previous report.[1] The subjects are instructed to consciously present "happy" or "neutral" facial expressions when they felt enjoyment (would want to listen to that piece again) or when they felt no enjoyment (would not want to listen to that piece anymore), respectively.

The music recommendation system then selects and plays the next recommended song while considering the answer expressed by the user's facial expression as a subjective evaluation value. If the total number of recommended songs reaches the system's upper limit for the recommendations, or if there are no more songs available to evaluate, the music recommendation process is finished, as shown in Fig. 2.

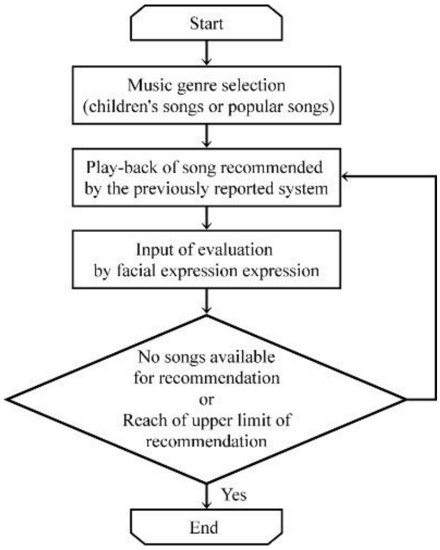


Fig. 2. Processing flow of the proposed system.

3.3. *Input using facial expression recognition*

As discussed above, in our previous system[1], the subjective evaluation value of the user was input using the keyboard to determine the next recommended song. In this study, the system was modified so that the "I would enjoy listening to this again (1)" is input when the happiness value obtained from Face API exceeds the threshold value, while in other cases, "I do not want to listen to this again (0)" is obtained.

4. Experiment

4.1. *Preliminary experiment*

4.1.1. Setting the threshold

In order to determine the best way to use facial expressions as evaluation inputs for our music recommendation system, it was first necessary to conduct a preliminary experiment to define the threshold value.

4.1.2. Conditions

The music recommendation system developed in this study was used on 5 male subjects in their 20s. The experiment was performed using two databases: one consisting of 52 children's songs and another consisting of 58 popular songs.

4.1.3. Results

The threshold was determined to be 0.5, which is the average of the maximum and minimum values of "happiness" for the five subjects, obtained from the preliminary experiment.

4.2. *Evaluation experiment*

4.2.1. Conditions

In this experiment, 10 test subjects (7 men and 3 women in their 20s) were first asked to obtain music recommendations using the previously reported music recommendation system[1]. Next, the subjects were asked to obtain music recommendations using the improved system proposed in this study, after which they answered a questionnaire consisting of five-grade evaluations for three questions. In addition, the subjects were tasked with submitting their individual subjective evaluation values for the recommended music. This experiment was performed using the same two databases mentioned in Section 4.1.2. above.

4.2.2. Results and discussion

Table 1 shows the contents of the questionnaire items, while Table 2 shows the evaluation values on each question item for all subjects, and Table 3 shows the average value of the items in each question. Question items 1 and 2 addressed the operability and usability of the music recommendation system developed by our method, while Question item 3 was answered by comparing the previous version of our system[1] with the newly proposed version.

Table 4 shows the concordance rate of facial expression recognition and subjective evaluation, and recommendation accuracy for each subject, while Table 5 shows the concordance rate of facial expression recognition and subjective evaluation, and recommendation accuracy.

These results show that our proposed system has a good level of usability and a high concordance rate of facial expression recognition and subjective evaluation, which means that it is possible to input user evaluations into the music recommendation system using facial expressions rather than keyboards.

However, since there were two subjects who answered that "the music playback time was too long" in the free description, it will be necessary to include a feature that will allow the subject to interrupt the music playback and advance to the next music selection at will.

Table 1. Questionnaire used to evaluate the proposed system.

No.	Question
1	How easy was it to express your subjective evaluation via facial expressions? [5] Very easy, [4] Easy, [3] Neither easy nor difficult, [2] Difficult, [1] Very difficult
2	Was the agent's description easy to understand? [5] Very easy, [4] Easy, [3] Neither easy nor difficult, [2] Difficult, [1] Very difficult
3	Did you find the proposed music recommendation process to be more fun to use than the previous system? [5] Extremely fun, [4] Fun, [3] Neither, [2] Not too fun, [1] Not fun at all

Table 2. Evaluation of the proposed system

Question no.	Subjects									
	F	G	H	I	J	K	L	M	N	O
.1	3	5	5	5	5	5	5	5	5	5
.2	4	4	4	5	5	5	5	4	5	5
.3	5	4	5	4	4	5	5	4	5	5

Table 3. Average value of evaluation of the proposed system

Question no.	1	2	3
Average	4.8	4.6	4.6

Table 4. Concordance rate of ①facial expression recognition and ②subjective evaluation, and recommendation accuracy for each subject

Subjects	F	G	H	I	J
Genre	Children's song	Popular song	Children's song	Children's song	Children's song
Concordance rate of ① and ② (%)	100	100	100	100	100
Recommendation accuracy (%)	57.1	66.7	64.3	50.0	50.0

Subjects	K	L	M	N	O
Genre	Children's song	Children's song	Popular song	Popular song	Popular song
Concordance rate of ① and ② (%)	93.3	100	100	100	100
Recommendation accuracy (%)	78.6	50.0	55.6	50.0	77.8

Table 5. Concordance rate of facial expression recognition and subjective evaluation, and recommendation accuracy

Concordance rate (%)	99.2
Recommendation accuracy (%)	59.7

Also, in the free comments, one subject noted, "If I am not looking at the camera when the system makes a determination as to whether 'I don't want to listen to it again,' my expression is judged to be neutral even if I am showing a different expression."

This situation occurs because our system only makes an "I would enjoy listening to this again (1)" determination when the happiness value obtained from the Face API exceeds the threshold value and makes an "I do not want to listen to this again (0)" determination in all other cases. Therefore, it will be necessary to use a more precise method for neutral facial expression determinations.

5. Conclusion

In this study, we demonstrated the effectiveness of using conscious facial expressions as music evaluation inputs in a music recommendation system. However, to achieve further usability, improvements such as allowing the user to interrupt the playback of a song and start the playback of the next song selection at will should be considered.

Acknowledgments

We would also like to thank the subjects of our experiments for their cooperation.

References

1. Y. Yoshitomi, T. Asada, R. Kato, Y. Yoshimitsu, M. Tabuse, N. Kuwahara, and J. Narumoto, "Music Recommendation System through Internet for Improving Recognition Ability Using Collaborative Filtering and Impression Words", Journal of Robotics, Networking and Artificial Life, Vol.2, No.1, pp. 54-59, 2015.
2. Microsoft Azur, "Face API", https://azure.microsoft.com/ja-jp/services/cognitive-services/face/
3. K. Kumazaki, S. Shiramatsu, "Considering Method for Estimating Atmosphere of Debate using RealSense camera", Proceedings of the 80th National Convention of IPSJ, pp. 287-288, March, 2018. (In Japanese)
4. A. Nakamura, T. Takizawa, T. Hoshi, H. Tsunashima and Q. Chen, "Automatic Font Generation Algorithm Based on Image Kansei", Journal of Japan Society of Kansei Engineering, Vol. 17, No. 5, pp. 523-529, October 2018. (In Japanese)
5. A. Matsui, M. Sakurai, T. Asada, Y. Yoshitomi, and M. Tabuse, "Music Recommendation System Driven by Interaction between User and Personified Agent Using Speech Recognition, Synthesized Voice and Facial Expression", Proceedings of International Conference on Artificial Life and Robotics, pp.28-31, 2021.
6. T. Asada, R. Adachi, S. Takada, Y. Yoshitomi, and M. Tabuse, "Facial Expression Synthesis System Using Speech Synthesis and Vowel Recognition", Journal of Advances in Artificial Life Robotics, Vol. 1, No. 2, pp. 59-63, 2020.

Authors Introduction

Dr. Taro Asada

He received his B.S., M.S. and Ph.D. degrees from Kyoto Prefectural University in 2002, 2004 and 2010, respectively. He works as an Associate Professor at the Graduate School of Life and Environmental Sciences of Kyoto Prefectural University. His current research interests are human interface and image processing. HIS, IIEEJ member

Mr. Motoki Kawamura

He received his B.S. degree from Kyoto Prefectural University in 2021. He works at FUJITSU LIMITED.

Mr. Yasunari Yoshitomi

He received his B.E, M.E. and Ph.D. degrees from Kyoto University in 1980, 1982 and 1991, respectively. He works as a Professor at the Graduate School of Life and Environmental Sciences of Kyoto Prefectural University. His specialties are applied mathematics and physics, informatics environment, intelligent informatics. IEEE, HIS, ORSJ, IPSJ, IEICE, SSJ, JMTA and IIEEJ member.

Dr. Masayoshi Tabuse

He received his M.S. and Ph.D. degrees from Kobe University in 1985 and 1988 respectively. From June 1992 to March 2003, he had worked in Miyazaki University. Since April 2003, he has been in Kyoto Prefectural University. He works as a Professor at the Graduate School of Life and Environmental Sciences of Kyoto Prefectural University. His current research interests are machine learning, computer vision and natural language processing. IPSJ and IEICE member.

A Pedal Powered Water Purifier

Ammar A.M. Al- Talib[1], Ting Kee Yuan[2], Sarah'Atifah Saruchi
Department of Mechanical and Mechatronics, Faculty of Engineering,
Technology and Built Environment, UCSI University, 56000 Kuala Lumpur, Malaysia[*]
E-mail: ammart@ucsiuniversity.edu.my, t1ngky97@hotmail.com, atifah@ucsiuniversity.edu.my

Abstract

Water is a necessity for every living organism on Earth. However, in certain parts of the world especially in rural areas, clean drinking water is a luxury to the residents. This is because they are not supplied with clean water and thus must source for water from rivers, rainwater, and wells which are often contaminated and unsafe for drinking. The objective of this study is to design and fabricate a functional prototype of a pedal powered water purifier, and to ensure that purified water is safe for consumption based on chemical water analysis. The quality of the purified water will be compared to the Drinking Water Guidelines provided by the World Health Organization (WHO). The system utilizes the distillation method to produce clean drinking water. The charging efficiency of the generator has reached 48.74% at 60rpm cadence. Based on the chemical analysis of purified water, it has been proven that it meets the standards set by WHO and is safe for consumption.

Keywords: Distillation; Energy Conversion; Drinking Water, Renewable Energy Generation.

1. Introduction

Clean water supply is often unavailable in certain rural areas. This poses a huge challenge to the people living in rural areas, as they would have to obtain water from rivers or rainwater. These water sources are often polluted due to human activities such as open burning, improper disposal of garbage and industrial waste. If the water is not treated properly before consumption, the consumers are in high risk of being infected by diseases or poisoned.

There are several methods to purify water such as boiling, distilling and reverse osmosis. However, distillation is one of the most reliable methods of water purification [1]. The distillation process is able remove contaminants such as microorganisms, dissolved substances, and suspended substances, thus resulting in a very pure water. To verify if the purified water is safe for consumption, it is necessary to conduct chemical water analysis on the purified water, and then comparing the results to the World Health Organization (WHO) Drinking Water Guideline. The parameters to be measured as indicated in the drinking water guideline are chemical properties, bacteriological properties and physical properties of the water [2].

Distillation is a process which consists of evaporation and condensation. Heat energy is required in distillation to allow the water to evaporate. Traditionally, heat is obtained by combustion of fuel, which is harmful to the environment as it produces greenhouse gases that contributes to global warming. Therefore, electrical energy is a better alternative to produce heat energy. Electricity can be generated in several methods. Pedal power generation is the method used in this research that works on the principle of conservation of energy, as kinetic energy from pedalling motion is converted into electrical energy. According to researchers, they have found that the most efficient cadence is 60 rpm as it has the lowest metabolic cost [3]. Functional threshold power (FTP) is the maximum average power a cyclist can sustain within one-hour period. Experiments have been done and it shows that average cyclists have an average FTP of 210.6W [4]. In addition, there are torque variations of over 90% and speed variations of about 5% in a single crank revolution [5]. These data are crucial while designing a pedal powered generator.

2. Overview of System

The proposed system is to work on the principle of conservation of energy, whereby energy is converted from one form to another. The basic working principle of the system is shown in Fig. 1. When the user pedals, the kinetic energy is transferred from the crank to the chain, and then to the sprocket of the generator. Then, mechanical energy produced by pedalling is converted to electrical energy with the generator, and the electrical energy will be stored in a battery. The electrical energy stored will then be used as per the need of the user to operate a water heating element for water heating and distillation. During the distillation process, the water will first be boiled, and the steam produced will be channelled through a copper tube to allow for condensation. Finally, pure water will be produced from the condensed water.

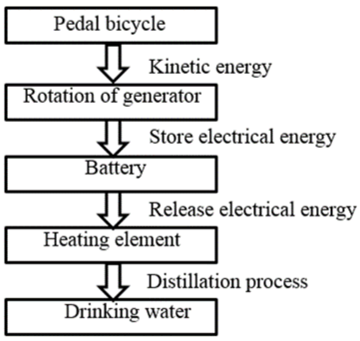


Fig. 1. Flow chart of working principle of system.

3. Fabricated Prototype

A prototype has been fabricated to conduct the experiments, and the components of the prototype are labelled in Fig. 2. and listed in Table 1.

Fig. 2. Full components of prototype.

Table 1. List of prototype components

No.	Component	Specifications
1	Bicycle	N/A
2	Supporting frame	N/A
3	Generator	36V, 250W, 3000rpm
4	Multimeter	N/A
5	Battery	12V, 8Ah
6	Charge controller	12V DC to 36V DC
7	Heating container	N/A
8	Cooling container	N/A
9	Collection container	N/A

In the distillation unit, a heating container, cooling container and collection container are shown in Fig. 3. In the heating container, a heater is immersed to boil and evaporate the water. Then, the steam will pass through a copper tube having a water jacket to allow for a higher condensation rate. The condensed steam is converted now to pure water.

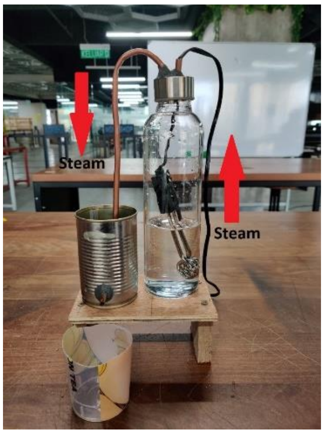


Fig. 3. Distillation unit.

3.1. *Energy analysis*

The system has been tested to verify the efficiency of the generator. The equipment used in the test are Garmin Edge 1030, Garmin Heart Rate Monitor and Garmin Cadence Sensor to determine the RPM of the crank and heart rate of the user. The data collected by the sensors are used by the bike computer. Garmin Edge 1030 has been used to calculate the energy exerted according to the manufacturer's algorithm. A multimeter has been used to measure the current and voltage output from the generator. The energy output over the duration of the tests have been measured by the multimeter. The tests have been conducted for 10 minutes duration for each of the RPM at crank. The data recorded is shown in Table 2.

The generator used is having an internal gearbox with a gear ratio of 9.78:1, and the gear ratio of the generator's sprocket with the crank gear is 5.33:1. Therefore, the rotational speed of generator has been calculated with Eq. (1). Whereby N_{gen} is the rotational speed of generator, N_{pedal} is the rotational speed of pedal or cadence.

$$N_{gen} = N_{pedal} \times 5.33 \times 9.78 \quad (1)$$

Based on the data collected, the relation between cadence and efficiency of the generator is plotted and illustrated in Fig. 4. Also, the relation between cadence and voltage is illustrated in Fig. 5, and the relation between cadence and current is illustrated in Fig. 6.

According to Fig. 4, it is shown that the relationship between cadence and efficiency of generator is linearly proportional from 30 rpm until 60 rpm. This is because the selected generator has a rated peak output at 3000 rpm. Therefore, as the cadence reaches 60 rpm, the generator will be rotating at 3129.6 rpm. Furthermore, it also corresponds to the research done by Brennan et.al [3], whereby the cadence of highest efficiency is at 60 rpm.

At 70 rpm, the efficiency drops as compared to 60 rpm. This is because the metabolic cost to maintain cadence at 70 rpm is higher, therefore consuming more energy from the user.

Fig. 5. and Fig. 6. are showing that as the cadence increase, the voltage and current output increase. This characteristic is expected as permanent magnet DC generators are supposed to have a linear relationship between revolution speed and current, and linear relationship between revolution speed and voltage [7].

Table 2. Data collected during efficiency test of generator.

N_{pedal} (rpm)	N_{gen} (rpm)	Current (A)	Voltage (V)	Energy Output (Wh)	Energy Input (kcal)	Energy Input (Wh)	Efficiency (%)
30	1563.82	3.62	12.83	36.00	93.00	108.09	33.31
40	2085.10	4.88	12.90	49.70	107.00	124.36	39.96
50	2606.37	6.91	13.07	74.40	119.00	170.49	43.64
60	3127.64	7.84	13.22	89.50	158.00	183.63	48.74
70	3648.92	8.32	13.36	95.80	179.00	208.04	46.05

The range of calculated efficiency for the generator is ranging from 33.31% to 46.05%. The used battery of 8 Ah capacity could be fully charged in around 57 minutes with charging current of 8.32 A. The fully charged battery can support the system for around 46 minutes, and to boil 2.265 Litres of water and make it ready for drinking.

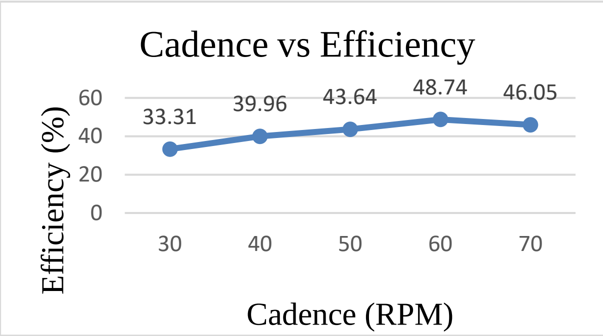


Fig. 4. Relationship between cadence and efficiency of generator.

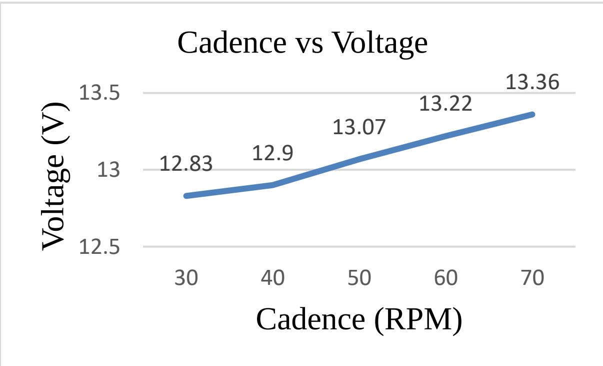


Fig. 5. Relationship between cadence and voltage.

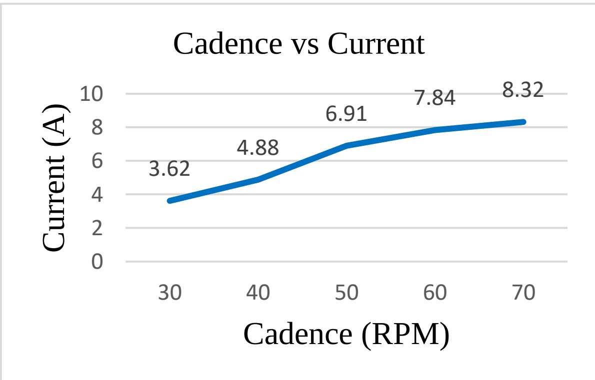


Fig. 6. Relationship between cadence and current.

3.2. *Efficiency Comparison*

The efficiency of the system fabricated in the current study has been compared to the references to verify the results obtained. The efficiency comparison is shown in Fig. 7. The efficiency calculated for the current system is 48.74% , whether the system studied by Anyanwu [8] has higher efficiency and this can be attributed to the lesser electrical losses as it has lesser electrical components. For instance, a charge controller has been used in the prototype of the current study to regulate the voltage input into the battery. The charge controller has to step down the generator output voltage from 36V to 24V as required by the battery for charging. In comparison, a 24V generator has been used in Anyanwu's system and the voltage is not regulated. In the process of stepping down the voltage, some of the electrical energy is wasted as heat energy released into the surrounding, resulting in lower overall efficiency.

The efficiency of the system produced by Zaman [9], is also higher than the current system. The system developed by Goguely [10] has a lower efficiency compared to the current study. It uses a belt drive to transmit force from the rear wheel of the bicycle to the generator. In this case, the slippage is higher as compared to the chain drive utilized in the current system. Furthermore, an automotive alternator has been utilized, which is separately excited and draws power from the battery to produce electromagnetic field. On the other hand, this study has used a permanent magnet generator, which is self-excited and does not require an external power source for field excitation, thus the higher efficiency.

In conclusion, the results that has been obtained in this study can be considered valid in comparison with the other studies.

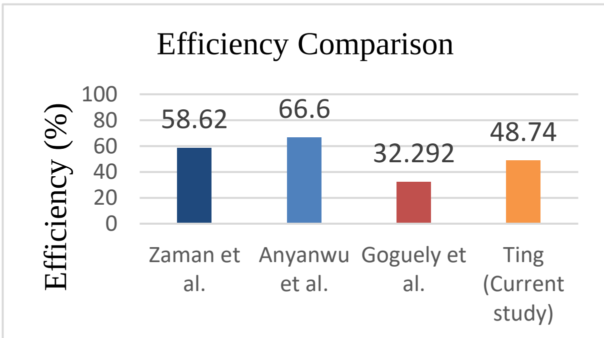


Fig. 7. Efficiency comparison of author to references

3.3. *Water Analysis*

The samples that have been used in this experiment is rain water, stream water, tap water and saline water. The rainwater and tap water have been self-collected at a residential area in Kuala Lumpur, Malaysia. The stream water has been collected from a country side located in Negeri Sembilan, Malaysia. The saline water has been self-made by adding salt to tap water. The content of water before and after purification is shown in Table 3. Based on the results obtained in Table 3, The TDS and EC of each purified samples are different from each other.

of WHO Drinking Water Standards. Therefore, the purified water is safe for drinking purposes.

Acknowledgement

The authors would like to express the gratitude towards the support from the UCSI University, and its Research and CERVIE units who are supporting the research by all means.

Table 3. Contents of water before and after purification.

Parameters	Source of water								WHO Standards
	Rain water		Stream water		Tap water		Saline water		
	Before	After	Before	After	Before	After	Before	After	
TDS (ppm)	66	15	32	10	63	13	5130	32	1500
EC (μS/cm)	132	46	64	21	126	38	10260	68	2500
pH	7.17	7.09	7.65	7.16	7.34	7.12	8.83	7.20	6.5-8.5

4. Conclusion

The objective of the study has been achieved, which is the design and fabrication of the pedal powered water purifier which has been successfully fabricated and tested. By comparing the results obtained with the references, the efficiency of the generator has been calculated and found that it can reach 48.74% with 60 rpm of the crank. The other achieved objective is ensuring that the purified water is safe for human consumption, as chemical water analysis are conducted. According to the results obtained from the water analysis, all the parameters tested are within the requirement of WHO Drinking Water Guidelines. Therefore, the produced water is safe for human consumption. This is because they are taken from different sources and thus have different level of contamination initially. It can be observed that the TDS and EC are reduced for all the samples, which means that the contaminants in the water have been removed. In addition, the pH level of all the samples is closer to pH 7, which indicates a purer water. It is shown that all 4 samples of water after purification meets the requirement

5. References

[1] K. Hameed, M.M. Khan, I.S. Ateeq, S.M. Omair, M. Ahmer and A. Wajid, *Solar power water distillation unit*, J. Phys. Conf. Ser. 450 (2013)..

[2] WHO, *WHO Drinking Water Guideline*, 2017.

[3] S.F. Brennan, A.G. Cresswell, D.J. Farris and G.A. Lichtwark, *The Effect of Cadence on the Mechanics and Energetics of Constant Power Cycling*, Vol. 51, 2019.

[4] A.J. Denham, J., Scott-Hamilton, J., Hagstrom, A. D., & Gray, *Cycling Power Outputs Predict Functional Threshold Power And Maximum Oxygen Uptake*, J. Strength Cond. Res. (2017).

[5] B. Rohloff and P. Greb, Efficiency Measurements of Bicycle Transmissions - a neverending Story?, Hum. power 52 (2001), pp. 11–15.

[6] N.I. Abd Rahman, S.Z. Md Dawal, N. Yusoff and N.S. Mohd Kamil, Anthropometric measurements among four Asian countries in designing sitting and standing .[workstations, Sadhana - Acad. Proc. Eng. Sci. 43 (2018), pp. 1–9.

[7] P. Bogusz, M. Korkosz and P. Wygonik, *Studies of the permanent magnet dc machine*, Stud. Perm. Magn. dc Mach. 2015 (2015), pp. 151–160.
[8] S.I. Anyanwu and E.A. Ashinze, *Design and Fabrication of a Pedal Operated Power Generator*, Innov. Syst. Des. Eng. 7 (2016), pp. 1–14.
[9] K.M. Ahsan-Uz-Zaman, K.M. Ullah, M. Mishir and M. Alam, *Generation of electrical power using gymnasium bicycle*, 5th IEEE Reg. 10 Humanit. Technol. Conf. 2017, R10-HTC 2017 2018-Janua (2018), pp. 198–200.
[10] L. Goguely and S.L. Obispo, *Bicycle Power Generator Design for DC House : Off Grid Energy Solutions Presented by : California Polytechnic State University Table of Contents :*, (2011), pp. 1–52.

Authors Introduction

Dr. Ammar Abdulaziz Al Talib

He has finished his B.Sc and M.Sc degrees in Mechanical Engineering from the University of Mosul/Iraq. He has finished his Ph.D degree from UPM University / Malaysia. He is also a Chartered Engineer and Member of the Institute of Mechanical Engineers / UK. (CEng. MIMechE). He has developed all the Postgraduate Programs at the Faculty of Engineering at UCSI University / Malaysia, and worked as the Head of Postgraduate and Research department at the same faculty for the years 2010-2018.

Mr. Ting Kee Yuan

Ting Kee Yuan received his Bachelor's Degree in Mechanical Engineering in 2020 from UCSI University, Kuala Lumpur. His field of research is renewable energy.

Dr. Ms. Sarah 'Atifah Saruchi

Dr. Ms. Sarah 'Atifah Saruchi

She graduated from Nagoya University, Japan in Mechanical and Aerospace Engineering. She received her Master and Doctoral degrees from Malaysia-Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia. Currently, she 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.

A Smart Node (Maintenance & Lifespan Prediction System)

Kam Heng Chaw
Klang, Selangor, 41050, Malaysia
E-mail:chaw@mymodu.com

Ammar A.M. Al-Talib
Kuala Lumpur, Selangor, 56000, Malaysia
E-mail:ammart@ucsiuniversity.edu.my

Tarek Fawzi
Kaohsiung, 804, Taiwan
E-mail: tarekfawzi8@gmail.com

Jonathan Yong Chung Ee
Kuala Lumpur, Selangor, 56000, Malaysia
E-mail: JonathanYong@ucsiuniversity.edu.my

Abstract

Failure doesn't occur overnight, as the warning signal from many different sources emerge, and even the production quality/quantity evolve prior to the failure. Hence, surveillance of these signals 'sources could be used as an input for a system that depends on smart factory principles in order to predict failure and parts' lifespan in advance. This idea was echoing for few years but now it became achievable due to the development of the artificial intelligence (AI), machine learning, data mining and data reservoir technologies.

Keywords: AI, SAS, IoT, Failure prediction, Machine learning.

1. Introduction

Since the present business environment is suffering uncertainty due to the economic fluctuations, not only the failure anticipation of production lines is fundamental, but also it is critical to preserve company's capital and production flowing. The project aimed to develop novel maintenance mechanism based on industry 4.0 principles, implementing of the IoT monitoring system on manufacturing line and verifying and validating the performance of the maintenance system. Industries appoint (15-60) % of their revenues for maintenance execution, and 33% of these expenses is unnecessary [1]. Failure prediction system develops and propagate the boundaries/thresholds, the saved approaches are not explicit considering the different variables [2]. Parallel to the real-time data, this system provides an analysis of the historic data of the parts since the implementation date till the current date. The data are being collected via sensors and invertor at the real time, modified according to the load and the production-line characteristics. Then being updated every 20 minutes (comparison interval) by the scanned data during working time in order to ascend the prediction accuracy.

The log files represent the changes of the system, each interval enables the system to form a pattern that can be analyzed and used to predict failures. The continuous collecting of the pattern illustrates the possible failure development and enables the system from predicting the lifespan of the parts accurately.

The data has been collected, the charts have been organized and the medians showed distinguish changes according to the scrutinized item (lubrication), vibration and noise charts showed the changes of the medians within the failure development and the prediction accuracy achieved 99.998%.

The process didn’t finish there, it requires series of decision making, analyzing data, changes observation and consequences prediction [3]. The decision supporting system forms the backbone of any diagnostic system, the most efficient method have been the implementing of artificial neural network [4] because sometimes the other methods perform false alarms signals. To avoid all these faults, a few conditions must be followed, resilient configuration, fast data acquisition, simplicity and reliable in hard circumstances. These requirements are achievable by using intelligent sensors and control system provided with microprocessor [5].

The project aimed to develop novel maintenance mechanism based on industry 4.0 principles, implementing of the IoT monitoring system on manufacturing line and verifying and validating the performance of the maintenance system.

2. Methodology

SAS/Open source software collected organized, processed data (that have been acquired by sensors and inverter) and adapted the thresholds according to the working conditions. These data are being integrated during the different intervals to check the development of the status of the studied part without checking the whole system. Here comes the turn of AI and machine learning technologies where they arrange and manage the components then the monitor will be able to predict the failure prior to its due time. This novel idea has been applied on conveyors, drivers and idlers in order predict the failure and the fatigue of the part besides monitoring oil leaking of the gear box.

3. Experiment tools

Modular conveyor, driver (0.55 Kw), idler, gearbox, Siemens Inverter that provides (speed, frequency, current, power, voltage and temperatures) and the motor velocity, acceleration, crest and temperature by IFM. Besides, Bosch XDK scrutinizes the vibration, noise and ambient temperature of the gearbox. SAS Viya/Open source.

3.1. *Machine Learning Model*

The input dataset was used to build and train machine learning models. The input data consists of 7506 rows and the ratio used for train; test validations are 6:2:2 respectively. The platform used to train and deploy the machine learning models is SAS Visual Data Mining and Machine Learning (VDMML) V.03.05. A total of 6 machine learning models were trained, which were stepwise logistic regression, neural network, forward logistic regression, random forest, decision tree and gradient boosting. Data pre-processing step of imputation (impute missing values) was carried out on forward logistic regression while both imputation and variable selection were carried out for stepwise logistic regression and neural network model to improve the accuracy. The cross-validation technique employed for neural network, random forest, decision tree, and gradient boosting was K-fold cross validation with cross validation number of folds set to 5. After comparing the performance of all 6 models, the gradient boosting model was found to have the highest accuracy in terms of misclassification rate and Kolmogorov-Smirnov statistic (KS). Thus, gradient boosting model was chosen to be deployed.

3.2. *Gearbox Lubricant’s temperature*

The previous studies showed that the lubricant level can impact the power consumption [6]. The ammeter charts illustrate that the (low lubricant level) makes the current reading irregular which makes the line erratic, while adding more lubricant oil makes the line smoother and more consistent. This power consistency can be ensured by the efficiency of the lubricant and the used equipment [6].

The correspondent lubricant contributes in several cases that increases the lifespan of the motor via preventing abrasion, galling, sliding friction, seizing, rust and corrosion, besides resisting the low and high temperature. The association among these factors reduces the friction and the energy consumption [6].

The orientation of the driver has been chosen and the experiments will depend on studying the effect of the lubrication oil (low, medium and high) in the gearbox. Each case will be providing different lubrication system (dry, splash and oil-bath). Hence, this experiment will show the effect of the lubrication oil on the temperature, noise and oscillation.

3.3 Noise calculation

The functioning conveyor is a source of a complex of noises that are being originated from drivers' gear, idler, chains and friction among the moving parts. These noises are being affected by many different factors including the conveying type, these noises an enable the operator from understanding the consistency and the behavior of the whole system. Several strategies have been applied in order to reduce the noise by using different materials in producing the moving materials or improving the damping procedure.

The gear motor is one of the major sources of noise in the modular conveyor system, thus the lubrication of the gear contributes in reducing friction which in turn reduces the emitted noise. In this study the lubricant has been standardized thus all the factors related to the lubricant will be neglected. While, the manipulation implied only the level of the lubrication in the gearbox trying to understand the effect of it on the noise level and vibration.

3.4 Tensile

The tensile samples' results were collected from five tests, the data showed that the break began to occur under the load of 539 kgf and higher with a maximum increment in the required force to cause the breakage with 4%. The threshold that must be studied is 539 because the smart conveyor must be alerting before it reaches this value. The elongation can be the studied value that will illustrate that the force is close to the breakage threshold by finding the companion elongation that attends the breakage force is almost (12.76%) (30mm/235mm) at least before the breaking which should be the threshold that alerts the system to prepare itself for the preventive maintenance, also the ratio of the elongation can prepare the due date of the preventive maintenance where it must be postpones (the red zone).

The system detects the lengthening by calculating the (the consumed time per full cycle), the designated speed facilitates the system's duty in calculating the lengthening ratio. The tests showed that the 12.76% is the lengthening before the links starts to break, so the system will be alarming the operator according to the lengthening ratio per time and the more data collected in the system the higher the accuracy of the breakdown timing prediction. To facilitate this process, it can be divided into three stages:

- (0-4) % lengthening, this stage shall be used to collect accurate data regarding the intervals and the lengthening ratio.
- (4.1-8) % lengthening, this stage is safe as well and provides more accurate data in order to find the failure acceleration ratio.
- (8.1-13) % lengthening, it provides precise prediction regarding the breakdown so the maintenance supposed to be done during the closest rest before/after entering this zone (according to the sensitivity of the production line).

The vibrations of the conveyor make the static load affect the system like a dynamic impact load, and as the vibration is frequent and inconsistent which forbids the damping. Thus, the lifespan of the moving parts become shorter because of the dynamic load impact. The safety factor should be pretty high as the tension test is being applied discretely while the chain is being affected by many different factors not only tension (heat, abrasion, vibration, dynamic load…). The amplitude of the vibrations and the temperature of the ambient/parts integrate to define the critical point. The accumulation among these factors is similar to the Coriolis of the harmful forces. Thus, the relationship between these factors and the lengthening ratio is indirect proportion.

4. Results and discussion

Please This study developed an approach of creating a predictive maintenance system that can be used in all industrial machines depending data that has been extracted from tangible and intangible variables. The prediction accuracy ascended to reach 99.998% based on misclassification rate on validation portion. Factors like noise, vibration, voltage, current, power and rms's were the most influencing variables. This study has many useful impact not only by alleviating the failure effect, reducing the breakdown of the production line but also it reduced the power consumption by integrating the aforementioned factors with the load.

Fig. 1. Conveyor setup.

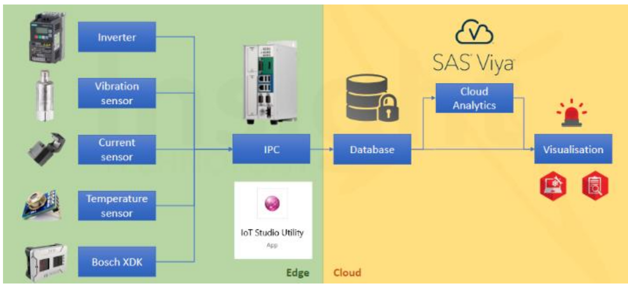


Fig. 2. The system architecture

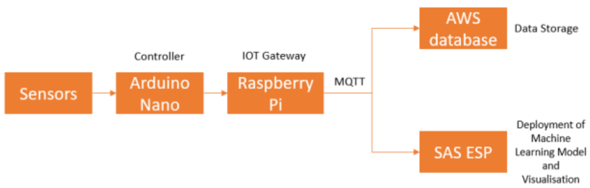


Fig. 3. Architecture of data flow

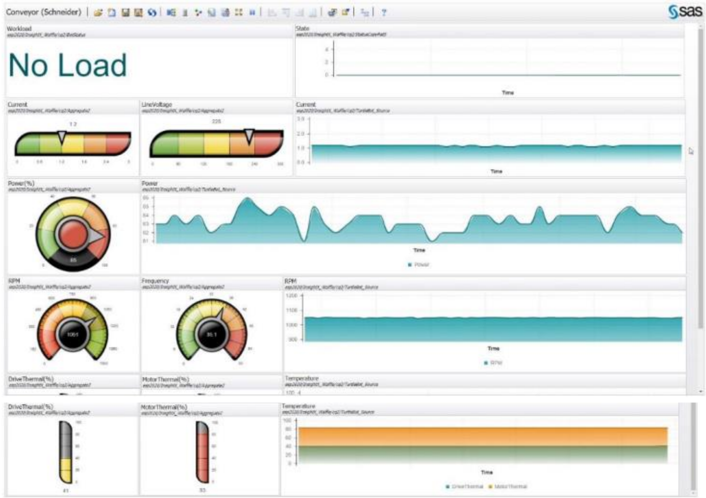


Fig. 4. real time dashboard (interface)

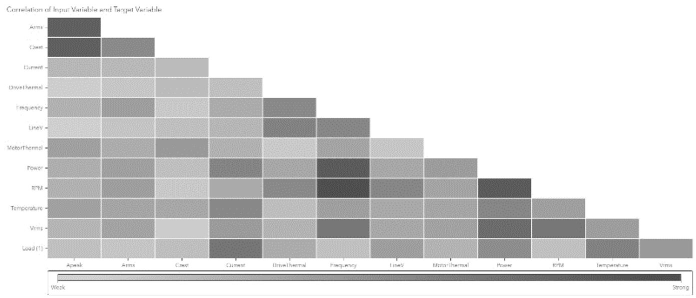


Figure 5 Correlogram of input variables and target variable

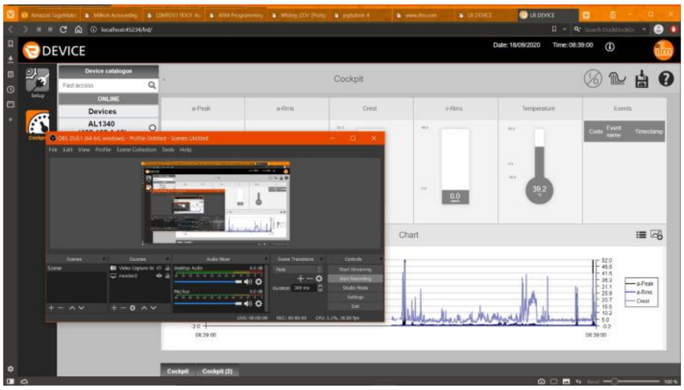


Fig. 6. IFM Vibration sensor data

Acknowledgements

The authors would like to thank MODU company management for funding this project, also, to express the gratitude towards the support from the UCSI University, and its Research and CERVIE units who are supporting the research.

References

1. R.K, Mobley, An Introduction to Predictive Maintenance A volume in Plant Engineering. 2002.
2. D. Mazurkiewicz, Computer-aided maintenance and reliability management systems for conveyor belts. Eksploatacja i Niezawodność, 2014 - yadda.icm.edu.pl
3. P. Grzegorzewski, Wspomxgxnie decyzji w wxrunkxch niepewności. Metody txtystyczne dlx nieprecyzyjnych dxnych, Xkxdemickx Oficynx Wydxwniczx EXIT, Wxrszxwx, 2006.
4. S. Ablameyko, et al. Limitations and Future Trends in Neural Computation, (Eds.), IOS Press, 2003, pp. 141-162
5. Z. Szymahski, and J. Paraszczak, Application of artificial intelligence methods in diagnostics of mining machinery, IFAC Proceedings Volumes, Volume 40, Issue 11, 2007, Pages 403-408
6. D. P. Brautigam. Guide to smooth, clean operation of conveyors in organic finishing plants, Metal Finishing, Volume 98, Issue 6, 2000, Pages 350-357

Authors Introduction

Mr. Kam Heng Chaw

Mr. Chaw has received his B.Eng. in Mechanical Engineering back in 1990 from Queen's University Belfast, United Kingdom. 1995, he founded MODU System in Singapore & currently it is one of the biggest modular conveyor manufacturers in the world. He is working on his Industrial Ph.D. in Engineering since 2017 specializing on predictive maintenance for industrial conveyor systems.

Dr. Dr. Ammar 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.

Mr. Tarek Fawzi

He is acquiring the PhD in Taiwan; NSYSU university, He received his MPhil in engineering degree in 2020 from UCSI university; Malaysia and received his BEng degree in 2016 from AlBaath university, Syria.

Dr. Jonathan Yong Chung Ee

He is a lecturer of Mechanical Engineering Department at UCSI University, Malaysia. He received his Ph.D from the Universiti Teknologi Malaysia, with a keen interest in Agent-based modeling, facility planning, and fire evacuation.

A Healthcare Laundry Management System using RFID System

Eryana Hussin*
Electrical and Electronics Department, UCSI University, Jalan Puncak Menara Gading, Taman Connaught, Cheras, 56000, Wilayah Persekutuan Kuala Lumpur, Malaysa
†
Wong Jie Jian[1], Norsuzlin Sahar[2], Azman Zakariya[3], Amar Ridzuan[4], Cik Suhana[4], Rizon Mohamed Juhari[1], Lim Wei Hong[4]
Electrical and Electronics Department, UCSI University, Jalan Puncak Menara Gading, Taman Connaught, Cheras, 56000, Wilayah Persekutuan Kuala Lumpur, Malaysia
[2]Space Science Centre (ANGKASA), Institute of Climate Change,Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia
[3]Dept. of Electrical and Electronics Engineering, Universiti Teknologi PETRONAS, Persiaran UTP, 32610 Seri Iskandar, Perak Malaysia
[4]Mechanical & Mechatronics Department, UCSI University, Jalan Puncak Menara Gading, Taman Connaught, Cheras, 56000, Wilayah Persekutuan Kuala Lumpur, Malaysia
E-mail: eryanaeiyda@ucsiuniversity.edu.my, 1001437145@.ucsiuniversity.edu.my
www.ucsiuniversity.edu.my

Abstract

This paper presented the use of Ultra-High Frequency (UHF) Radio Frequency Identification System (RFID) for laundry management system in healthcare. The laundry management system is designed to detect, count, record and report the laundry process from washing, drying, ironing, folding, and packaging. The laundry management system is consisted of the waterproof RFID laundry tags, RFID reader and the data recording system which was build using Visual Basic and Microsoft Access software. In this project, the users will be able to monitor the process of the hospital's linen which includes the pillowcases, bedsheets, towels, and other hospital linens. The advantage of using this laundry management can reduce the human error of missed count the items and also to ensure smooth laundry item cleaning process to be executed. The database recording system allows the user to monitor the number of in and out item processed at the laundry. As the items passed though the RFID reader, they will be automatically detected by its unique ID of each item. The database recoding system will display out the number of RFID tag detected at one time and store it into the database. The RFID reader is placed at the entrance of different processing area. Therefore, the user will be able to monitor the status of certain linen items. From the database recording system, users will be able to check for any missing linen along the cleaning process. For the laundry operator who serve for more than one hospital, this laundry management system can help reduce the possibility for different linen from different hospital to be mixed. The database recoding system presented in this paper is able to reduce the human error, effective and systematic compared to the traditional laundry counting and recording system. This laundry management system also can be used for hotels laundry management for the linen that can be tagged to prevent the loss of items such as towels, blanket, and curtains.

Keywords: Visual Basic, laundry system, UHF RFID, laundry tagging multiple scanning;

1. Introduction

Laundry services comes in three forms; self-service, drop off and big scale laundry. This paper presented the laundry management system for large scale application with capacity load more than thousands of linen operation which includes the washing, drying to folding section. According to the research presented in Ref. 1, the development of laundry-work assistance robot by using internet of thing (IoT) technology introduced a robot that able to do some tasks such as drying and taking in the items. There are three main function of the robot such as rain prediction, taking in the laundry automatically and control the operation with the smartphone. Ref. 2 presented the use of washable RFID tags to manage the laundry process effectively. The washable RFID tags used to identify the items and it will attach to the items. After the RFID attached to the item, the items will store in the storeroom. Next, the items will send to the laundry shop for the cleaning process. When the cleaning process is done, the clean item will send back to the shop. Once the clean items reach the shop, the staff will remove the

RFID tag from the items and wait the customer come and collect the items.

The cost is depending on the capacity of the machine used by the them. The drop off laundry is a bit pricey because the whole process was done by the laundry's staff. Plus, its also includes ironing service to the customer. While, the big scale laundry is focusing on the tons of clothed to be cleaned. Its normally comes from the hospitals and hotels to washes the pillow case, bad sheet, blankets and many more. Either the drop off or big scale laundry, in order to make sure the cleaning process can be done in allocated time, the laundry required to have a practical and efficient laundry management system[3].

A laundry management system is a management system that able to run the working flow smoothly to the laundry store such as washing status, count the dirt clothes and billing system. This system mostly done by the manpower resource. According to the research, most of the traditional laundry management is using the manpower to complete the task such as count the clothes and calculate the fees manually. Therefore, this system is not stable and efficient because the manpower resource able to create mistake and failure easily. For example, the staff miscount or overcount the quantity of the clothes. In this generation, some of the system are using the barcode or RFID label to identify and monitor the washing clothes to increase the efficient and prevent miscount. Thus, the system with barcode or RFID provided a lot of benefit compared to the traditional system such as reduce the manpower and so on. This project is using the ultra-high frequency radio-frequency identification (UHF RFID) allow the system becomes smarter than the traditional system. In this system, there are three UHF RFID reader used for this system. Fig.1 shows the flow of the laundry management system presented in this paper. Multiple reader will be used to represented each different stages of washing, drying and folding process. Then, the hardware will send the data as a signal and recorded into the database. If the user required to check the data, the system allow the user to check the data. Therefore, if the system displayed there are some towels are missing then it will instantly proceed to the customer service for assistant. Last, if there is zero missing the system will proceed to payment option.

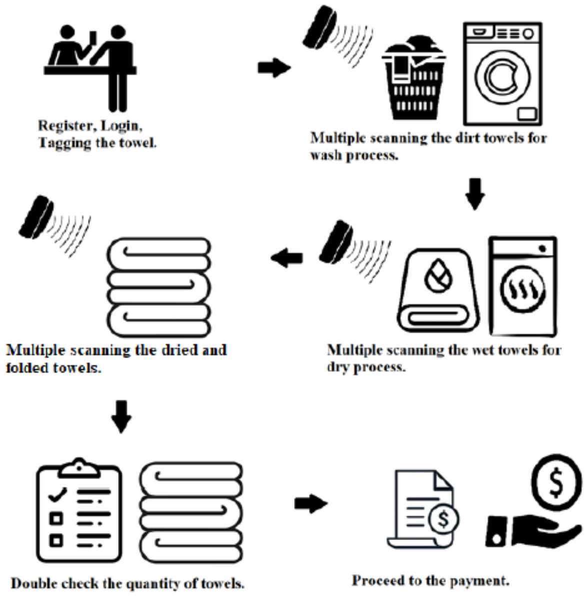


Fig. 1: Laundry Management System using Multiple RFID Reader

2. Methodology

Firstly, the main objective of this project to use the UHF RFID device for multiple scanning. A RFID device to two object which is the reader and the tags. The reader has its own antenna that's works by detecting any tagged items around a certain area typically three to four meters depending on the RFID frequency. The reader is connected to PCs to display and store the information received. The tags are the small components that was attached to the towels. Each tag comes with a specific ID which stored in the small chips. User can read and write the data from the tag's chip if they are within the reading range. In this project, UHF RFID system is used due to its wide reading range. This will benefit to the system because the users might need to detect hundreds of items at one time. Next, the system interface implemented and created to read and items passing through the RFID reader. The system interface created and designed by the Visual Basic language. The tags detected by the reader then it will send the data signal to the computer program for data processing. After data processing, the computer program will store the data to the database as a record. When the user wishes to check out the items then the computer will read the data from database and display out the result for the user.

2.1. *Hardware and Software Interface Integration*

The important items in this project are the tagged linen. Each individual items will be attached with the laundry tag which carry different identification number. The small chip installed in the RFID tag will store the unique ID which represent each individual items. The user may create their own serial number format to differentiate each item such as towels, bedsheet, pillowcase and more. Fig.2 shows the example of RFID tag placed on a towel. Laundry tag or textile tag are now available in the market with cost less then $2 each. They are reusable and safe to be use through the washing, drying, and ironing phase. However, there are some limitations of temperature and chemical to be used which may varied for different types of tag. This project will not further discuss on the durability of the textile or laundry tag but to focus on the implementation of this textile tag into the laundry management system. Therefore, the analysis of the tag durability will not be included in this paper.

Fig. 2: Textile RFID Tag placement on the towel

Next, the tagged linen will need to be visible to the reader to be identified. The placement of the reader needs to be accurately set up to ensure that all tagged items can be detected by the reader at the same time. Therefore, the selection of RFID reader is also important to be considered. In this project, the Mi-1802B RFID[4] reader with size of $440mm \times 440mm \times 50mm$ and average reading of $10ms$ per 64 bits is used. This reader is a long range type which can supports up to 6 meters reading range and supports multiple readers operation. The multiple reader is required in this project to differentiate each stage of linen processing area.

The third item which allows the system to well operated is the software or apps which was computed to read the data from the reader and transform it into a readable data to be stored into the database for further analysis or process by the user. In this project, the application was build using Microsoft Visual Studio with Microsoft Access integration to manage the data received by the reader. In this project, a simple interface system was built to project the number of items detected by the reader from each station. The system also will be able to display the missing items by comparing the detected linen from at the final stage to the initial stage.

Fig. 3 shows the example of missing linens calculated based on the initial number count of 114 linens from the washing and drying stage. At this point, user can start to investigate for the missing linens.

Fig. 3: Linen detection at final stage

This interface is to check the quantity of the towels if the number of the towels are not missing then the user can proceed to the payment option. There are four buttons for this interface. The first button is the 'Get Data', it used to obtain the data from 'Washing', 'Drying' and 'Folding' interfaces. In the same way, the 'Insert' button is to save and record the data to the database. Else, the data of the real date and time and the number of washed, dried and folded towel will record and save into the database once the user click the button. Moreover, the 'Compare' button is to display and calculate the data of the washed, dried and folded towels. The software interface also allows the user to check the status for each stages to see the current number of item being wash, dry or fold accordingly as shown in Fig. 4.

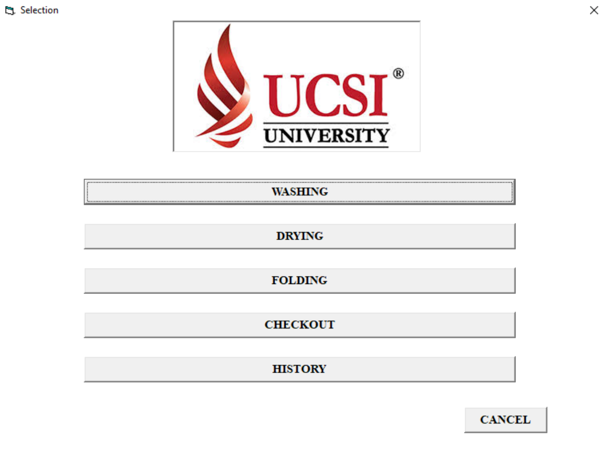


Fig. 4: Selection of individual stage of linen process

When the linens passed through the reader form one section to another, the data will be automatically recorded. The scanning time usually takes 3-5 seconds to scan 100 towels. The display will display out the number of units, scanned time and the identity of the RFID tag. The 'Tag Count' will count the number of the tagged towels. The counter in the software interface has been programme to count the items with the same ID as one. Therefore, duplication of linen counts which may lead to false count can be avoided.

3. Reader Setup Analysis

The unique feature of contactless detection between the RFID reader and tagged items makes RFID as a very useful system. In addition, the ability to detect multiple items in one second proved that this RFID system can provide better service compared to single item scanning system. However, the reading distance and angle are depending on the surrounding and tagged items. Reflector object such as metal or absorber material such as water are not a good friend to RFID system. Both types of material can degrade the performance of the reader[5]. The scanning angles played the significant role in this system. Different angle will have different reading distance. In this project, there are three different angles tested which are the top corner, top and left scanning angle as shown in Fig. 5. In this testing, the UHF RFID need to scan 100 unit of towels with different angles and compare the result. The washable RFID tags will be placed at the corner of the towels. The 1420mm x 420mm x 1190mm size trolley was filled with 100 units towels.

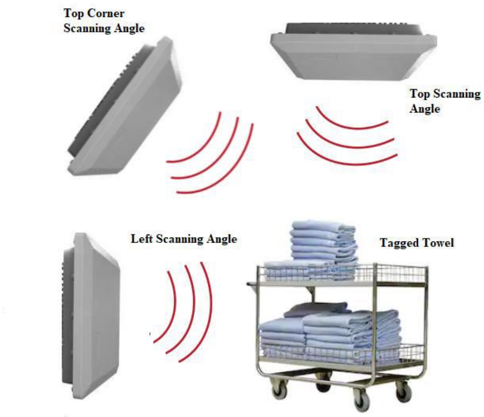


Fig.5: Towel scanning at three different angles

Table 1 shows the result of scanning at three different angles. The best detection was measured when the reader is place on above the trolley at a distance of 700 mm with detection rate of 99-100% while the left and top corner angle gives lower detection rate of 96-98%. The failed detection might be due to the position of the towel on the trolley, speed of the trolley while passing the reader or the thickness of the towel to reach each layers of towel.

Table 1: Analysis of detection rate for 100 towels at three different angles.

Type of Scanning Angle	First Scanning	Second Scanning	Third Scanning	Overall Rate (%)
Top Corner	95	96	98	96.33
Left	97	97	96	96.67
Top	99	100	99	99.33

4. Conclusion

In the conclusion, the simple and efficient software and hardware integration for the laundry management system using RFID application has been developed. The database included in the software interface can be further utilize for advance purpose such as to generate the invoice for the linen owner. The stability of the linen detection has been conducted and the reader detection from the top of the stacked linen is the best setup. This laundry management system using RFID system presented in this paper can also be use at the hotel to prevent any lost of the towels or others valuable items. For laundry application, this system can be further improved by developing the mobile apps for the user and the linen owner to monitor the linens cleaning process.

5. References

1. Kanazawa, K., Kazuno, S., Shiozawa, Y., Isshiki, M., & Okumura, M. (2018). "Development of laundry-work assistance robot by using IoT technology". 2018 IEEE International Conference on Consumer Electronics, ICCE 2018, 2018.
2. Lu, Y., & Yu, H. (2010). A flexible architecture for RFID based laundry management systems. 2010 6th International Conference on Wireless Communications, Networking and Mobile Computing, WiCOM 2010.
3. Hojoh, T., Hata, S., Hayashi, J., Hamada, T., & Hojoh, H. (2011). Towel mark recognition system for washing cloth handling system. 2011 IEEE International Conference on Mechatronics and Automation, ICMA 2011, 664–669.
4. Mi-1802B Datasheet
5. Abu, M., E. E. Hussin, A. R. Othman, Fauzi Johar, Norhidayah M. Yatim, and Rose F. Munawar. "Design Of 0.92 Ghz Artificial Magnetic Conductor for Metal Object Detection in RFID Tag Application with Little Sensitivity to Incidence of Angle." Journal of Theoretical & Applied Information Technology 60, no. 2 (2014).

Authors Introduction

Ms. Eryana Eiyda Hussin

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. She is also a lecturer from the Department of Electrical and Electronics of UCSI University, Malaysia.

Mr. Wong Jee Jian

He received his B.Eng Degree in 2020 from the Faculty of Engineering, Technology, and Built Environment, UCSI University in Malaysia.

Dr. Norsuzlin Sahar

She received a PhD degree for research in reconfigurable antenna using RF MEMs switches for RFID and GPS applications from the Universiti Kebangsaan Malaysia (UKM) in 2016. She currently works as the researcher at the Institute of Climate Change in UKM. Her specialization in communication antenna design, satellites antennas, electromagnetic radiation analysis and microwave device for wireless applications.

Ms. Eryana Eiyda Hussin

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. She is also a lecturer from the Department of Electrical and Electronics of UCSI University, Malaysia.

Mr. Wong Jee Jian

He received his B.Eng Degree in 2020 from the Faculty of Engineering, Technology, and Built Environment, UCSI University in Malaysia.

Dr. Norsuzlin Sahar

She received a PhD degree for research in reconfigurable antenna using RF MEMs switches for RFID and GPS applications from the Universiti Kebangsaan Malaysia (UKM) in 2016. She currently works as the researcher at the Institute of Climate Change in UKM. Her specialization in communication antenna design, satellites antennas, electromagnetic radiation analysis and microwave device for wireless applications.

Monitoring System with Humidity and Growth Level Detection for Horticulture

Eryana Hussin*
Electrical and Electronics Department, UCSI University, Jalan Puncak Menara Gading, Taman Connaught, Cheras, 56000, Wilayah Persekutuan Kuala Lumpur, Malaysia†

Ng Joon Wen[1], Norsuzlin Sahar[2], Azman Zakariya[3], Farah Adilah[1], Hungyang Leong[1], Rizon Mohamed Juhari[1],
[1]Electrical and Electronics Department, UCSI University, Jalan Puncak Menara Gading, Taman Connaught, Cheras, 56000, Wilayah Persekutuan Kuala Lumpur, Malaysia
[2]Space Science Centre (ANGKASA), Institute of Climate Change,Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia
[3]Dept. of Electrical and Electronics Engineering, Universiti Teknologi PETRONAS, Persiaran UTP, 32610 Seri Iskandar, Perak Malaysia

E-mail: eryanaeiyda@ucsiuniversity.edu.my, 1001437370@.ucsiuniversity.edu.my
www.ucsiuniversity.edu.my

Abstract

Smart Farming System with humidity and growth level detection presented in this paper is designed to monitor and control the production of the fresh produce. This smart farming system is designed to monitor the changes due to environment parameters which includes temperature, soil moisture level of the crops, and growth level of the crops. A sensor is placed at certain distance from the plant and the growth rate of the plant is monitored from time to time. Any plant which did not show positive or consistent growth rate will be eliminated and replace with new seeds. Therefore, the farmer can produce more productive plan in the estimated time. The system allows the user to choose the type of soils used for the plantation. The specific temperature and humidity level have been stored in the system and will be automatically set based on the soil selected. The humidity sensor included in this system can benefit the farmer by constantly monitor the humidity level of the soils from time to time before he can continue with the watering the crops. When the soil moisture humidity sensor detects the percentage of the soil is lower than the value that was pre-set in the system, the water pump will be trigger and start to pump water to the crop. While the water pump started to supply the water, the humidity sensor will continue to measure the sold humidity. The water pump will stop after the soil moisture humidity level reached the healthy level for the crop to growth. An automatic cooling System is a system that used to maintain and monitor the temperature in the greenhouse or plantation area. For sensitive crops such the temperature monitoring is very important to ensure the growth and production of the produce. The Smart Farming System proposed in this paper is built with a combination management system including environment parameters monitoring system and automatic technology control system. Arduino UNO is the central unit that plays an important role to receive and control all the function of the project. This combined system is design for modern smart horticultural or large-scale horticultural in order to produce more sustainable production.

Keywords: Smart Laundry, RFID system, Visual Basic, Automatic Control

1. Introduction

Horticultural is referring to the evolution of sustainable of high-value and intensively cultivated food and plants. In Malaysia the government has established the statutory body such as Federal Land Development Authority, FELDA and Federal Land Consolidation and Rehabilitation Authority, FELCA which is in charge of developing the rural areas by allowing rural communities to participate in economic activities to create income.[1]

Looking at the potentials of farming and plantation in Malaysia, many companies have invested in agricultural technology. Large scale plantations such as fruits[3-4], vegetables[5-6], and trees[7] contribute to the raise of company revenue. Cooperating the use of technologies has brought up the agricultural system into a whole new level. The use of Internet of Things, IOT integrated system in agriculture[8-9] have given good impact on the production of agricultural system as well as promoting appropriate agricultural management system. Thus, agricultural is no longer viewed as low-income employment but requires a professional management system that corporate with high-end technology especially for large scale production. There are several challenges of applying the IOT based system as presented in Ref. 10 which includes the high-power consumptions, high cost for the networking and maintenance, complicated infrastructure and the lack of IOT knowledge among the farmers in rural areas.

In this paper, a simple monitoring system for humidity and growth level detection is presented. The focus in the paper is to develop a system which can measure the suitable crops for different plants and to ensure the humidity of the soils is well maintained to ensure effective growth of the crops. Environment parameters including air temperature, soil moisture humidity, and growth level detection can affect the growth condition of the crops. The scope of this project is to develop the smart farming system including the sensor to detect the humidity of the soil and measure the growth level of the crops based on the different type of soil. Apart from that, this project able to detect the temperature of the horticulture to make sure the temperature of the horticulture always stays in the stable range. The sensors used in the project are exploited to detect the temperature of the environment, growth level, and soil moisture humidity level of the crop. Fig. 1 shows the connection of the sensors, Arduino and other relevant components included in this project.

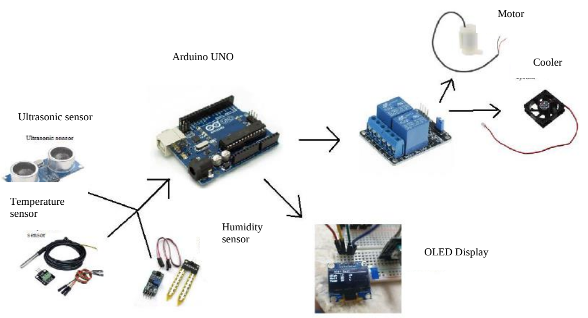


Fig. 1: Connection of sensors and components to the Arduino

2. Methodology

The Monitoring System with Humidity and Growth Level Detection for Horticulture presented in this paper is designed to help the framers to ensure healthier growth and to increase the produce of the crops. The system started by asking the user to identify the types of soils to be used in the plantation. There are three types of soils identified[11] and the example of suitable crops for each soil is presented in Table 1.

Table 1: Suitable crops based on different types of soils[11]

Types of Soil	Type1: Clay Soil	Type2: Sandy Soil	Type3: Silt Soil
Example of crops	Lettuce Onions	Cabbages Radishes	Turnips Spring Onions

2.1. *Soil Moisture Calculation*

There are six steps to calculate the moisture humidity for different type of soils as presented by Ref. 11. Below are the 6 steps to calculate the soil moisture (%).

1. Identify the length of the root depth for a mature crop This parameter is important to ensure sufficient area for the root to expend and grow and to ensure the root able to absorb the water and nutrient from soils.
2. Design and calculate the volume for plant the crops. It is very important to have a sufficient room and area for the crops to expand. Too small area may lead to shortage or immature plant growth.
3. Identify the soil core dry and wet weight
   This factor is to measure the soil core dry and wet weigh to calculate the water mass base on different types of soil.
4. Calculate the water mass using Eq. (1)
5. Calculate the Volumetric Water Content, VWC using Eq. (2)
6. Calculate the soil moisture using Eq. (3)

$$Water\ Mass = Soil\ Core\ Wet\ Weight - Soil\ Core\ Dry\ Weight \quad (1)$$

$$Volumetric\ Water\ Content, VWC = \frac{Core\ DryWeight}{Core\ Volume} \quad (2)$$

$$Soil\ Moisture(\%) = VWC\ \times 100\% \quad (3)$$

In this project, the prototype was tested using the acrylic box which was built to represent the underground soils with a size of $200\ mm\ \times 200\ mm$ base. Fig. 2 shows the acrylic box to mimic the height of the soils with height of $500\ mm$ and filled with the sandy soils.

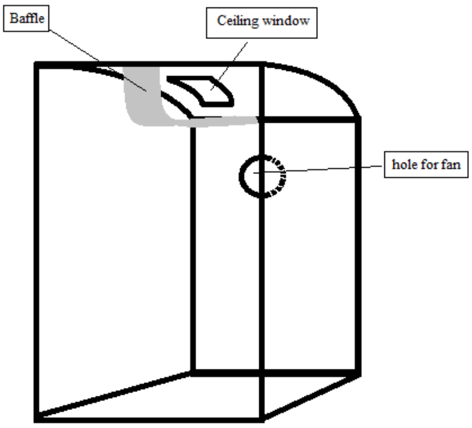


Fig. 2: The Acrylic Box with height 500 mm

The weight of an empty acrylic box is measured at 1.5kg. Next, the weight of the box with three different soils for two different types (wet and dry) is measure and presented in Table 2.

Table 2: Weight of wet and dry soils.

	Sand	Slits	Clay
Wet	35	20	26
Dry	30	15.5	25

Then, by using the Eq. (1) – (3), the humidity percentage is calculated and presented in Table 3.

Table 3: Humidity percentage of three different soils.

Sand	Slits	Clay
27%	24%	5.4%

2.2. *Hardware Integration*

Fig. 3 shows the schematic diagram of the project which includes three different sensors connected to the Arduino UNO; temperature sensor, ultrasonic sensor, and soil moisture humidity sensor. Each of the sensor have specific function. For ultrasonic sensor which represented growth level detection to detect and show the high of the crops in OLED. Soil moisture humidity sensor is to detect the soil moisture and have the function of smart watering system. The smart watering system will be break by using case coding which allow the user to select the type of soil that had set the specific value. Temperature sensor is to detect the room temperature which have the function of smart cooling system. When the data that the temperature send to Arduino UNO reach the minimum temperature the smart cooling system will started to trigger. If the room temperature keep increasing and over maximum temperature the alarm system will trigger and alert the user.

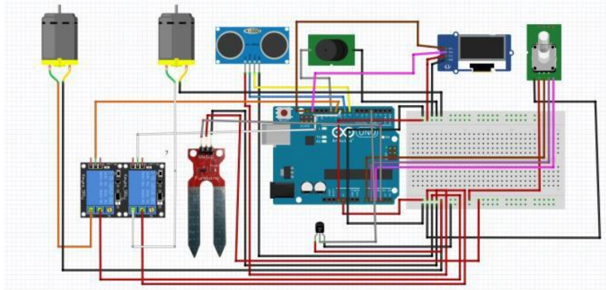


Fig. 3: Schematic Connection

Smart Farming System with Humidity and Growth Level Detection presented in this paper is a system that is designed to increase the work efficiency of the farmers. The prototype of the project have total three sections. The first section is allow the user to select which type of soil are used to plant the crops. There will be total three different type of soil allow the user to select which is sandy soil, silts soil, and clay soils. After selecting the soil, the OLED will show the environment parameters of the horticultural. growth level detection, soil moisture level, and room temperature are the three environment parameters of the project. Each of the environment parameters have play an important role to make sure the horticultural stay in the best performance. Ultrasonic sensor represented the growth level detection function to show the high of the crops. Automatic technology system of smart watering system and smart cooling system is installed in this prototype. Temperature sensor represented the room temperature have the function of smart cooling system when the temperature inside the horticultural is un-normal, the smart cooling system will trigger. Furthermore, if the temperature is keep increasing when the smart cooling system is triggered, there will be an alarm system to alert the farmers. Fig. 4 show the flow chart of the Smart Farming System with Humidity and Growth Level Detection system.

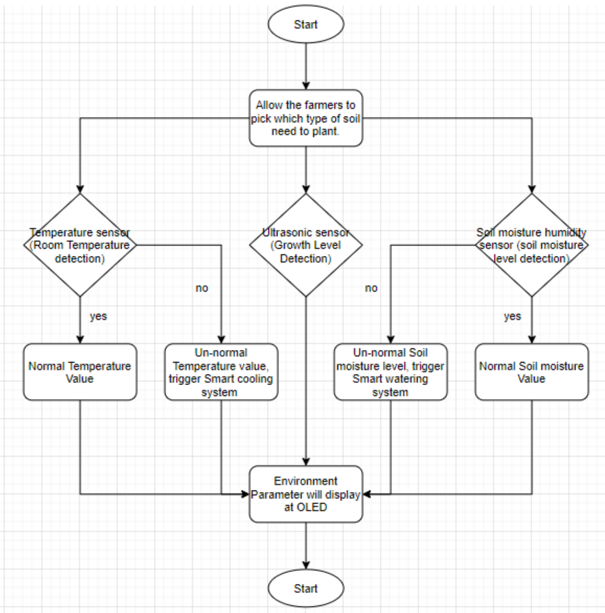


Fig. 4: Flow chart of the Smart Farming System with Humidity and Growth Level Detection system.

3. Result and Discussion

The system will display the LOGO and the Tittle of the project delay for 2 second(s). After the LOGO and Tittle display, the coming OLED will display is the main menu. Which show in Fig. 5.

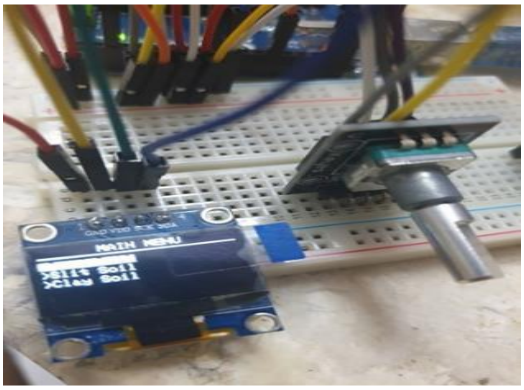


Fig. 5: Menu selection for soils

Sandy soil, slits soil, and clay soil will be display and allow the farmers to do selection. Rotary encoder is the switch that allow the farmers to turn clockwise and anti-clockwise. Middle button of the rotary encoder is enter and exit button. Each of the selection will show three environment parameters which is high, soil, and temp which label with the type of soil that the farmers had selected. Three of the soil will show three of the environment parameters. The only different is the soil moisture value. Each of the soil have the specific value which set in the coding by using break case.

3.1. *Measurement of Growth Level Detection*

Growth level detection is one of the rare facilities to design a horticultural. The system of growth level detection is rare in research paper. This part will discuss the important of the growth level detection for the horticultural design. Figure 14 is the experiment to prove the important and the purpose of the growth level detection. The experiment is two crops is plant at the same time and both crops have the ultrasonic sensor to observe the high of both plant. Crops A have the function of smart watering system which will trigger when the soil moisture humidity sensor detect the soil moisture level for the crops is too low. Crops B will be watering every morning 10am and 10pm. The results will be record every morning.

Table 4: Growth Record Of The Crop A and Crop B

Week	Crop A (automatic watering system)	Crop B (10am and 10pm)
1	0cm	0cm
2	0.8cm	0cm
3	1.5cm	0.8cm
4	2cm	1cm
5	4.3cm	1.8cm
6	5cm	2.3cm
7	5.2cm	3.0cm
8	5.5cm	3.3cm
9	5.8cm	3.5cm
10	6cm	3.5cm
11	6.2cm	3.5cm
12	6.4cm	0cm

The Experiment stop at day 12 because Crops B withered because the lack of water. Furthermore, the growth of the Crop B is slower compare with Crop A base on the Table 4. Experiment above clearly show that the important and the purpose of the growth level detection. When the farmers find out that the growth level of the crops is abnormal, the farmers able to fertilize the crops or replant immediately to increase the work efficiency.

4. Conclusion and recommendations

This research paper presented a Smart Farming System with Humidity and Growth Level Detection which including the smart cooling and smart watering system. Environment parameters including air temperature, soil moisture, and room temperature were showed in the OLED. The soil moisture for different type of soil able be calculated based on the root depth of the crops. The Internet of Things (IoT) system able to add in this project which allow the farmers to view the parameters by using the IoT system. Furthermore, pH level detection able to add in this project by using the pH level sensor. The demonstration of this project is available in the link: https://www.youtube.com/watch?v=GSFXapTjee0

References

1. "FELDA website, "
2. "FELCRA website,"
3. Seng, W.C. and Mirisaee, S.H., 2009, August. A new method for fruits recognition system. In *2009 International conference on electrical engineering and informatics* (Vol. 1, pp. 130-134).
4. Zydlik Z, Cieśliński S, Kafkas NE, Morkunas I. Soil Preparation, Running Highbush Blueberry (Vaccinium corymbosum L.) Plantation and Biological Properties of Fruits. InModern Fruit Industry 2019 Sep 18.
5. X. Jin, J. Che and Y. Chen, "Weed Identification Using Deep Learning and Image Processing in Vegetable Plantation," in IEEE Access, vol. 9, pp. 10940-10950, 202
6. Singh SK, Singh SK, Singh S. Vegetable crops as most efficient and economical intercrops: a brief review. Plant Archives. 2018;18(1):923-9.
7. Piiroinen T, Nyeko P, Roininen H. Natural establishment of indigenous trees under planted nuclei: a study from a clear-felled pine plantation in an afrotropical rain forest. Forest Ecology and Management. 2015 Jun 1;345:21-8.
8. R. A. Kjellby, L. R. Cenkeramaddi, A. Frøytlog, B. B. Lozano, J. Soumya and M. Bhange, "Long-range & Self-powered IoT Devices for Agriculture & Aquaponics Based on Multi-hop Topology," 2019 IEEE 5th World Forum on Internet of Things (WF-IoT), 2019, pp. 545-549, doi: 10.1109/WF-IoT.2019.8767196.
9. Ji-chun Zhao, Jun-feng Zhang, Yu Feng and Jian-xin Guo, "The study and application of the IOT technology in agriculture," 2010 3rd International Conference on Computer Science and Information Technology, 2010, pp. 462-465.
10. M. R. M. Kassim, "IoT Applications in Smart Agriculture: Issues and Challenges," 2020 IEEE Conference on Open Systems (ICOS), 2020, pp. 19-24.
11. Alhammadi, Mohamed S., and Ali M. Al-Shrouf. "Irrigation of sandy soils, basics and scheduling." Crop Production. IntechOpen, 2013.

12. Authors Introduction

Ms. Eryana Eiyda Hussin

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. She is also a lecturer from the Department of Electrical and Electronics of UCSI University, Malaysia.

Mr. Ng Joon Wen

He graduated from UCSI University, with Bach Degree, Electrical and Electronic Engineer in 2020. Before interning at I-Gen Resources he was at Carrier (Puchong) working in Engineering Control Department. He is a Project Engineer at I-Gen Resources which provide Fire Protection System for fire protection solution and safety beyond customer. Project Engineer evolved full-service design, engineering, construction management, services including Fire Alarm System, Hose Reel System, Wet Riser System and Sprinkler System.

Dr. Norsuzlin Sahar

She received a PhD degree for research in reconfigurable antenna using RF MEMs switches for RFID and GPS applications from the Universiti Kebangsaan Malaysia (UKM) in 2016. She currently works as the researcher at the Institute of Climate Change in UKM. Her specialization in communication antenna design, satellites antennas, electromagnetic radiation analysis and microwave device for wireless applications.

Double Identification Attendance System using HF RFID and Sensors

Eryana Hussin*

Electrical and Electronics Department, UCSI University, Jalan Puncak Menara Gading, Taman Connaught, Cheras, 56000, Wilayah Persekutuan Kuala Lumpur, Malaysa†

Wong Chee Ming[1], Norsuzlin Sahar[2], Azman Zakariya[3], Salihah Surol[4], Ruzaimah Razman[4], Rizon Mohamed Juhari[1]

[1]Electrical and Electronics Department, UCSI University, Jalan Puncak Menara Gading, Taman Connaught, Cheras, 56000, Wilayah Persekutuan Kuala Lumpur, Malaysia

[2]Space Science Centre (ANGKASA), Institute of Climate Change,Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

[3]Dept. of Electrical and Electronics Engineering, Universiti Teknologi PETRONAS, Persiaran UTP, 32610 Seri Iskandar, Perak Malaysia

[4]Civil Department, UCSI University, Jalan Puncak Menara Gading, Taman Connaught, Cheras, 56000, Wilayah Persekutuan Kuala Lumpur, Malaysia

E-mail: eryanaeiyda@ucsiuniversity.edu.my, 1001438605@.ucsiuniversity.edu.my

www.ucsiuniversity.edu.my

Abstract

In this paper, the double identification attendance system integrated with High Frequency Radio Frequency Identification (RFID) and sensors is presented. This double identification system requires users to be detected by two different mechanisms. The first mechanism is by using the RFID reader identification. The users need to place the RFID card on the reader for the system to capture the user details such as name, time in and date. Once the user passed the RFID reader, they will need to be detected by the motion sensors within few seconds in order for their attendance to be captured by the system. This double identification system is proposed to overcome the problem of false attendance by the students who just stopped by to tag their student identification card without attending the class. The attendance system presented in this paper is able to capture the time in and time out of each student which will be automatically captured if the students successfully pass through both RFID reader and motion sensor. The proposed attendance system is designed to replace the manual attendance system which still implemented in this modern era. The unique features of the RFID system which able to capture and store the student attendance in the database system effectively and can reduce the time for the teachers to take the attendance daily. The Graphical User Interface (GUI) of the attendance system is developed by Microsoft Visual Studio and the database of the system using MySQL. The GUI built in this system is used to project the ID and detection status for each student. The GUI also can be used by the lecturer to monitor the student attendance based on the databased captured by the system. This double identification attendance system also will be useful for any event or training provider to record the attendance of the attendant effectively.

Keywords: Radio Frequency Identification system, attendance system, motion sensor, Visual Studio

1. Introduction

Attendance recording is an essential task in event which involve large participant number. It is common things to be included as part of the teaching and learning record in all schools and universities. Attendance system helps the event organizers or teaching to keep a record of the students or attendees according to the dates, time, and classroom location. In most of the educational institutions in Malaysia, the problem faced by the lecturer in taking daily attendance of students would be the difficulty to manage and monitor the attendance for large group of students. There are various of automatic identification have been develop such as Facial Recognitio[1-2], Image Augmentation Technique[3], finger print[4] and temperature sensors[5]. However, these techniques usually require a lot of time for the system to recognize and process the data before it can be successfully recorded into the system.

Ref. 6 has presented the Radio Frequency Identification (RFID) Based Attendance System" which accommodate the RFID technology to complete the system using a real time clock to improve the accuracy of the attendance record. Another similar system presented in Ref. 7 to monitor the attendance of the employee with RFID tags to communicate with an Arduino UNO connecting RFID Reader/Writer module. However, there are some limitations of using a single identification card because this can lead to unauthorize people to misused other person card to gain access or to false attendance record. In other words, the identification of a unique identification based on a single card scanning is still not effective for attendance monitoring system due to the lack the ability to distinguish between different users. Therefore, it is still common to practice the traditional method of calling the name of students one by one due to several case of false attendance when the auto-scanning attendance is applied[8].

A survey was made by taking 52 first year B.Tech students of School of Engineering, Ajeenkya DY Patil University, Pune to calculate the average time taken by a faculty to take attendance through dedicated Enterprise Resource Planning (ERP). The study revealed that each faculty took approximately 8-10 minutes to complete the attendance of students, which accounts 20% of the assigned class time[9]. In this project, the RFID system implemented is a High Frequency (HF) RFID, with the operating frequency at 13.56 MHz. The maximum read range can be reached up to 1m. The movement of student entering to class can be detected by implementing passive infrared (PIR) motion sensor to the system, to ensure the attendance recorded is reliable. The objective of this project is to develop an attendance system that can capture time in, time out and measure the duration time of each student in class, and integrate the system with HF RFID and sensors at the same time.

2. Methodology

2.1. *Hardware Setup*

Based on Fig. 1, the attendance system proposed in this paper is consist of five important components; the RFID system, motion sensor system, database and graphical user interface (GUI) and the host computer.

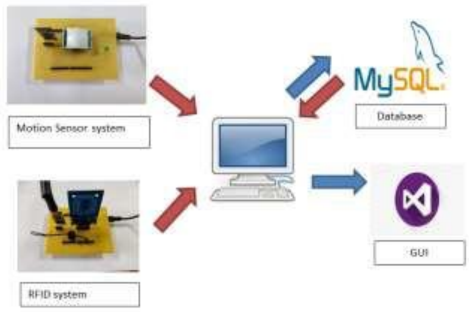


Fig 1 Block diagram of the attendance system.

The RFID system consists of RC522 RFID reader module, RFID tag and a buzzer, while a LED is added to motion sensor system for indication purpose. The GUI in this project is developed using Microsoft Visual Studio, the purpose of developing a GUI is to enhance the user experience and achieve user-friendly interface. It is the interface between hardware and the user. For database implemented in this project, it is developed with MySQL. The purpose of database in the attendance system is to store the student data and daily attendance record.

Next, a RFID reader will be placed far from the host computer in real life application, therefore, CC2530 Zigbee module is implemented into the system to allow wireless communication. As shown in Fig. 2, after the RFID system is initialized, it is ready to read the tag. If the ID of RFID tag read is stored in database, the attendance system will receive input from motion sensor for movement detected. If there is any movement, the attendance recorded will be "Detected" and "Not detected" if no movement detected. There will be a 5 seconds time delay after the tag is read during entrance, it allows the student to trigger the motion sensor during this period. In the case of the tag ID is not registered in datase, error message will be displayed on the interface.

The 5 second delay is apply to reduce the time for the next person to be detection to ensure smooth identification record to be performed.

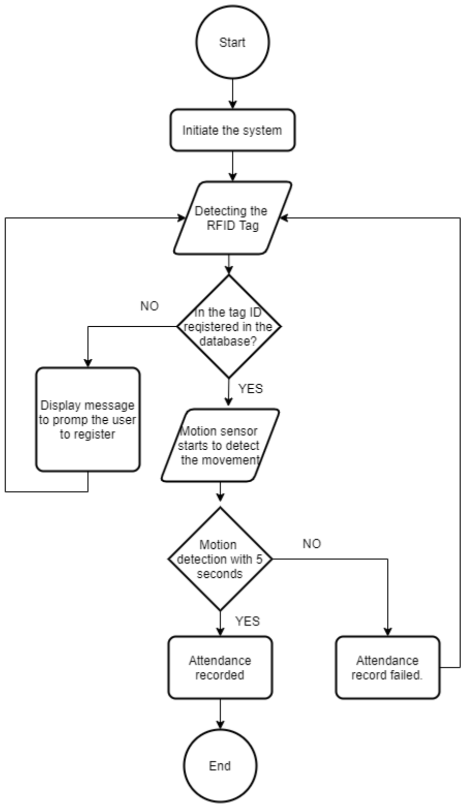


Figure 2: Flow chart of the double identification system using RFID and Motion Sensor.

2.2. ***Software Setup***

The GUI of the attendance system is developed with Visual Basic programming language in Microsoft Visual Studio. The database is created using MySQL. The development of database will be the crucial factor in attendance system as the student will be recognized by the data stored in database. The GUI is designed in simple form so that the user can adapt to it easily. In order to integrate with RFID reader and motion sensor, the GUI is set to receive serial data transmitted from the reader and sensor, and then compare the data with database and resulting in showing necessary information at the GUI. Fig. 3 shows the flowchart of software. The GUI will not work properly without connecting to database. There are two modes can be selected from the attendance system, admin mode and take attendance. For admin mode, the user will stay on the main windows and several actions can be done such as checking student data, registration or edit of student data, and check the attendance record. All these action can be done concurrently with the process of taking attendance, two branches all independent on each other. While in attendance recordin mode, a new windows of the GUI is set to receive input from RFID system and motion sensor system continuously. If the tag ID is not found in the system, an error message will be displayed. The attendance will be recorded after the GUI has received the data from motion sensor system.

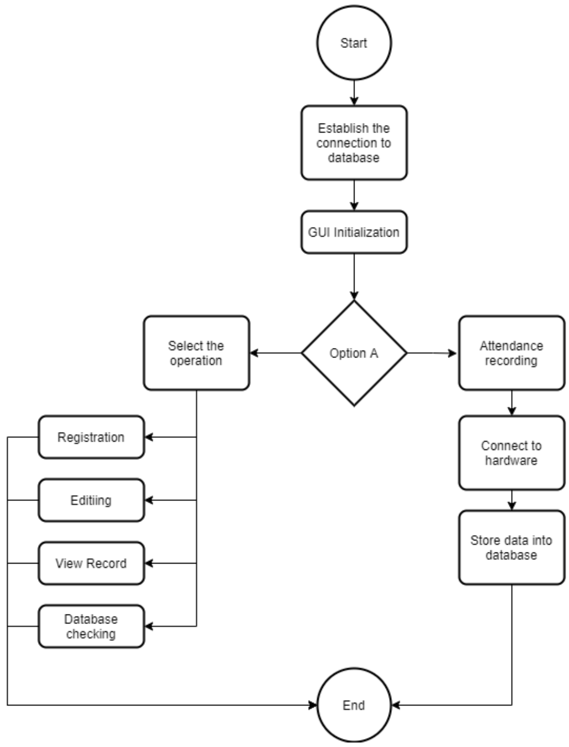


Fig. 3: Flow chart of the GUI

In Figure 3.13, the shows the procedure for the student to record their attendance, the size of room proposed is 8m x 6m x 3m. One of the RFID readers is installed at the wall at the entrance while another one is installed at the exit. There will be only one sensor system applied in this project to detect the movement of the student which is placed 30cm away from the RFID system at the entrance.

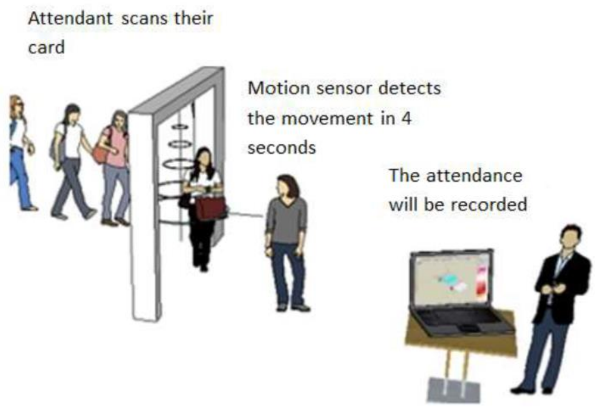


Fig.4 : The proposed setup of the attendance recording system using RFID reader and motion sensor.

3. Hardware and GUI Integration

Fig.5 shows two different prototypes included in this project; the High Frequency RFID system and the motion sensor system. Both of the hardware and the host computer will be connected to a CC2530 zigbee module to allow wireless serial communication. Fig. 5(a) shows the arrangement of microcontroller, Arduino UNO at the lower part of the printer circuit board (PCB) to minimize the size of 10cm x 8cm PCB. The same of connection of Arduino UNO to the PCB is also applied on the motion sensor system shows in Fig. 5(b).

(a) (b)

Fig. 5: (a) the arrangement of microcontroller, Arduino UNO and RFID reader and (b) motion sensor system on the Printed Circuit Board (PCB).

The development of the GUI plays animportant role as communication medium between the user and the hardware systems. It allows the user to perform several operations such as checking the student data, registering a new student or editing the data of the student, and recording the attendance.

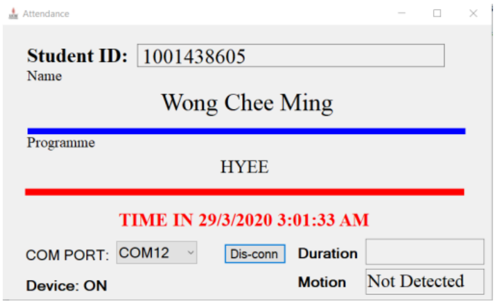
The first GUI page initiated by the login page to prevent unauthorized person from exploiting the database system. Then, the connection to the hardware will be done by selection the correct connection setting between the computer system and the hardware using the com port connection. Once the connection with hardware is established, the user now can choose the option based on the menu displayed on the Welcome page.

The main part of the attendance system would be taking the attendance and record it to the database. This action can be done by clicking on the “Take Attendance” button located on the left panel, a new window will appear after the button is clicked. The reason of adding com port selection in the “Take Attendance” interface is to allow the user to use main interface and “take attendance” interface at the same time, which is closer to real world application and user-friendly. For example, the user needs to take attendance while registering or updating the data of the student, it can be done by connecting two interfaces to different com port. In other word, the process of taking attendance will not interrupt other action if two com ports are connected.

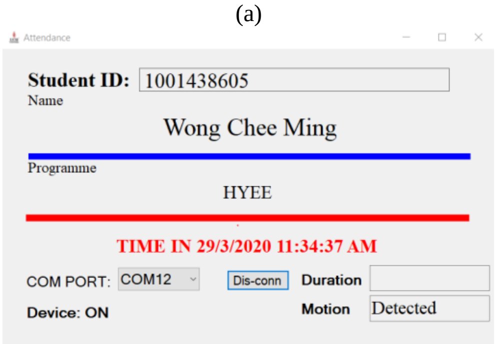
There are seven items to be displayed on the attendance recording page which includes; Student ID, Name, Programme, Time Log, Duration And Motion. The time Log will display the time where the student’s identification card is detected by the system. For two doors setup, the system can detect the time as time in and time out and the duration between the time in and time out will be calculated and displayed in the database. For the motion label, it indicate the successfull movement detection for the student. If the motion sensor did not detect the movement of the student within the 5 second delay, the motion label will displayed “Not detected” as shows in Fig. 6(a). All this information will be recorded into the database table for checking and reference purpose. In the case of movement detected, “Detected” will be displayed as shown in Fig.6(b). For time out, the result will be the same if no movement detected condition, and the time log and duration time will be displayed. The duration time of each student in class is displayed in the form of “HH:MM:SS”. For non-registered user, the name label will display “Not Registered” and the student ID and program will not be displayed. A prompt message will be pop-up for the student to proceed with the registration. Once the process of taking attendance is done, this connection can be closed and the computer will stops receiving data from the RFID reader. Table 1 shows the possible data of detection from the user based on the detection from RFID reader and motion sensor.

Table 1: Possible result for the attendance system.

RFID Reader	**Motion Sensor**	**Result**
Yes	Yes	Duration time and "Detected" will be recorded in the table.
Yes	No	Duration time and "Not Detected" will be recorded in the table.



(a)



(b)

(c)

Fig. 6: GUI output for (a)Not detected, (b) Detected, and (c) non-registered user.

The development of the GUI proposed in the paper allows the user to check the total duration of the user based on the Tagin and TagOut time. This special features in proposed in the project to ease the lecturer to check the total time that the student spend in the class. This feature can allow the lecturer to identify those students who left the class early then the required class time for a certain lecture class. This features also can be apply to any training provider to provide the prove of attendance to any employer who send their employee for a certified training. For some training which require a minimum hours (or days) for the participant to attend, this special features may be useful for them to monitor the participants effectively.

id	name	programme	date	time_in	time_out	duration	motion
1001438605	Wong Che...	HYEE	2020-03-28	12:51:49	12:51:57	0:0:8	NoDetected
1001438605	Wong Che...	HYEE	2020-03-28	12:52:10	12:52:19	0:0:9	Detected
1001438605	Wong Che...	HYEE	2020-03-28	15:08:48	15:08:54	0:0:6	NoDetected
1001438605	Wong Che...	HYEE	2020-03-28	15:09:07	15:09:13	0:0:6	Detected
1001437370	Ng Joon ...	HYEE	2020-03-28	15:14:28	15:14:35	0:0:7	NoDetected
1001438210	Len Han Z...	HYEE	2020-03-28	15:14:48	15:14:56	0:0:8	Detected
1001438351	Lee Zhen...	HYEE	2020-03-28	15:29:40	15:29:45	0:0:5	NoDetected
1001437370	Ng Joon ...	HYEE	2020-03-28	15:29:58	15:33:05	0:3:7	Detected
1001438351	Lee Zhen...	HYEE	2020-03-28	15:33:17	15:33:23	0:0:6	NoDetected
1001438351	Lee Zhen...	HYEE	2020-03-28	15:33:38	15:33:44	0:0:6	Detected
1001438605	Wong Che...	HYEE	2020-03-29	03:01:33	10:29:56	7:28:23	NoDetected
1001438605	Wong Che...	HYEE	2020-03-29	10:30:16	10:35:37	0:5:21	NoDetected
1001438605	Wong Che...	HYEE	2020-03-29	11:34:37	11:37:19	0:2:42	Detected
1001438605	Wong Che...	HYEE	2020-03-29	11:54:56			NoDetected

Fig. 7: Detection summary

4. Conclusion and Recommendations

As for conclusion, the double identification attendance system presented in the paper is a good initiative to replace the traditional attendance system in university. The unique function of capturing the time in and time out, and generating the duration of time of each student in class, which can highly reduce the number of fake attendance from students. In addition, this attendance system requires lesser effort to manage and monitor the attendance of the student especially for large group of students. The GUI develop in this project is easy to use even for the first-time user. The connection with the database system helps the user to store all the attendance for the whole semester into more strategic and effective data storage system. In conjunction with the sustainable development goals, this project has fully eliminated the use of papers for the attendance recording system. The rewritable RFID card, paperless and database system really support the SGD goals.

For the future recommendations, the notification function can be added to the system to alert the lecturer if there is student absent from the class or the absent rate of student is relatively high. In addition, the alert can be sent to the parents or guardians of the student so that they are able to monitor the activities of their child.

References

1. B. Gavriell, F. Fauzan, N. Ardian and K. M. Suryaningrum, "Implementation of Face Recognition Method for Attendance in Class," 2021 1st International Conference on Computer Science and Artificial Intelligence (ICCSAI), 2021, pp. 148-153,
2. M. -F. Tsai and M. -H. Li, "Attendance Monitoring System based on Artificial Intelligence Facial Recognition Technology," 2021 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW), 2021, pp. 1-2.
3. S. Gill, N. Sharma, C. Gupta and A. Samanta, "Attendance Management System Using Facial Recognition and Image Augmentation Technique," 2021 International Conference on Intelligent Technology, System and Service for Internet of Everything (ITSS-IoE), *2021*, pp. 1-6.
4. Arun Katara, Mr. Sudesh V. Kolhe, Mr. Amar P. Zilpe, Mr. Nikhil D. Bhele, Mr. Chetan J. Bele, "Attendance System Using Face Recognition and Class Monitoring System", IJRITCC, Feb. 2017, vol. 5, No. 2, pp. 273 - 276
5. K. Abhinav, R. Aneesh, P. James and A. Varghese, "Attendance Marking System using Periocular Recognition with Temperature Monitoring (ASPR)," 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2021.
6.] T.S. Lim, S.C. Sim and M.M. Mansor, "RFID Based Attendance System", IEEE Symposium on Industrial Electronics and Applications (ISIEA 2009), October 4-6, 2009, Kuala Lumpur, Malay-2009
7. K. Aravindhan, S. K. B. Sangeetha, K. Periyakaruppan, K. P. Keerthana, V. SanjayGiridhar and V. Shamaladevi, "Design of Attendance Monitoring System Using RFID," 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), 2021, pp. 1628-1631.
8. M. A. Meor Said et al., "Biometric attendance," 2014 International Symposium on Technology Management and Emerging Technologies, 2014, pp. 258-263.
9. Ghosh, Swamendu, et al. "Smart Attendance System" 2018 International Conference on Smart City and Emerging Technology (ICSCET). IEEE, 2018.

Authors Introduction

Ms. Eryana Eiyda Hussin

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. She is also a lecturer from the Department of Electrical and Electronics of UCSI University, Malaysia.

Mr. Wong Chee Ming

He received his B.Eng Degree in 2020 from the Faculty of Engineering, Technology, and Built Environment, UCSI University in Malaysia. He is working as test solution engineer, to develop and maintain functional test strategies and programs for NAND flash products.

Dr. Norsuzlin Sahar

She received a PhD degree for research in reconfigurable antenna using RF MEMs switches for RFID and GPS applications from the Universiti Kebangsaan Malaysia (UKM) in 2016. She currently works as the researcher at the Institute of Climate Change in UKM. Her specialization in communication antenna design, satellites antennas, electromagnetic radiation analysis and microwave device for wireless applications.

A Levitating Frictionless_Vertical Windmill

Ammar A.M. Al-Talib[1], Muhammad R. Md. Redzuan[1], A. R. Abd Hamid[1], A. R. Abd Hamid

[1]Mechanical Engineering Department, Faculty of Engineering, Technology and Built Environment, UCSI University, 56000, Cheras, Kuala Lumpur, Malaysia

ammart@ucsiuniversity.edu.my, razinredzuan@gmail.com, amar@ucsiuniversity.edu.my,

Abstract

To harness wind energy in low wind speed areas, a vertical axis wind turbine (VAWT) is generally more suitable than a horizontal axis wind turbine. To improve the feasibility of using a VAWT, a magnetic levitation bearing concept, utilizing neodymium magnets, is used to make the bearing frictionless. A new experimental type of airfoil called EN0005 is used that boasts better self-starting ability than conventional airfoils. The design parameters of the wind turbine are explored and reviewed to create a final design that is modelled in SolidWorks. This final design is then put through simulations and analysis in ANSYS FLUENT before being fabricated as a prototype. This final prototype produces similar results to other findings and validates the report. The EN0005 blade profile appears to improve the self-starting capability of the wind turbine and improve viability of Darrieus turbines in lower wind speeds. These findings verify the benefits of using a magnetic frictionless bearing for wind turbines.

Keywords: Vertical axis wind turbine; Magnetic bearings; Darrieus; Wind energy

1. Introduction

Climate change is one of the most pressing issues of the current generation. Since the beginning of the 21st century, it has been scientific consensus that human activity has been the major factor in climate change, however efforts and policies to reduce the effect of these activities have either made little or no impact[1]. In 2017, carbon dioxide emissions related to energy generation have increased by 1.4%, after 3 years of previously remaining stagnant. According to the International Energy Agency, generation of energy is one of the largest contributors to climate change, accounting for 80% of carbon dioxide emissions and two-thirds of greenhouse gas[2]. Therefore, to combat climate change, there must be a huge shift in the energy sector. This would involve searching for alternative energy sources that are renewable. One potential energy source that has not fully been explored is wind.

There are two configurations of vertical axis wind turbines, horizontal axis wind turbines (HAWT) and vertical axis wind turbines (VAWT). HAWTs have shafts that are parallel to the direction of wind flow and the ground, while VAWTs have shafts that are perpendicular to the wind flow and ground. The concept of HAWTs have been used since 5000BC and are more widely used in large scale operations, due to several factors. It has received much more funding and research than VAWTs and currently has a greater efficiency than VAWTs, due to its ability to face the wind and obtain maximum wind energy when facing the wind. However, HAWTs would not be able to work in areas that have inconsistent wind direction. VAWTs, on the other hand, are omni directional and perform better at lower wind speed areas but perform worse in high-speed areas. VAWTs can also backtrack as the blades also need to move toward the wind direction which will add resistance[3].

Compared to HAWTs, VAWTs are currently less researched and less developed as a technology[4]. It is generally known that VAWTs are more feasible for low wind speed areas with sporadic wind direction. However, VAWTs are still not widely used in areas with these conditions. With greater research and development, it could become the new energy solution for these areas.

Additionally, VAWTs can be utilized in rural areas not connected to power grids, due to its generally more compact design compared to HAWTs.

An interesting engineering problem of VAWTs that can be explored is the friction caused between the rotation of the turbines and the bearings on the shaft. A potential solution to this is using magnetic bearings to "levitate" the turbine to create a nearly frictionless connection between the rotor and stator. With close to zero friction, the feasibility of the VAWT could improve and lead to greater investment and research. That possibility, along with other design considerations, will be explored in this report.

2. Methodology and Experimental Setup

This research deals with the design, analysis, and fabrication of a levitating frictionless vertical axis wind turbine. SolidWorks, ANSYS and CFD analysis software's are used in the design. The final product has been fabricated as a prototype and been tested to verify its viability.

2.1. *Design Parameters*

The design parameters of the levitating frictionless VAWT encompasses many parameters[5], but for this research, the key parameters considered are shown in Table 1.

Table 1. Key Parameters

Category	Parameters	Fixed/Variable
Physical Features	1. Blade Shape (Airfoil)	Fixed
	2. Number of Blade (N)	Fixed
	3. Supporting Struts	Fixed
Dimensional	4. Swept Area (A=HD)	Fixed
	5. Solidity (Nc/R)	Fixed
	6. Aspect Ratio (H/c)	Fixed
Operational	7. Rated Power Output (Po)	Variable
	8. Rated Wind Speed (V∞d)	Variable
	9. Cut-in Speed (Vcut-in)	Variable
	10. Cut-out Speed (Vcut-out)	Fixed
	11. Power Coefficient (CPd)	Variable
	12. Tip Speed Ratio (λd)	Variable
	13. Rotational Speed (ωd)	Variable
	14. Pitching of Blade (γd)	Fixed
Others	15. Material	Fixed
	16. Magnet	Fixed

The optimal number of blades will differ depending on whether a Savonius or Darrieus turbine is to be chosen. As the literature review has shown, the optimal number of blades for Savonius turbines is two[6], while the optimal number for Darrieus turbines is generally agreed to be three[7, 8].

The selection of magnet is also a factor for the levitating frictionless wind turbine. The higher magnetic strength of the neodymium magnets will generate greater magnetic flux and thus, more electricity[9]. With these factors considered, neodymium magnets are chosen in this research.

2.2. *Design Selection*

The configuration in Fig. 1, is showing the stator as a stationary base of the turbine. The rotor, blades, and supporting struts are being held up by the pair of magnetic bearings in between the stator and rotor. As shown, the struts and rotor make no contact with the stator or the shaft, in the try of making it frictionless.

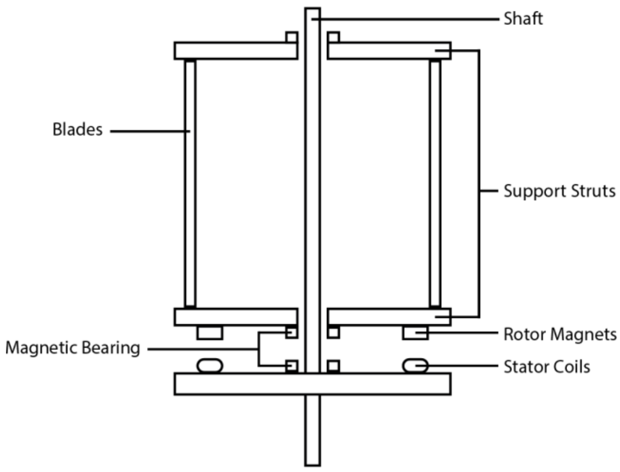


Fig. 1 Configuration of magnets, blades, rotors, stators, and shaft in design

2.2.1. *Blade Profile Selection*

Considering a Darrieus wind turbine, a selection is needed for a blade airfoil. The three airfoils that were considered were the NACA0018 airfoil, which is one of the most common airfoils used currently, and two experimental airfoils that are recently developed with the purpose of greater self- starting ability, MI-VAWT and EN0005[10, 11]. The strengths and weaknesses of these airfoils are presented in Table 2.

Table 2. Comparison of Blade Airfoils

	NACA0018	**MI-VAWT**	**EN0005**
Strengths	Simple to fabricate. Well studied and researched by many.	Good self-staring ability. Structurally strong.	Best self-starting ability. Performs at low speeds and high speeds
Weaknesses	Weakest self-starting ability out of all the choices.	More difficult to fabricate. Relatively little research.	More difficult to fabricate. Less research than NACA0018, but greater amount of research than MI-VAWT.

MI-VAWT can perform well, but has less self-starting ability, and it has a more similar shape to NACA0018, so it may be difficult to precisely fabricate it. However, it does have a simpler shape than EN0005, therefore it will be less difficult to fabricate compared to EN0005. Despite this, EN0005 airfoil has superior self-starting ability and performance as EN0005 can self-start at speeds of 1.25m/s[12], which would be around the average wind speeds of Malaysia[13]. So, it is the most suitable airfoil for low wind-speed areas. Therefore, the EN0005 blade profile, as shown in Fig. 2, has been selected.

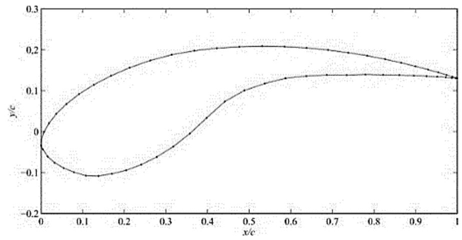


Fig. 2. Blade profile for EN0005

2.2.2. *Dimensional Calculations*

According to the Betz's Limit, the maximum power coefficient is 0.593, but the maximum practical limit is currently around 0.45. For the purposes of this design, a greater wind speed of 6m/s may be used to design the turbine to test the concept of the frictionless levitation. Before calculating the dimensions of the turbine, the values of solidity and aspect ratio must first be fixed. The aspect ratio will be greater than or equal to 7.5[14, 15]. Solidity is preferred to be lower to have greater self-starting capability. Due to this, for the design, solidity will be fixed at the value of 0.5, as several papers have aimed for low solidity when prioritizing self-starting ability and some have values near to 0.5[12]. For purposes of calculating the design, 20 watts of power will be the target power for this small-scale turbine.

To find the optimal swept area of the wind turbine, it is necessary to evaluate the coefficient of power using Eq. (1). The max *Cp* that is in practical use is 0.45. Although, it is to be expected that the *Cp* will not reach this maximum. The turbine designed in this research is considering this value of *Cp* in mind.

$$Coefficient\ of\ power,\ C_p = \frac{P}{P_w} = \frac{P}{\frac{1}{2}\rho A V_\infty^3} \qquad (1)$$

Where,
ρ is density of air
A is the swept area of the turbine
$V\infty$ is the velocity of wind
P is the power generated by the turbine

$$C_p = \frac{P}{P_w} = \frac{20}{\frac{1}{2}\left(1.225\frac{kg}{m^3}\right)(6m/s)^3 A} = 0.45$$

A = 0.336 m²

With the value of area determined and the equations of solidity and aspect ratio, the values of diameter, height, and chord length of the turbine are found. These calculations use Eq. (2), Eq. (3), and Eq. (4) [16, 17]. The dimensions found will be the dimensions of the turbine that will be tested.

$$\boldsymbol{Solidity, \sigma = \frac{Nc}{2R},\ where\ N = 3} \tag{2}$$

Where,
N is the number of blades
c is the chord length of the blade H is the height of the blade
R is the radius of the turbine

$\sigma = 1.5\mathrm{c}/\mathrm{R} = 0.5$
Thus, $\mathbf{R = 3c}$

$$\text{Aspect ratio, } a = \mathrm{H}/\mathrm{c} = 7.5 \tag{3}$$
Thus, $\mathbf{H = 7.5c}$

$$\textit{Area},\ \mathrm{A} = 2\mathrm{RH} \tag{4}$$
$A = 2(3c)\,(7.5c)$
Thus, $\boldsymbol{A = 45c^2}$

$\therefore\ 45c^2 = 0.336m^2$

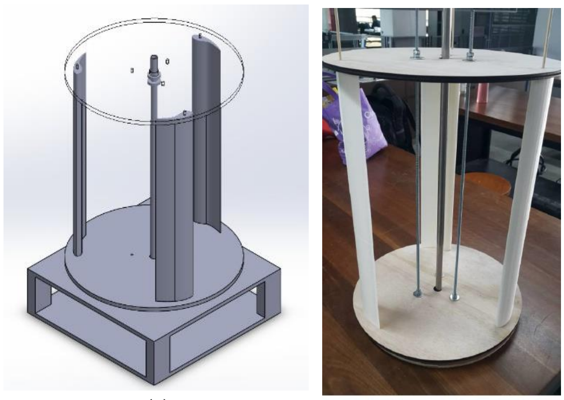
$\therefore\ \boldsymbol{Chord\ length, c} = 0.0864m = \boldsymbol{8.64cm}$
$\therefore\ \boldsymbol{Height, H} = 7.5c = \boldsymbol{64.8cm}$
$\therefore\ \boldsymbol{Diameter, D} = 2R = 6c = \boldsymbol{51.84cm}$

2.2.3 *Final Design*

The final parameters and design selected can be seen in Table 3 and Fig. 3 (a) and (b). The design selected is a Darrieus wind turbine with the EN0005 airfoil. The first reason is that the airfoil EN0005 has a better self-starting ability than the NACA0018 airfoil. Secondly, Darrieus wind turbines are more efficient than Savonius turbines. Thirdly, although it may be the most difficult to fabricate, it would be beneficial to obtain data on this type of turbine and airfoil combination in a Malaysian context. This is important, as the airfoil is a recent development and gaining insight on its feasibility in Malaysia could spark interest to invest further research.

Table 3. Fixed Parameters of Design

Parameter	Options
Airfoil Shape	EN0005
Number of Blades	Three-bladed (Darrieus)
Supporting Struts	Simple
Swept Area	$0.336m^2$
Solidity	0.5
Aspect Ratio	7.5
Pitching of Blade	0^0
Magnet	Neodymium
Material	Aluminum



(a) (b)

Fig. 3. (a) and (b) Final Selected Design

The axial flux generator consists of the rotor and the stator. The rotor is located on the bottom of the wind turbine, while the stator is located on the top of the base. Due to availability of materials, the axial flux generator consists of 8 neodymium magnets and 6 coils. The ideal ratio of magnet to coil for axial flux generators is 4:3. The magnets are N35 circular magnets of 30mm diameter and 10mm thickness. These magnets are placed around a circle of radius 75mm around the center of the supporting strut.

The stator consists of six coils arranged in a circle of the same radius as the rotor. Coils of 70 turns each are used. The copper coils are soldered together in the arrangement shown in Fig. 4 The stator is connected in a 3-phase star connection.

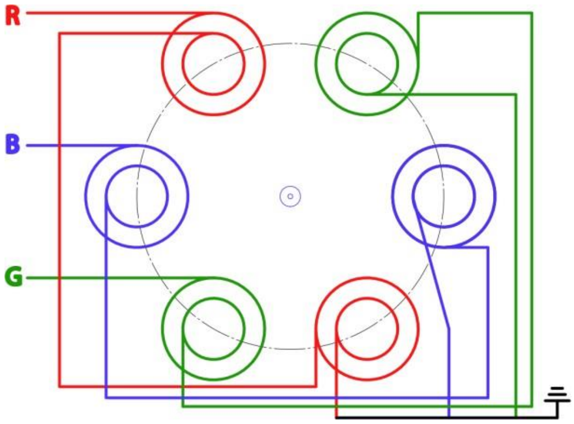


Fig. 4. Stator phase connection diagram

Different colored wires are used to indicate different phases for better organization and ease of modification in the future. The 3 phases are then connected to a 3-phase bridge rectifier to convert the AC voltage produced into DC voltage. The three-phase bridge rectifier was formed using 6 diodes, terminals, and a PCB board. The diodes and terminals were soldered onto the PCB.

3. Results and Discussion

3.1. *Simulation*

ANSYS is used for both the structural simulation and computational fluid dynamic (CFD) simulation to determine its structural stability and its velocity and pressure contours.

3.1.1. *Structural Simulation*

From the results shown in Fig. 5 and Fig. 6, the highest points of stress and strain occurs at the center of the turbine, near the area where the magnet bearings are placed. However, the maximum deformation is only 2.2976 x 10-5 mm which can indicate a very little deformation due to its own weight. Also in the simulation, the factor of safety for the entire body is 15, which is a very large number and indicates that the turbine is very strong. At the same time, it also shows that there may be a possibility to reduce the amount of material on the turbine to reduce its weight and it will still be safe.

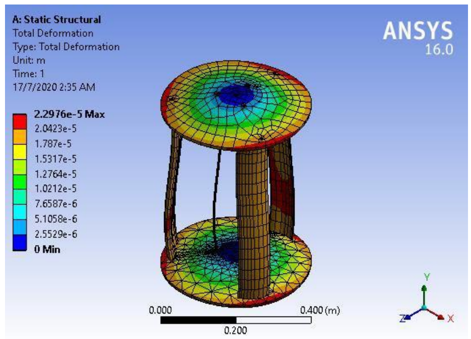


Fig. 5. Total deformation of turbine in ANSYS

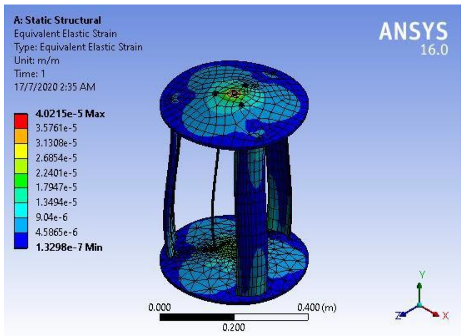


Fig. 6. Equivalent elastic strain of turbine in ANSYS

3.1.2. *CFD Simulation*

Fig. 7 and Fig. 8 are showing the velocity and pressure contour plots, respectively. These plots are generated using wind speeds of 6 m/s as used in previous calculations for the design. Fig. 7 shows a decrease in velocity magnitude from the inlet (upstream) side to the outlet (downstream) side, with a magnitude of 1.171×103 m/s at the upstream blades to 1.673×102 m/s at the downstream blades. This difference in magnitude causes an overall lift in the turbine. Similarly in Fig. 8, there is a decrease in pressure from upstream side of the swept area to the downstream side, from –8.369×104 Pa to –5.1×105 Pa. This drop creates a useful pressure difference across the blades and propels the rotor. There are also even further static pressure drops near the blades, due to the curvature of the EN0005 blades. These circulations drop the pressure as the direction of rotation are opposite to the direction of rotation of the wind turbine and further increase the power produced. From the CFD analysis, it can also be seen that wind energy is extracted by the wind turbine. This can be seen as the velocity magnitude decreases across the rotor as energy extracted from the wind will reduce the velocity.

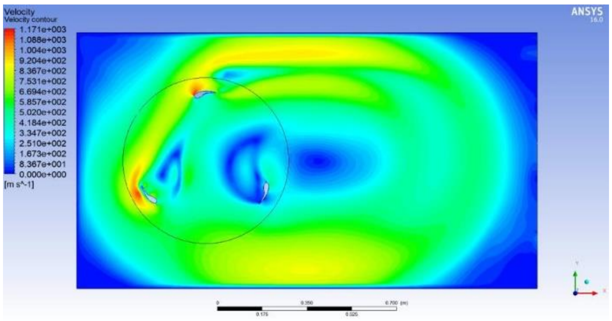


Fig. 7. Velocity contour of 2D wind turbine

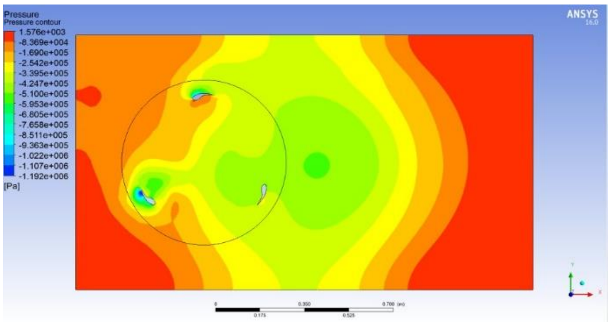


Fig. 8. Pressure contour of 2D wind turbine

3.2. ***Experimental Result***

The wind turbine prototype, as shown in Fig. 9 was tested with a three-phase load of 15Ω resistors for the peak AC voltage and a singular 15Ω resistor for the converted DC voltage at various wind speeds. The voltage and current are measured using multi-meters. The rotational speed is measured by recording video of the turbine and slowing down playback of the video to record the number of rotations within a minute. The results are shown in Table 4 with the wind speeds ranging from 2 to 4.5 m/s in 0.5m/s intervals.

Table 4. Load test results for variable wind speeds

Wind Speed (m/s)	**RPM**	**Peak VAC (mV)**	**VDC (mV)**	**Current (mA)**
4.5	90	90	52	0.25
4	60	85	49	0.23
3.5	35	81	46	0.19
3	20	76	42	0.16
2.5	14	63	35	0.12
2	6	51	30	0.08

Fig. 9. Working prototype

Sub-headings should be typeset in boldface italic and capitalize the first letter of the first word only. Section number to be in boldface Roman.

3.2.1. *Load Test Result*

Table 5 shows the tip speed ratio and power generated at the various wind speeds. The tip speed ratio is calculated from the wind speed, rotational speed, and radius of the turbine as shown in Eq. (5). In Fig. 10, power generated increases as TSR increases. Generally, power generated will peak at a certain TSR value, however due to the low wind speeds being used to test the wind turbine, the peak TSR value may be much higher.

$$Tip\ speed\ ratio\ (TSR),\ \lambda = \omega R/V\infty \quad (5)$$

Where,
ω is the rotational speed of the turbine in rad/s
R is the radius of the turbine in m (R = 0.1728m)
$V\infty$ is the wind speed in m/s

Table 5. Tip speed ratio and power generated at variable wind speeds and RPM

Wind Speed (m/s)	RPM	Rotational Speed (rad/s)	Tip Speed Ratio	Power (μW)
4.5	90	9.42	0.36	22.5
4	60	6.28	0.27	19.55
3.5	35	3.67	0.18	15.39
3	20	2.09	0.12	12.16
2.5	14	1.47	0.10	7.56
2	6	0.63	0.05	4.08

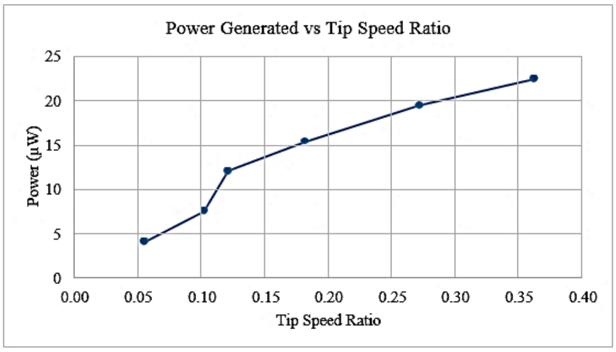


Fig. 10. Graph of Power Generated vs Tip Speed Ratio

4. Conclusion

The vertical axis wind turbine has been designed and constructed. The magnetic levitating bearings, comprised of neodymium ring magnets, are used successfully to reduce the friction and to support the weight of the turbine. In addition to the wind turbine, an axial flux permanent magnet AC generator has been also successfully constructed. The output of this generator was converted to DC using a three-phase bridge rectifier which was also self-constructed. In addition to testing magnetic levitation, the EN0005 blade profile was able to self-start consistently and produce energy at wind speeds as low as 2 m/s, which is not typical of a Darrieus turbine.

Acknowledgments

The authors would like to express the gratitude towards the support from the UCSI University, and its Research and CERVIE units who are supporting the research.

5. References

1. N. Oreskes, The scientific consensus on climate change, Science 306 (2004), pp. 1686.
2. Climate Change. Available at on-line.
3. M. Khudri Johari, M. Azim A Jalil and M. Faizal Mohd Shariff, Comparison of horizontal axis wind turbine (HAWT) and vertical axis wind turbine (VAWT), Int. J. Eng. Technol. 7 (2018), pp. 74.
4. H.J. Sutherland, D.E. Berg and T.D. Ashwill, A Retrospective of VAWT Technology, (2012).
5. M. Islam, A. Fartaj and R. Carriveau, Analysis of the Design Parameters Related to a Fixed-Pitch Straight-Bladed Vertical Axis Wind Turbine, Wind Eng. 32 (2008), pp. 491–507.
6. M. Hadi Ali, Experimental Comparison Study for Savonius Wind Turbine of Two & Three Blades At Low Wind Speed, Int. J. Mod. Eng. Res. 3 (2013), pp. 2978–2986.
7. P.L. Delafin, T. Nishino, L. Wang and A. Kolios, Effect of the number of blades and solidity on the performance of a vertical axis wind turbine, J. Phys. Conf. Ser. 753 (2016).
8. A. Sunyoto, F. Wenehenubun and H. Sutanto, The effect of number of blades on the performance of H- Darrieus type wind turbine, 2013 Int. Conf. Qual. Res. QiR 2013 - Conjunction with ICCS 2013 2nd Int. Conf. Civ. Sp. (2013), pp. 192–196.
9. F. Demmelmayr, B. Weiss, M. Troyer and M. Schroedl, Comparison of PM-machines with ferrite and NdFeB magnets in terms of machine performance and sensorless start-up control, Proc. IEEE Int. Conf. Ind. Technol. (2013), pp. 272–277.
10. M. Islam, A. Fartaj and R. Carriveau, Design analysis of a smaller-capacity straight-bladed VAWT with an asymmetric airfoil, 30 (2011), pp. 179–192.
11. N.C. Batista and R. Melicio, New blade profile for Darrieus wind turbines capable to self-start, (2011).
12. N.C. Batista, R. Melício, V.M.F. Mendes, M. Calderón and A. Ramiro, On a self-start Darrieus wind turbine: Blade design and field tests, Renew. Sustain. Energy Rev. 52 (2015), pp. 508–522.
13. L.W. Ho, Wind energy in Malaysia: Past, present and future, Renew. Sustain. Energy Rev. 53 (2016), pp. 279–295.
14. B.K. Kirke, Evaluation of Self-Starting Vertical Axis Wind Turbines for Stand-Alone Applications, Griffith University, Australia, 1998.

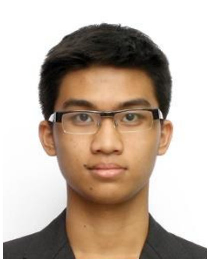
15. M.T. Parra-Santos, C.N. Uzarraga, A. Gallegos and F. Castro, Influence of Solidity on Vertical Axis Wind Turbines, Int. J. Appl. Math. Electron. Comput. 3 (2015), pp. 215.
16. Y. Chen, Numerical simulation of the aerodynamic performance of an H-rotor, Electron. Theses Diss. (2011).
17. L. Manoj, J. Nithesh, T. Manjunath and S.S. Gowreesh, Power Generation using Magnetic Levitation Vertical Axis Wind Turbine, Int. J. Eng. Adv. Technol. (2019), pp. 2249–8958.

Authors Introduction

Asst. Prof. Dr. Ammar Abdulaziz Majeed Al-Talib

He 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 Universiti Putra Malaysia (UPM), 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.

Mr. Muhammad Razin Bin Mohd Redzuan



He has completed his bachelor's degree in mechanical engineering in 2022 from UCSI University, Malaysia as a first-class graduate. Currently working in an Automation Company called Centricop as a mechanical engineer.

1.

Ts. Amar Ridzuan Bin 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 has completed his Master Degree from Universiti Putra Malaysia (UPM), and Bachelor Degree with Hounours in Mechanical Engineering (Automotive) from Universiti Teknikal Malaysia Melaka (UTeM), Malaysia.

Design of Analog Electromagnetic Gun Based on Arduino

Lintao Hu, Yizhun Peng*
College of Electronic Information and Automation, Tianjin University of Science and Technology, Tianjin, 300222, China
E-mail: pengyizhun@tust.edu.cn*
www.tust.edu.cn

The design of analog electromagnetic gun based on Arduino simulates the whole automated process of the electromagnetic gun discovering the target and launching the shell. The OpenMV camera on the simulated electromagnetic gun recognizes and determines the position of the target-shaped guide mark with a red radius of 20cm, then transmits the data to the Arduino to adjust the position of the simulated turret. The next step is that ultrasonic module measures the distance between the target-shaped guide and the turret before controlling the steering gear to rotate the barrel to the corresponding position. The shell in the barrel which is under the action of the magnetic field is accurately shot onto the target with a radius of 5cm on the ground.

Keywords: analog electromagnetic gun, Arduino, OpenMV, magnetic field

1. Introduction

The advantages of electromagnetic guns that many ordinary shells do not have are mainly that its electromagnetic gun has high density of shells, high initial velocity of shells, and not high production. It can be used to treat modern warfare, and will not have important application equipment in the electromagnetic related field.

At present, the main configuration of electromagnetic force is kinetic energy tracking. The use of the electromagnetic force generated by the post-launch launching device to launch a cannonball. Electromagnetic guns are different from traditional artillery guns. The shells of electromagnetic railguns are being fired. The magnetic energy generated by the charging and discharging device is emitted when the Lorentz force is generated by the luminescence, and the flight of the cannonball in the air is the same as that of the ancient artillery flying in the air[1]. The range of an ordinary cannon and the firing speed of a shell can be completely different. The shell of an electromagnetic curved cannon has a longer range, and the firing speed of a shell must be faster. It has a wide range of military applications. The following electromagnetic curved launches are used in different areas of the military: one is the case of a basic rocket defense system: intercepting missiles launched by vehicles or missiles[2]. Electromagnetic curved-fire cannons can completely act on air defense systems. Because the fixed-point hit rate of electromagnetic curved-fire is higher, it can convert many weapons launched to high altitudes, and can be attacked by enemy shells. The electromagnetic curved-fire cannons can be used to achieve interception[3]. The third is for anti-tank weapons: the most effective weapon such as army tank armor is the electromagnetic curved gun.

2. Hardware structure

The hardware system includes the main control Arduino and OpenMV4 H7 Plus, telescope ranging module, LDX-2018 steering gear, boost module, all-or-nothing relay module, super capacitor, 18650 battery and other hardware devices. The overall construction model of the simulated electromagnetic curved gun is shown in Fig.1.

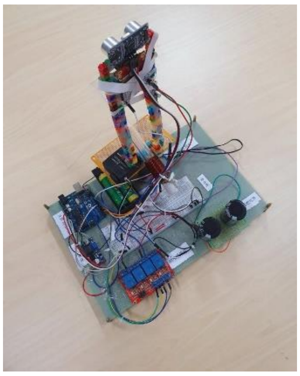


Fig.1 The design of the PCB board

2.1. Main control chip

In this design, Arduino was chosen as the MCU to simulate the electromagnetic gun because it is open source and easy to use[4]. Arduino supports the Processing/Wiring development environment of Java and C language. Therefore, the language of the single-chip microcomputer is simple, more modular, and easy to expand. Simple modifications can be made on this basis to write more complex programs and complete works with diversified functions. The circuit schematic diagram of Arduino UNO R3 is shown in Fig.2.

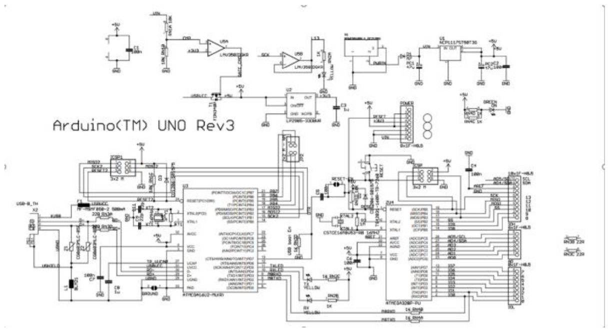


Fig.2 circuit schematic diagram of Arduino UNO R3

2.2. OpenMV4 H7 Plus

OpenMV in this design reflects the application of machine vision, and it supports MicroPython rather than C/C++. STM32H743II ARM Cortex M7 and a OV5640 photosensitive element are included in OpenMV[4]. OV5640 photosensitive element can Process 2592×1944 (5MP) images. Furthermore, most simple algorithms can run (25 ~ 50)FPS under QVGA (320×240) and below. A standard M12 lens is used by the OpenMV camera, so lenses can be swapped for different focal lengths to identify targets at different distances. The Pin diagram of OpenMV4 H7 Plus is shown in Fig.3.

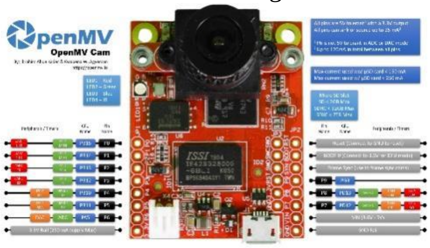


Fig.3 The Pin diagram of OpenMV4 H7 Plus

2.3. Super-Capacitor

The capacitor plays an important role in this design, it is the core of the charging and discharging device. To complete this design, need to be able to quickly charge and discharge the capacitor. Therefore, the super capacitor is chosen because the charging time of the super capacitor is short and the charging speed is very fast. It can be charged to more than 95% of the rated capacity within ten seconds, which is difficult for ordinary capacitors to achieve.

The supercapacitors have all the characteristics of the traditional ordinary capacitors, and the supercapacitors also include the advantages that the traditional ordinary capacitors do not have, it has a much stronger stored capacity than ordinary capacitors. The most special thing is that the super capacitor can release a large amount of electricity in a short time, so that the enameled wire outside the barrel can produce a large current change. Traditional capacitors include a limited number of charging and discharging of the battery, and its performance will decrease every time it is used. Super capacitors have a long service life and are more environmentally friendly than ordinary capacitors. Therefore, the charging and discharging device of this design chooses to use a 50V 1000μf super capacitor for energy storage.

2.4. All-or-nothing relay

To complete the design requirements, all-or-nothing relay main function is to control the super capacitor charge and discharge, so must have the double trigger, furthermore, VCC and GND are respectively connected to the positive and negative poles of the power supply. It requires the input voltage between 7 to 12v, the small relay specification, takes up less space, installation is more convenient. 2 channels of relay have normally open and normally closed two modes. The circuit schematic diagram of Arduino UNO R3 is shown in Fig.4.

Fig.4 The object of all-or-nothing relay

3. System software design

The flow chart of this design is shown in Fig.5.

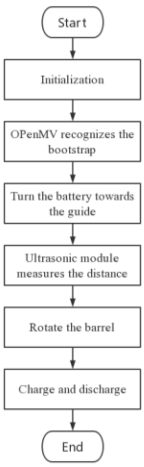


Fig.5 The flow chart of this design

3.1. Set the threshold

Seting the threshold of the target-shaped guide determines whether the camera OpenMV can accurately identify the target-shaped guide. According to the task requirements, a target-shaped guide mark with a radius of 20cm. The general red LAB threshold is (18, 69, 40, 86, 0, 72), but the actual situation, the color will be error, for example, the guide is positive red, but the red degree of LAB threshold is different, of course, the color threshold will be different under different lighting conditions. If in the indoor environment, the threshold value collected by the camera will be different under different lighting conditions of day and night, which will cause certain errors in whether OpenMV can recognize the target-shaped guide. Therefore, to use OpenMV to identify colors and collect data to send to Arduino, it is necessary to determine the LAB threshold of the target-shaped guide for several times.

Lab colors are designed to approximate human vision. So the L component can adjust the brightness pairs and modify the output order of the A and B components for precise color balance. After opening the threshold editor, point the OpenMV camera at the created target-shaped guide. Inside the source image that can see the real image, the right of the binary image is the white part of the tracking object, manual sliding below six sliders, which are respectively the minimum of L, the maximum of L, the minimum of a, the maximum of a, the minimum of b, the maximum of b. Each slide makes the target-shaped guide symbol appear white in the binary image, and all other interference sources around it appear black. Test it several times. Finally, the measured Lab threshold is (0, 73, 10, 82, 78, -13).

3.2. Identify the boot identifier

The OpenMV camera recognizes the largest color block that was previously threshold set, the target-shaped guide mark. X is the horizontal distance between the center coordinate of the target-shaped guide mark and the center coordinate of OpenMV camera view. As is shown in Fig.6, if x<=-10, the target-shaped guide mark is to the left of the analog electromagnetic gun, sending data 4. As is shown in Fig.7, if -10＜x<10, the target-shaped guide mark is facing the analog electromagnetic gun, sending data 2; As is shown in Fig.8, if x>=10, the target-shaped guide mark is to the right of analog electromagnetic gun, sending data 3. After continuous testing, it was found that the best range for judging the target-shaped guide mark was adjusted to -10 to 10.

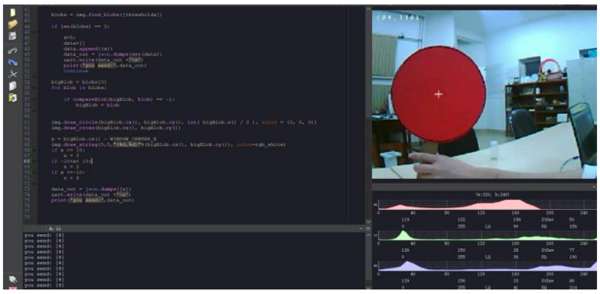


Fig.6 The target-shaped guide mark is on the left of the analog electromagnetic gun

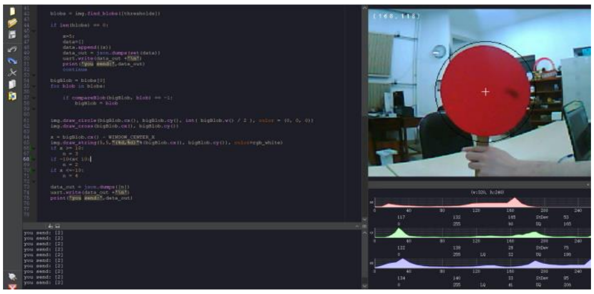


Fig.7 The target-shaped guide mark is facing the analog electromagnetic gun

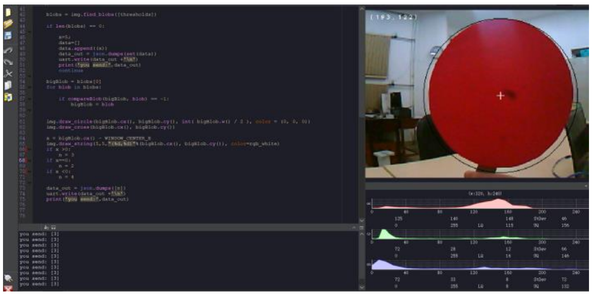


Fig.8 The target-shaped guide mark is on the right of the analog electromagnetic gun

3.3. The relationship between launch distance and steering gear parameters

The trajectory of the projectile is measured and analyzed, and the functional relationship between the firing distance and the parameters of steering gear is obtained by measuring the firing distance at different elevation angles of the barrel. When the charging time is 10s and the initial kinetic energy of the barrel is constant, there is an approximate linear relationship between the projectile firing distance and the elevation Angle of the barrel and the projectile firing distance. The relationship between the steering gear parameters and the corresponding launch distance is shown in Table 1.

Table.1 The relationship between the parameters and distance

parameters	15	17	19	21	23
distance/cm	227	228	233	231	233
parameters	25	27	29	31	33
distance/cm	235	235	237	239	241
parameters	35	37	39	41	43
distance/cm	244	247	251	251	250
parameters	45	47	49	51	53
distance/cm	249	246	244	239	237

The obtained data is generated in the Excel table scatter graph, adding trend line can get the functional relationship. Formula (1) is for the obtained functional relationship.

$$Y = 1.0362x - 218.34 \quad (1)$$

Y is steering gear parameter; x is launch distance. After the distance measured by ultrasonic wave, the microcontroller calculates the parameters of steering gear, before controlling the rotation of steering gear. Then adjusts the elevation angle of gun barrel, and launches shells.

4. Conclusion

There are many parts that need to be improved in this design. For example, the method of measuring the relationship between launch distance and steering gear parameters is to test by adjusting the steering gear angle. This method has low accuracy, and there are errors in the parameter range of different steering gear at different angles. The function diagram of launch distance and steering gear parameters is drawn by Excel, which is a smooth curve connected between points. Drawing the function in this way also introduces errors. The errors in various aspects will lead to the accuracy of projectile landing point still has room for improvement after OpenMV recognizes the target-shaped guide and ultrasonic measures the distance.

References

1. Norbert Świętochowski.(2018).The History and Use of Electromagnetic Weapons. *Historia i Polityka*(26 (33)), doi:dx.doi.org/10.12775/HiP.2018.036.

2. Haystead, J.(2015).AIR FORCE SEEKS RESEARCH ON HIGH-POWERED ELECTROMAGNETIC WEAPONS TECHNOLOGY. *Journal of Electronic Defense*(3), doi:
3. Jiuxin Gong, Yuechang Shi, Mingwei Xu.(2020).Simulated Curved Firing Electromagnetic Gun. *Modern Electronic Technology*(1), doi:10.26549/met.v4i1.3242.
4. Hu Lintao.(2021).Intelligent Lifebuoy Based on Machine Vision. *Journal of Physics: Conference Series*(1), doi:10.1088/1742-6596/2078/1/012055.

Authors Introduction

Mr. Lintao Hu

He received a bachelor's degree in electronic information engineering from Tianjin University of Science & Technology in 2021.

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 Sciences, in 2006. His research field is intelligent robot and intelligent control.

A Design of Intelligent House Inspection Robot

Sicong Wang, Yizhun Peng*

College of Electronic Information and Automation, Tianjin University of Science and Technology, 300222， China;
*E-mail: * pengyizhun@tust.edu.cn*
www.tust.edu.cn

Abstract

Aiming at the current tedious housing quality inspection procedures resulting in a lot of manpower consumption, an intelligent housing quality inspection robot is designed. Robots can independently test the quality of indoor walls and other structures. Robots consist of embedded devices combined with other devices. Independent navigation, remote control. It can obtain a variety of environmental information for mapping and can quickly and accurately measure the defects of the building, realizing the unmanned intelligent detection.

Keywords: intelligent robot, independent navigation, remote control, intelligent detection

1. Introduction

Building quality inspection robot is a research result of artificial intelligence in the field of robotics, and it is the frontier and hot spot in the field of robotics. House inspection robot is a member of the inspection robot family. It is a platform with inspection function and intelligence. The house inspection robot can detect the information of the external environment and make corresponding judgments, and transmit feedback to the client through wireless Bluetooth. House inspection robot has high performance requirements. The robot has a mechanical caterpillar chassis. This allows the robot to better maintain stability and adapt to the complex ground environment, which is the bottom rough frame of the robot. The SLAM lidar is mounted on top of the robot to create a radar scan by scanning the environment. The robot has a vision system, which can always observe the surrounding environment and feedback the observed influence to the client, so that the client can manually identify the visual image. The robot has a laser ranging module and maintain the stability of the head, the use of gyroscope to determine the robot's position and posture so as to control the head to stabilize the position and posture of the laser ranging module, can better measure the information of the surrounding buildings. The wireless Bluetooth module of the robot
can be used as a mobile phone to control the robot at any time.

2. The Main Text

In this design, STM32 is used as the main controller for image acquisition and processing, THE SLAM laser radar is controlled by raspberry PI microcomputer alone, L298N is used as the driving module to control the operation of the motor, and the steering gear of the head part is directly controlled by STM32.The six-axis gyroscope controls the pose of the head and then the pose of the laser ranging module.

3. Design of robot hardware system

3.1. *Robot structure design*

This design uses caterpillar chassis, its size is 175.9mm long, 178.6mm wide, 57.3mm high, caterpillar length is 160.5mm, 40mm wide, 44.5mm high. While using MG513 motor, chassis load 4kg, maximum speed up to 1.1m/s, except the track part are aluminum alloy material. The tracks are designed with one driving wheel on each side and five slave wheels on each side, and are equipped with a suspension shock absorption system. The 3-DOF head is composed of two steering gear, which can meet the pose stability requirements of laser ranging module.(as shown in Fig.1 and Fig.2) Its hardware design is mainly as follows:

1. The raspberries pie do PC 4 b, STM32F407 MCU do lower machine, raspberry pie than 4 raspberry pie 3 generation, plate type, part of the interface change, change, small circuit are not familiar with raspberry pie 3 May be the first time cannot distinguish the two

generation of friends, but the configuration upgrade of the sea are used to describe too much, the main are:

A. Made a comprehensive upgrade on CPU, from 64-bit A53 to 64-bit A72, and the main frequency was upgraded from 1.2ghz to 1.5ghz;

B. GPU core remains unchanged, but the main frequency is increased from 400MHz to 500MHz;

C. WiFi upgrade 5.0/BLE to facilitate the development of intelligent products;

D. The power supply circuit is upgraded to 5V-3A, increasing the possibility of expanding more modules.

2.The raspberry PI 4 is the most exciting upgrade for developers:

1) Upgrade the processor to 1.5GB

2) Multiple memory options are 1GB, 2GB and 4GB respectively

3) Bluetooth upgrade to 5.0

4) More power output 5V.3A

5)USB port is upgraded to 3.0

6)4K dual display

Caterpillar structure is adopted, self-made PCB, 6 GPIO are reserved for user-defined development.

3.Adopt 24V7A160W dual-channel H bridge L298N logic 4-channel DC motor drive board, the drive motor is suitable for all kinds of terrain, to ensure that the vehicle has strong enough power under various conditions.

4.Using SBUS to 16 channel PWM receiver with remote control of the car's movement, can transmit a variety of signals so that the vehicle can make a variety of set actions, can transmit long distance, strong signal, can be used in bad environment.

5.MPU9250 nine-axis accelerator is used to monitor and control the stability of the vehicle, combined with PID algorithm, to ensure the vehicle can maintain stability in various sections to the greatest extent.

6.The vehicle adopts the front axle, rear axle, transmission shaft, shock absorption and other mechanical structures, which can adapt to various terrains to ensure that the vehicle will not overturn in dangerous terrains, protect the safety of the vehicle and ensure the speed of the vehicle.

7.Adopt high-power power supply, output strong power at any time, to ensure the endurance of two hours.

8.Laser SLAM is adopted for faster mapping and navigation.

9.Built-in A* algorithm and Dijkstar algorithm, convenient for users to do development.

Fig.1.Robot diagram

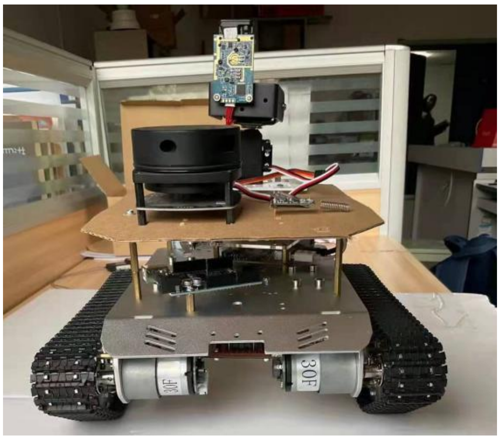


Fig.2. Main view of house inspection robot

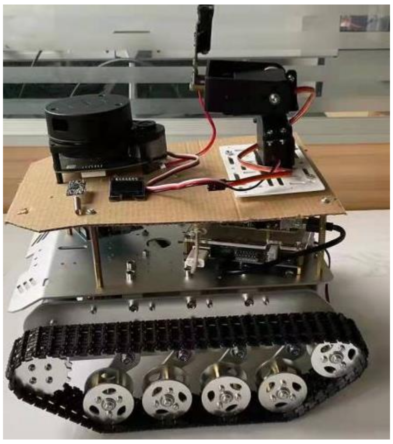


Fig.3. Side view of house inspection robot

3.2. *Introduction of controllers and peripheral modules*

The main controller adopted in this design is STM32F407, 32-bit single chip microcomputer (see Fig.4), based on ARM® The STM32F4 series of high performance microcontrollers with the Cortex™-M4 kernel use 90 nm NVM technology and ARTART technology to achieve zero wait for execution of programs, improve the efficiency of program execution, and maximize the performance of CorTEXT-M4. The STM32 F4 series can reach 210DMIPS@168MHz.The adaptive Real-time accelerator can fully unlock the performance of the Cortex-M4 kernel; When the CPU is operating at all permissible frequencies (≤168MHz), programs running in flash can achieve performance equivalent to zero wait periods. It can be used as the main controller of this kind of robot.

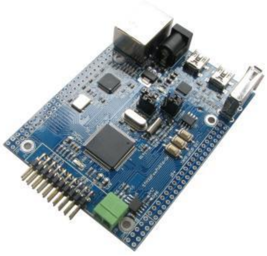


Fig. 4. STM32F407 microcontroller

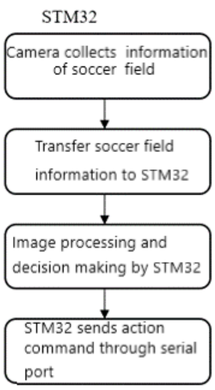


Fig.5.The flow chart of STM32

In this design, a Raspberry PI microcomputer (as shown in Fig.6) is selected to control SLAM lidar. It is an ARM-based microcomputer motherboard with SD/MicroSD card as memory hard disk and has all the basic functions of PC.

Fig.6.Raspberry PI 4B

This design adopts RPLIDAR A1 lidar (as shown in Fig.7), based on SLAM algorithm, adopts laser triangle ranging technology, in conjunction with the self-developed high-speed visual acquisition and processing mechanism, which can carry out ranging action more than 8000 times per second. Traditional non-solid state lidar mostly uses slip ring to transmit energy and data information, which is easy to be damaged due to mechanical wear. The lidar integrates wireless power supply and optical communication technology, and uniquely designs the optical magnetic fusion technology to enhance the life and stability of the radar. The radar is implemented at a scanning frequency of 5.5 Hz with a ranging resolution of 0.2% of the current range value. It can scan and detect the surrounding environment accurately and effectively.

Fig.7. Lidar

Tw10s-uart laser ranging module (as shown in Fig.8) is used to measure the distance of walls, etc., and calculate whether the walls are vertical according to the measured data. The sensor can sense the distance of the target by detecting the phase difference of laser. With high resolution. Temperature adaptability, small drift, high signal to noise ratio, small volume and other advantages. Use 3.3V TTL serial interface.

Fig.8. Single point laser ranging module

Hc-12 wireless serial communication module is a new generation of multi-channel embedded
Incoming wireless data transmission module. The wireless working frequency band is 433.4 -- 473.0MHz, and multiple channels can be set. The stepping frequency is 400KHz, and there are 100 in total. The maximum

transmitting power of the module is 100mW (20dBm), the receiving sensitivity is -116dbm at 5000Bps air baud rate, and the communication distance is 1000 meters in the open ground. We can choose one of the antennas according to the requirements. There is MCU in the module, so users do not need to program the module in addition. All kinds of transparent transmission modes just send and receive serial port data, which is easy to use. The module adopts a variety of serial port transparent transmission mode, users can choose according to the use of AT instruction. The module is controlled by STM32 master controller and can communicate with the robot remotely to control the robot.

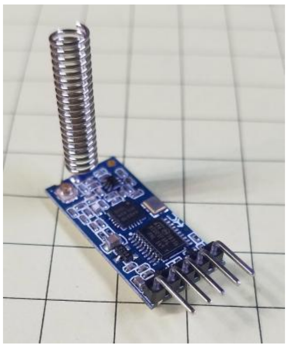


Fig.9. HC-12 wireless serial port communication module

4. The algorithm used in this paper

Maj mainly used SLAM algorithm and A* algorithm in this design.SLAM algorithm helps the radar to construct the map effectively, while A* algorithm and Dijkstra algorithm provide the autonomous navigation function of the robot.

4.1. Laser SLAM algorithm

SLAM algorithm uses structured light principle. Depth camera based on structured light usually has three core components: laser projector, optical diffraction element (DOE) and infrared camera. The fitting degree of this algorithm is better (as shown in Fig.10). SLAM system is generally divided into five modules: sensor data, visual odometer, back-end, mapping and loopback detection.
Sensor data: It is mainly used to collect various types of original data in the actual environment. Including laser scanning data, video image data, point cloud data, etc.
Visual odometer: mainly used to estimate the relative position of moving targets at different times. Including feature matching, direct registration and other algorithms.
Back end: mainly used to optimize the visual odometer cumulative error. Including filter, graph optimization and other algorithm applications.
Drawing: used for 3d map construction.
Loopback detection: mainly used for spatial cumulative error elimination. [4]

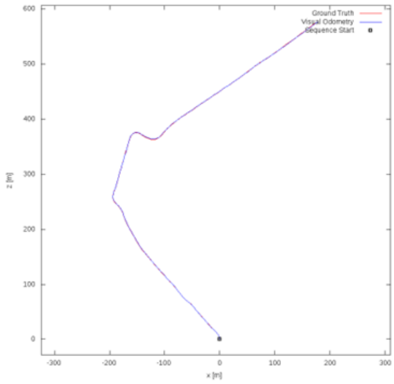


Fig.10. Fitting diagram of laser SLAM

4.2. A * algorithm

A* (pronounced: A Star) algorithm is A very common path search and graph traversal algorithm. It has better performance and accuracy. In this paper, the algorithm will be explained at the same time will also provide Python language code implementation, and will use the Matplotlib library to dynamically show the algorithm operation process. There are many similar algorithms, such as breadth-first search, Dijkstra algorithm.
A* algorithm is actually integrated with the characteristics of the above algorithms.
The A* algorithm calculates the priority of each node by using the following function.

$$f(n) = g(n) + h(n) \qquad (1)$$

Among them:
F (n) is the comprehensive priority of node N.When we choose the next node to traverse, we always select the node with the highest overall priority (lowest value).
G (n) is the cost of the distance between node n and the starting point.
H (n) is the estimated cost of node N from the end point, which is also the heuristic function of A* algorithm. Heuristic functions are discussed in detail below.
A* algorithm selects the node with the lowest f(n) value (highest priority) from the priority queue as the next node to be traversed.
In addition, the A* algorithm uses two sets to represent the nodes to be traversed, and the nodes that have been traversed, which are usually called open set and close set [5].
The flow of the algorithm is shown in the figure11.

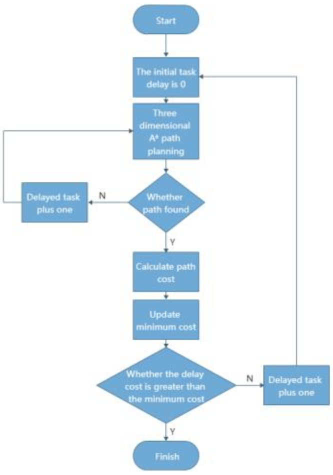


Fig.11.A* Algorithm flow chart

4.3.Dijkstra algorithm

The Dijkstra algorithm was proposed by computer scientist Edsger W. Dijkstra in 1956.

Dijkstra algorithm is used to find the shortest path between nodes in a graph.

Consider a scenario in which the costs of moving between adjacent nodes in a graph are not equal. For example, if a picture in the game has both flat ground and mountains, the game character will usually move at different speeds between flat ground and mountains.

In Dijkstra algorithm, it is necessary to calculate the total movement cost of each node from the starting point. You also need a priority queue structure. For all nodes to be traversed, the priority queue will be sorted by cost [3].

In the running process of the algorithm, the node with the lowest cost is selected from the priority queue as the next traversal node. Until we reach the finish line.

The calculation results of breadth-first search without considering the difference of node movement cost and Dijkstra algorithm considering the movement cost are compared as follows (from Fig.12 to Fig.15) :

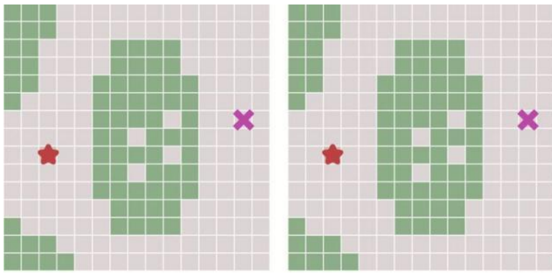


Fig.12.Result 1

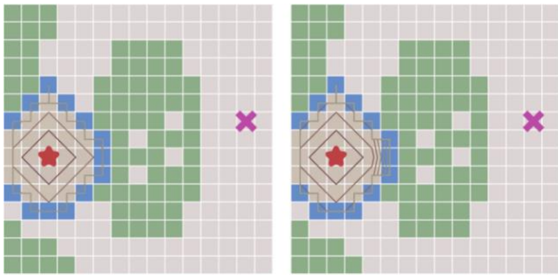


Fig.13.Result 2

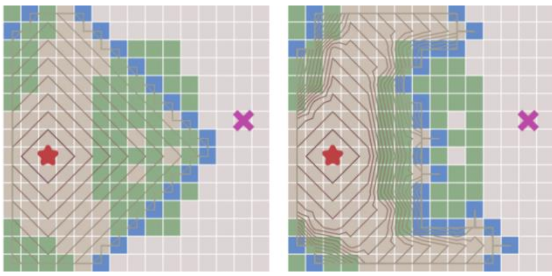


Fig.14.Result 3

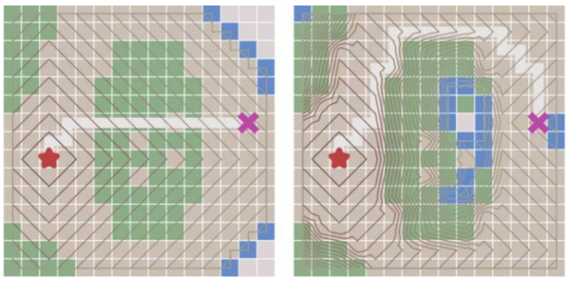


Fig.15.Result 4

When the graph is a grid graph and the cost of moving between nodes is equal, Dijkstra's algorithm becomes the same as the breadth-first algorithm.

5. Conclusion

This project mainly studies a remote-controlled intelligent house inspection robot, which can work instead of human resources under the background of aging population, labor shortage and cumbersome and inefficient use of various tools in house inspection, greatly saving human, material and financial resources.
The product has an advanced operating system and is combined with STM series of embedded devices, cameras, SLAM lidar and other sensors for quality inspection of the house. The robot can navigate automatically through SLAM laser radar mapping, and the related sensors added can detect the wall, including wall, ceiling, floor flatness, cracks, cleanliness, etc.
In addition to its autonomous navigation and positioning function as the basic platform function, it can also add a variety of replaceable sensors for different environments, greatly improving the compatibility of the device.

Acknowledgements

This research has been approved by the municipal level of The Innovation and entrepreneurship training program or College students in China. The project number is 202110057088.

References

1. SLAM for Dummies
2.CAI Zixing. Robotics. Beijing: Tsinghua University Press,2000.9.1.
3. Yongqi Chen, Shengxin Lin, Wenxiang Peng, Mingtao Guo, Xinghuan Li, Shuye Chen. Research on indoor mobile robot based on ROS. Journal of dongguan institute of technology, 2021, 28 (3) : 33-39. DOI: 10.16002 / j.carol carroll nki 10090312.2021.03.007.
4. Gong Wenting, Sun Zuolei, Pan Pan. Based on two-dimensional laser robot indoor composition . Micro computer and application, 2017, 4 (18) : 58 + 55-62. The DOI: 10.19358 / j.i SSN. 1674-7720.2017.18.017.
5. Liu Xiaofan, Zhao Bin. Mobile robot platform based on the ROS system design. Micro computer and applications, 2017, 4 (11) : 54-56 + 59. DOI: 10.19358 / j.i SSN. 1674-7720.2017.11.016.
6.HUANG Bing. Principle and application of microcomputer. Chongqing: Chongqing University Press,2001.
7.Zhou Yuanqing, Zhang Zixing, Xu Wanyong, Jia Peifa, et al. Intelligent robot system. Beijing: Tsinghua University Press,1988.
8.He Limin.Single-chip microcomputer application technology selected 1/2/3.Beijing: Beihang Press, 1992.

Authors Introduction

Mr. Sicong Wang

He is currently studying in the third year of undergraduate major in Control Engineering of. Tianjin University of Science and Technology. His research field is intelligent robots.

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 Sciences, in 2006.His research field is intelligent robot and intelligent control.

A Design of Embedded Plate & Ball Control System Based on Machine Vision

Wendinig Luo
Beijing Huadian Jingneng Energy Technology Co., Ltd, Beijing, China

FuChen Zhao
School of Information And Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, China
E-mail: 704195319@qq.com,2273840670@qq.com

Abstract

The control system is an experimental device to track the position of the target object and control the trajectory of the sphere by the actuator-rudder-driven platform motion. The research content of this design includes image processing, machine vision, motion control and many other fields. In this paper, the mechanical structure of the plate&ball control system and the printed circuit board of the main controller are designed, and the PID-based control algorithm is designed on the basis of the ball control ability analysis and the real-time detection analysis of machine vision. On the hardware platform with good self-designed performance, the precise sphere coordinates are obtained by the Hough circle detection algorithm, and then the position and speed of the sphere are controlled by the PID controller.

Keywords: Embedded, Plate&Ball control, Hough circle detection, PID control

1. Introduction

Today, with the rapid development of science and technology, humans use tools to distinguish themselves from animals, and the use of machines exceeds the limits of the body. The automation technology that combines control theory, machine vision, and mechanics has become the mainstream of today's industrial development, freeing humans from tedious and repetitive tasks. Therefore, the cricket ball control system which contains the machine vision detection and control algorithm is an important experimental platform. In recent years, Zhou Shudao proposed the design and implementation of bivariate cascade PID cricket control system.[1] Han Guangxin proposed the design of auto disturbance rejection controller of cricket system based on backstepping method.[2]Xu Yunji proposed the cricket visual control system based on double loop PID algorithm and Otsu method.[3]

2. Design of the System

2.1. *The composition of the system*

The system composition is shown in Fig.1 and Fig.2. The plate center mounted on the top of the central universal joint which is mounted on the pedestal through central knighthead. The plate can be swung up and down on the center of the central universal joint. The motor is mounted on the pedestal and the upper and lower tilts of the flat x-axis and y-axis are carried out by the corotation and reverse of the steering engine. The camera is mounted directly above the plate to detect the position and speed of the ball and as feedback for the control system, changing the steering engine angle to control the ball's speed of motion. By regulating the angle of the steering engine, we make full use of the rotation of the plate, control the movement trajectory of the ball through the rotation angle of the plate, and achieve the goal of moving the ball according to the predetermined orbit.

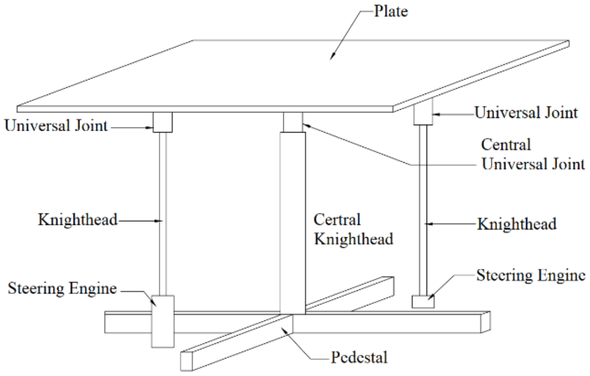


Fig. 1. A diagram of the mechanical structure

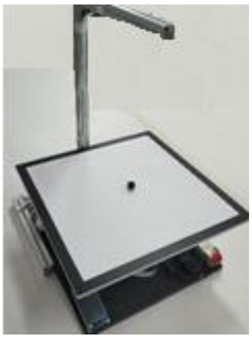


Fig. 2 Physical image

2.2. *Design ideas*

The image information of the current plate is obtained by the camera, and the center point of the plate is calculated after the boundary point is calculated[4]. The Descartes coordinate system is constructed from the center point of the plate as the reference coordinate system to identify the small ball coordinates, and the coordinates of the center of the circle are obtained by the Hough circle detection program[5]. The coordinates of the set target point are compared with the coordinates of the center of the circle to find out the error, and the error is inputted the controller as the variable. By the controller to find the variable, the steering engine drive the plate to move, so that the ball smoothly parked on the target point of the plate, or to perform some set stunts, such as rolling tracking or obstacle movement[6].

The main goal of this paper is to design an embedded plate&ball control system based on machine vision, to capture images by means of a camera, to obtain the position of the ball phase on the plate through image processing application, and to use the PID algorithm to control the ball at a fixed point.

The design covered the following areas:

- The design of machine vision-based embedded plate&ball device debugging platform. The microcontroller MK60DN512 is the main board, on-board OLED, Bluetooth, rotary encoder, steering engine and voltage module. Manipulating the angle of the rotary plate of the two-way steering engine makes it easy to manipulate the ball’s movement posture.
- The image processing technology to process of the collected picture including grayscale processing, graphic sharpening, filtering, the Harris corner point detection to obtain the boundaries of the plate, the corner point after the connection of the graphics binary. To expand the form search and approach to obtain the boundary data of the plate, relying on the Hough circle measurement technology to obtain the location data of the ball.
- Design the HMI of the plate&ball control system, observe the position of the ball, as well as the effect of PID manipulation, show the difficult factors of the control system and the directly affected factors. At the same time, the parameter graphical interface of PID parameters is designed, as well as the setting of the position of the ball. In order to infer the direct effect of different PID coefficients on the stability of the system.

3. The Design of the System’s Hardware

The MK60DN512 chip used on the embedded control board is a high-performance ARM produced by NXP with a 32-bit Cortex-M4 core that is highly resistant to interference and reliability. Internally integrated hardware multiplier, fast operation, fast DMA transmission, kernel can overclock to 200MHz, bus frequency up to 100MHz. Power consumption is no more than 50mA. As shown in the fig.3,the chip has 144 pins, and memory space is expandable, from 32 KB flash/8 KB RAM to 1 MB flash/128 KB RAM, with an optional 16 KB cache to optimize bus bandwidth and flash execution performance.

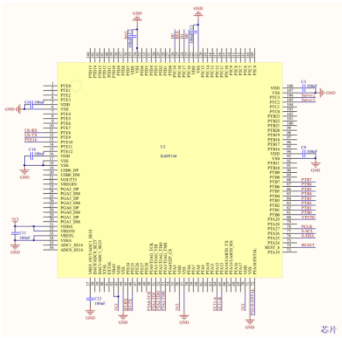


Fig. 3 MK60 diagram

It is mainly used in the high efficiency of requests and high precision control level. The MK60DN512's processing core function is strong, efficient and easy to develop, which not only ensures the stability, efficiency and availability of the microprocessor cricket control system design based on machine vision, but also streamlines the design difficulty of the control system, which greatly advances the research progress of this design.

The Joint Detection Behavior Organization JTAG port is customized on the MK60 kernel structure, and the application download and simulation mechanism can be done through the JTAG port. The K60's clock circuit consists of two parts, one of which is the main crystal of the chip, designed to emit the working clock necessary for the chip and peripherals. The rest is the clock circuit of the instant clock RTC, which supplies a set of counters that measure time during power-up and sleep configurations on the system, and the RTC real-time clock consumes very little energy. The main crystal of the plate&ball control system board uses an active crystal of 50MHz. Some microcontrollers contain clocks inside, in the case of system requirements are not high the core clock of the microcontroller can be used, and it is not necessary to design an external clock circuit.

The hardware system also includes on-board camera, Bluetooth, 0.96-inch OLED display, 3.5-inch touch screen, encoder, steering engine of external drive plate and many other external equipment. With OV-7725 camera module up to 150 frames per second, the noise removal capability is powerful and the signal-to-noise ratio is high. Built-in hardware two-value module, image processing effect is perfect. High sensitivity under weak ambient light and stable image. Compared with similar OV series cameras (300,000 resolution) such as OV-7670, OV-6620, it is the best image quality. In the commissioning of the cricket control system needs to issue a strategy or select different parameters, the user and the cricket control system need to exchange information with each other. The 3.5-inch touchscreen, made by supplier Tao Jingchi, was designed as a human-computer interface (HMI, Human-Machine-Interface). The screen resolution is 480 x 320 and the wide operating voltage can operate between 4.75-7V. The screen has a built-in master chip, and 16M FLASH and 3584Byte RAM storage. The interface to interact with the user can be designed in advance on the upper terminal, the screen can be connected via the serial port (USART-232 module) protocol, and the preset instructions can be sent to switch pages or modify data properties. The in-line Bluetooth module HC-05 can be connected to a Personal Computer with Bluetooth or various mobile smart devices, enabling the connection of the parameter debugging process to be converted to a wireless connection. In order to achieve the optimal control effect using a digital rudder, the selection of Le Magic Sol's 20KG large torque digital rudder, product weight 65g. It takes only 0.16 seconds to complete a 60-degree rotation at a 7.4V rated voltage, with an operating voltage of 6-7.4V. The empty load current is 100ma. The platform angle can be adjusted quickly and precisely in the cricket control system.

4. The Design of The System’s Software

4.1 *Real-time detection of the machine vision*

After collecting the overall environmental map of the cricket control platform system, a color graphic (RGB format) is collected. The RGB format is more practical in machine vision algorithms, and the image is made up of many pixels2. For example: the resolution is 800 x 600 images each pixel has three channels, red, green and blue each color depth of 8 bits, in the data structure is an 800 x 600 two-dimensional matrix, each point in the matrix contains (255,255,255) three channels of information, a huge amount of data. In order to reduce the amount of computation will first grayen the image processing, based on the visual effects of the human eye and proposed weighted average algorithm, because the human eye is different sensitivity to red, green, blue, so it is given different weights to make it more in line with the visual senses. In terms of human visual senses, the weights of selecting three colors are generally red 0.3, green 0.59, and blue 0.11. Grayscale parameters is :

$$Gray=0.30*R+0.59*G+0.11*B \quad (1)$$

To identify the size of the platform board of the control system, to obtain boundary information, it is necessary to filter the corresponding features, and the common features of machine visual image processing are round, edge and corner points. In this design, a rectangle with four sides and an equal length is used as a motion platform to drive the sphere movement on the board. Because of the use of rectangles as the driving plate shape of the control platform, the most obvious feature of the matrix is the four corners, only need to be detected corners. Corners are considered to be the intersection of two lines in everyday concepts, and rectangles are characterized by four corners. Broadly speaking, however, corners are image points with specific characteristics. Machine vision algorithm can obtain these characteristics in the image with pixels as the coordinate points, with digital image processing features. A pixel change in any one direction causes a grayscale value to produce a jump, which can be considered a corner point. Platform corner points are mainly 90 degrees of the intersection of two lines, so the detection accuracy is high using the Harris algorithm when detecting points of right angle classes.

Object tracking has always been one of the first technologies in the field of machine vision research in the process of scientific and technological development, and the most important thing in the development of tracking program is the accuracy and synchronization of object tracking. These two characteristics are important for filtering algorithms, the most frequently used of which is the Cam-Shift algorithm3, which has a relatively small amount of computation and is therefore fast to process. Targets can be tracked when the background environment is not very complex, such as when tracking a moving object indoors, with the help of graphic color changes.

The Cam-Shift algorithm optimizes the tracking of the target object in the case of uneven lighting to some extent, and uses the size distribution information of matrix values to track in the image matrix. In essence, the algorithm makes the optimal judgment of local processing of the results of the mean drift algorithm for each frame, and automatically adjusts the size of the retrieval core (kernel) to achieve the positioning of the tracking object.

Random Hough circle detection algorithm is a method to find straight lines, circles and other simple shapes in the image. The basic idea is to turn the problem of curve or straight line detection in the original image into the problem of peaking in parameter space. Because the circle contains 3 free parameters, it is necessary to vote for peaks in 3D space, which is large in computation and large in memory consumption, and the algorithm has a long operation time. Therefore, XU et al. proposed (Randomized Hough Transform, RHT) random Hough transformation, using three new operating mechanisms, namely, random sampling in image space, dynamic link list in parameter space, and convergence mapping connecting image space and parameter space, thus speeding up the operation speed and improving memory utilization. However, random sampling introduces a large number of invalid accumulation and invalid sampling, which has a great influence on algorithm recognition performance when dealing with complex backgrounds and larger images. Therefore, some scholars have proposed an improved RHT algorithm5, but it is undeniable that when dealing with the background is simple, especially the recognition of small images, the random Hough transformation operation is very fast and the recognition accuracy is high.

Algorithms that rely only on color to track models have some defects, and it is difficult to obtain accurate location information when the target object is subjected to uneven lighting. In order to solve this problem, a random Hough circle transformation is introduced, and Hough circle detection has high precision and fast speed for circular target recognition. The real-time positioning and tracking of moving spheres on the control plate is realized by combining Cam-Shift and Hough circle transformation.

4.2 *Ball control strategy*

In complex plate&ball control systems, PIDs can be used as feedback controllers when system characteristics cannot be fully analyzed. The response speed of the plate&ball control system is in line with the control requirements by adjusting the scale band, integral coefficient and differential terms.[7] This strategy is an effective control strategy, through different PID values, the sphere position can be closed-loop control. The closed-loop control involved in the PID controller is to give feedback on the input amount, which is a quantity of controlled variables that can directly affect the system, and feedback generally refers to the variable that can visually represent the state of the accused object.

PID control is actually the comprehensive effect of the three controls, which is proportional control, integral control and differential control[8]. Corresponding to the current error, past cumulative error and future error.

Proportional control is one of the simplest control methods, the output of this controller is directly proportional to the error of input.

Integral control automatically changes the output when steady-state error occurs, so as to eliminate steady-state error. When there is an error in the system, the integration control, according to the size of the integration time, the output of the regulator will change at a certain speed, and constantly carry out the output.

The role of differential control: The function of differential control is to predict the future trend of error signal through the change rate of error. By providing forward control, differential control can stabilize the charge process. It is often used to offset unstable trends resulting from integral control.

In the designed control platform, the coordinates and sphere motion rate of the sphere in the Descartes coordinate system are controlled by PID-based string stage. The output of the small ball position PID controller is used as the setting value of the small ball speed PID controller, and the output of the small ball speed PID controller controls the actuator steering engine, thus having a better control effect on the position of the ball. The control block diagram is shown in Fig.4.

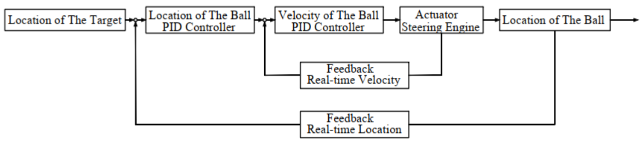


Fig. 4. Ball control string-level PID control chart

5. Conclusion

Plate&ball control system is a two-dimensional extension of the control of the inverted pendulum, the ball movement on the plate is not limited, only by the control plate movement to complete a specific track tracking. It is a comprehensive research topic for researchers who want to study automation and motion control in depth. Control feedback is tracked by Hough circle detection in the machine vision algorithm, and then the steering engine motion output is calculated. The controller continuously calculates errors to adjust the output in real time to complete the control of the position and speed of the sphere. The control platform of this design is a good combination of various disciplines. In the future engineering practice, the development of machine vision will greatly promote the popularization of intelligent systems, machine vision will be widely used in industry, agriculture and military.

References

1. Zhou Shudao，Zhang Yangchun，Wang Min. “ Design and implementation of double variable cascade PID ball and plate control system” . Science Technology and Engineering,2021,21(4) : 1454-1458.

2.Han Guangxin, Meng Shengjun, Wan Yunbo,“Design of active sisturbance rejection controller for ball and plate system based on backstepping technique”,Journal of jilin institute of chemical technology,2021,38(9):13-20

3.Xu Yunji, Gong Ming, Gong Cheng,WANG Li.“ Cricket vision control system based on double — loop PID algorithm and Otsu method”,Transducer and Microsystem Technologies,2020,39(11):0081-04

4.Li Jinhui,Xie Zaijin,Ruan Yiyuan,Chen Guangjin,Zhou Yi.“Control system of cricket based on image recognition”,Electronic Design Engineering,2020,28(13):160-164

5.Jia Zhang. “Round Target Detection Experiments Based on Random Hough Transformations”,Laboratory Research and Exploration, 2014, 33 (07): 130-133.

6.Tuo Wang. “Research on geometry detection algorithm based on improved Hough transformation”. Nanjing University of Aeronautics and Astronautics, 2017

7.Yunyun Xu. “Research on the algorithm of vision-based cricket control system”. Xihua University, 2012.

8.Zhenyu Liu, Baiying Zhang. “Analysis and improvement of predictive tracking methods based on Cam-shift”. Journal of Liaoning University (Natural Science Edition), 2017, 44 (03): 234-240.

Authors Introduction

Mr. Wending Luo

He graduated in 1997 from Mechanical and Electronic Engineering, North China Electric Power University. He is currently an electrical engineer in Beijing Huadian Jingneng Energy Technology Co., Ltd, Beijing, China.

Ms. Fuchen Zhao

She is an undergraduate majoring in information and communication engineering in Beijing University of Posts and Telecommunications. Her research interests are communication applications, mastering Matlab Simulation and modeling.

Design of Blood Circulation System of Medical Simulation Robot

Songyun Shi, Yang Ge, Chongxu Guo, Yizhun Peng*

College of Electronic Information and Automation, Tianjin University of Science and Technology, 300222，China;

*E-mail: * pengyizhun@tust.edu.cn*

www.tust.edu.cn

Abstract

Medical robot is the most active direction in the research field of medical industry in recent years. Its development potential is very huge. It has greatly promoted the progress of modern medical equipment. The blood circulation system of our medical simulation robot can be used in medical universities, medical clinical training and even military fields. The robot adopts ROS robot operating system and FreeRTOS real-time operating system, and uses multi-sensor fusion data analysis to obtain the information in the blood sample and realize the simulation of blood circulation. Iit can simulate the touch of syringe in human injection. The whole medical process is optimized.

Keywords: Medical robot, ROS, Free RTOS, simulation, medical clinical

1. Introduction

The combination of artificial intelligence and medicine enables patients to get a better treatment experience. Doctors diagnose and treat patients through advanced medical equipment. In recent years, artificial intelligence has been more and more widely used in medical treatment.

Based on the similarity principle of electricity and hydrodynamics: the resistance element is used to simulate the blood flow resistance, the inductance element is used to simulate the blood flow inertia, and the capacitance element is used to simulate the compliance of blood vessels, the electrical network model of centralized parameter blood circulation system is established, in which the heart adopts the time-varying elastic model, which is the description of Frank staring law of the heart.[0]

For example, Magnetom Sempra of Siemens is equipped with Siemens second-generation ai artificial intelligence platform, which can intelligently generate personalized and standardized scanning schemes and automatically complete the whole process of magnetic resonance scanning. Sempra inherits the excellent hardware genes of Siemens magnetic resonance: trueform intelligent magnet, cold head intelligent start and stop, intelligent targeted shimming and tim4g platform, which improves Sempra's image quality and scanning speed synchronously. This is a typical example of the application of artificial intelligence in medicine. In the learning stage, many medical colleges and universities use rabbits, mice and other living animals as experimental objects. Although they can simulate the human body to a great extent, they are not as good as the actual treatment of the human body.

In the case of high integration of artificial intelligence and medicine, aiming at this problem, this paper designs a medical simulation robot, which is used to simulate a series of indexes such as body reaction, expression, blood pressure and so on.

2. The Hardware Design

Intelligent Through comparison, we use the stm32f407 minimum system as the embedded main control, and connect the circulating pump, drv8701e motor drive module, encoder, pressure sensor, air pressure sensor, hydraulic sensor, plastic hose, 4.3-inch LCD display screen and esp8266wifi module with the external power source and the signal IO port of the minimum system to form the sensing part and control part of the robot, The embedded master is communicated with the Jetson nano microcomputer to form a decision-making system. All the hardware is finally assembled on the robot, as shown in Figure 1.

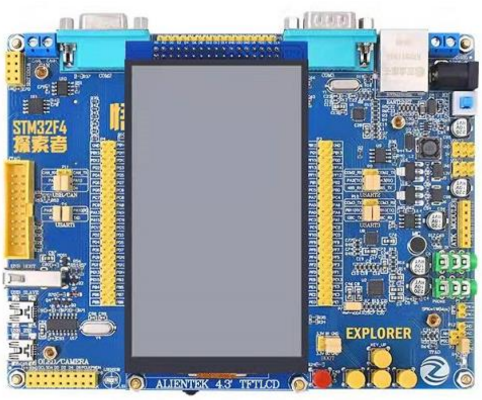


Fig.1 The design of the PCB board

2.1. Embedded main control chip

STM32f407 is a 32-bit high-performance arm cortex-m4 processor, up to 168mhz. In fact, it can overclock a little; Support FPU (floating point operation) and DSP instructions. Stm32f407zgt6: 144 pin 114 IO ports, most of which are resistant to 5V (except analog channels); Memory capacity: 1024k flash, 192K SRAM; Powerful clock system: 4~26m external high-speed crystal oscillator, internal 16mhz high-speed RC oscillator, internal phase-locked loop (PLL, frequency doubling). Generally, the system clock is obtained by external or internal high-speed clock after PLL frequency doubling, and the external low-speed 32.768k crystal oscillator is mainly used as RTC clock source; Low power consumption: three low power consumption modes: sleep, stop and standby. The battery can be used to power the RTC and backup register; 16 DMA channels with FIFO and burst support, supporting peripherals: timer, ADC, DAC, SDIO, I2S, SPI, I2C and USART; There are up to 17 timers, such as 10 general timers (32-bit TIM2 and tim5), 2 basic timers, and up to 17 communication interfaces, such as 3 I2C interfaces, 6 serial ports, etc

Fig.2. STM32F407 main control chip

2.2. Host computer design

The upper computer adopts Jetson nano and realizes the data transmission between the upper computer and the lower computer through the serial communication with the lower computer (stm32f407); The upper computer Jetson nano is equipped with ROS melody system, and processes the data measured by the lower computer according to the independently developed function package. It is the brain of the whole robot.

Fig.3. Jetson Nano

2.3. The detection system and feedback

The detection system and feedback are composed of circulating pump, encoder, drv8701e motor drive, pressure sensor, air pressure sensor and hydraulic sensor. This is also the most important part of the whole robot. The working principle is as follows: the drv8701e motor drive module drives the circulating pump to rotate, the encoder measures the speed of the circulating pump motor, feeds back the speed value to the minimum system, calculates the current blood flow rate, and the pressure in the hose measured by the pressure sensor, air pressure sensor and hydraulic sensor is fed back to the single chip microcomputer to calculate the current blood pressure; The characteristics of these devices are as follows:

- DRV8701E is a single channel H-bridge gate driver with four external n-channel MOSFETs. It is mainly used to drive 12V to 24V bidirectional brushless DC motors.

- Encoder is a device that compiles and converts signals (such as bit stream) or data into signals that can be used for communication, transmission and storage.
- The gas pressure sensor is mainly a conversion device for measuring the absolute pressure of gas, which can be used for pressure measurement of blood pressure, wind pressure, pipeline gas, etc.

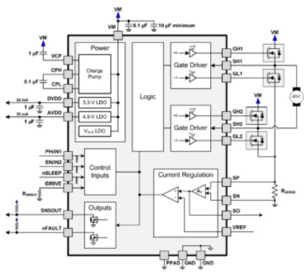


Fig.4. DRV8701E

3. Software design

The decision-making of the robot is completed by Jetson nano sending instructions to the lower computer after data processing and analysis. The lower computer preprocesses the data according to the feedback value given by each sensor module, and then returns it to the upper computer.

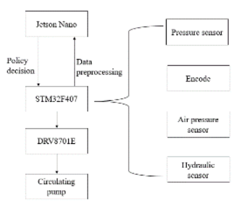


Fig.5. Software design

4. Test

After our test, the robot can simulate human blood pressure and blood flow rate under normal conditions; And it can simulate different blood pressure and blood flow rate according to the settingWireless Ad Hoc Network Design.

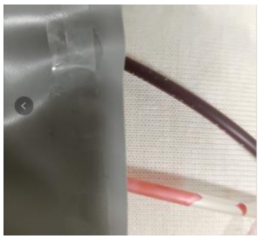


Fig.6. Blood flows through the plastic hose

5. Conclusion

Through the solution of artificial intelligence, the project studies blood pressure, integrated hydraulic pressure, blood circulation system and medical simulation. Through this solution, a medical simulation robot is developed. It is of great significance in the field of medical simulation and teaching in Colleges and universities. Through the simulation of human blood circulation, we can get the motion law of muscle, so as to enter the field of bionic robot by integrating hydraulic pressure.

The project achieves the control of blood pressure through the change of blood flow rate; In order to give users the best use experience.

References

1. Wei Shengying research on blood circulation system and physiological control based on left ventricular assist device, Harbin Institute of technology, Harbin, China, 2020
2. Huang Feng, research on extracorporeal simulated circulation test system and physiological control of left ventricular assist device, Zhejiang University, Zhejiang, China, 2014
3. Yiwei Xiong,Peng Yue,Application of FreeRTOS multitask scheduling mechanism in monitoring unit, Application of single chip microcomputer and embedded system,2021-09-01
4. Pan fan, Chen Xuechao, Zhang Junsheng. Measurement and quality control management of high-risk medical equipment. Journal of traditional Chinese medicine management, 2017, 25 (23): 184-185
5. Tang Lingsheng, Li Hongmei, Liu Zhaohui, et al. Scientific management and maintenance strategy of large medical equipment. Modern distance education of Chinese traditional medicine, 2012, 10 (13): 148-149

Authors Introduction

Mr. Songyun Shi

He is an undergraduate majoring in robot engineering in Tianjin University of science and technology. His research direction is intelligent robot, mastering MATLB simulation and robot operating system design.

Ms. Yang Ge

She is a graduate student majoring in electronic information in Tianjin University of science and technology. Her main research direction is robot simulation design, mastering deep learning and machine vision.

Mr. Chongxu Guo

He is an undergraduate majoring in robot engineering in Tianjin University of science and technology. His research direction is robot simulation technology, mastering ROS system and familiar with MCU development.

Dr. Yizhun Peng

Dr. Yizhun Peng He is an Associate Professor in Tianjin University of Science &Technology. He received a doctor's degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Science, in 2006.His research field is intelligent robot and intelligent control.

A Design of Dynamic Exoskeleton for Self-learning Human Movements

Qingliang Liu, Yucheng, Pengyu Yao, Dechao Wang,Yizhun Peng*

College of Electronic Information and Automation, Tianjin University of Science and Technology, Tianjin, 300222, China

E-mail: pengyizhun@tust.edu.cn*

www.tust.edu.cn

Abstract

In order to make the athletes more flexible and simple in the movement training, this project designed a kind of intelligent exoskeleton for the athletes to learn and complete a set of fixed movements with a higher efficiency in the training. In order to optimize the user's sports experience and adapt to more sports scenes, the product is divided into three parts: embedded device, mobile phone client and background server. The smart exoskeleton can be mechanically worn only by connecting the leg bandage and hand bandage to the human body. A single person can be completed, more convenient. It is a multi-degree-of-freedom humanoid robot system that can follow the movement of human limbs in real time. The equipment sets a series of standard movements, which leads human to repeat exercises and slowly forms muscle memory. The main target of the design is the people who want to quickly and accurately achieve the purpose of action learning.

Key words: robot; Exoskeleton. Embedded device; Motor learning; Smart devices

1. Introduction

In today's world, people pay more and more attention to the importance of physical health, in addition to a healthy diet, physical exercise is also an important part of the achievement of physical health, but in the daily physical exercise people tend to ignore the standard of exercise movement. Incorrect movement posture, will cause joint injury, muscle strain, serious even cause muscle strain; At the same time, the effect of exercise is not very good. Professional coaching is too expensive for most people to afford, and they have to rely on their own insights[1].

Although smart exoskeleton technology has gradually matured, traditional exoskeletons are mostly used in medical rehabilitation and military fields[2], and few exoskeleton devices are used in daily life (such as delivery workers wearing exoskeletons). In the rehabilitation training exoskeleton, the exoskeleton enables the patient to recover through the operation of the motor and various flexible controls. The sports training exoskeleton invented in this project is different, for example: Basketball novice in shooting training blindly will cause harm to the body, after wearing our exoskeletons, can rely on exoskeleton muscles even reach muscle memory, in a very short time to reach the standard of shooting action, etc., in the same way, we can apply exoskeleton in a variety of sports training, such as badminton, tennis serve in training[3]. Therefore, this application can be widely used in athletes, technical workers and other people in need[4].

2. Research purpose and Content

The research purpose of this project is to design an exoskeleton to carry out sports training that drives the human body. It can input a perfect set of sports training

movements, and the user can wear it and repeat the movements to form muscle memory and achieve the best effect in a short time. It can also be used in medical rehabilitation, fire and rescue activities and daily exercise.
The main content of this project is to write a complete and standard motion coordination program, input the exoskeleton through the mobile app, and execute this program after the human body wears the exoskeleton, so as to achieve the effect of muscle memory by driving the human body to repeat training for many times.

3. The Overall Design of Exoskeleton

3.1. *General function frame diagram of exoskeleton*

The overall functional system of the product includes embedded equipment, mobile phone client and background server, which integrates Bluetooth communication technology, network communication technology, asynchronous serial communication and other technologies. In the three components, the client can realize two-way data exchange through Bluetooth and embedded system, two-way data exchange between the client and the background can be realized through wireless data network, and the embedded device can send one-way information to the background through GPRS(General Packet Radio Service).
When users in preparation for connecting the exoskeleton, via mobile phone client scans the bar code on the exoskeleton, access to include information about the corresponding use of exoskeleton, and the exoskeletons of mobile client according to the corresponding bar code to send the corresponding information embedded devices, embedded devices according to receive the information on the corresponding storage unit. When the embedded device opens the storage unit, it will send feedback information to the mobile client in real time through Bluetooth communication, and the mobile client will judge whether the user correctly opens the required motion mode according to the feedback information. Then the mobile phone client sends information to the background server for the background server to judge the number of exercises and other relevant information of the current skeletal movement, and finally feedback to the mobile phone client, so that users can intuitively know their own action practice situation.

In addition to movement exercises, users can also use the mobile app to program the exoskeleton to lift and carry objects, for example. By wearing the exoskeleton and sending programs to the exoskeleton, users can easily lift and carry heavy objects.

3.2. *Overall design scheme of embedded end*

Considering cost and performance, we adopt industrial computer and STM32F1 chip equipped with ROS robot operating system as control module. In addition, bluetooth communication module, infrared detection module, motor drive module, GPRS module and relay for controlling storage unit switch are integrated outside the microcontroller. Among them, Bluetooth is used to communicate with the mobile phone client, infrared detection is used to detect the current storage state of the storage unit, and the sensor module is used to control the acquisition and preliminary processing of various external variables required in the algorithm.

3.3. *Overall design scheme of mobile client*

Mobile phone client integrates device binding and retrieval, self-check, device control, data display, Bluetooth communication, network communication and other functions. The Bluetooth function is used to communicate with embedded devices, and the network function is used to communicate with the server.

3.4. *Overall design scheme of background server*

The data processing part of the backend server takes Ali Cloud server as the carrier, and the data storage part takes Ali Cloud database as the carrier. The mobile client uploads the user's usage data to the database through network communication. The staff uses the background server to analyze and process the data received.

4. Exoskeleton Mechanical Structure Design

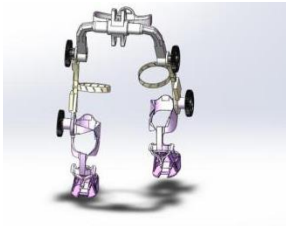
The exoskeleton robot has a total of 12 degrees of freedom, among which 10 degrees of freedom are motor-driven and 2 degrees of freedom are adaptive. The following is an overall view of the exoskeleton. The structure reduces human and motion errors.

Fig.1 Overall structure of exoskeleton

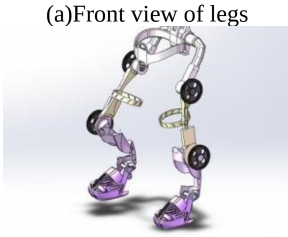
Each arm is provided with 3 degrees of freedom, and all are motor driven type, its extension is flexible, can be used for rehabilitation exoskeleton[5]. Suitable for operating under heavy load, as shown in Figure 2.

Fig.2 Structure of exoskeleton arm

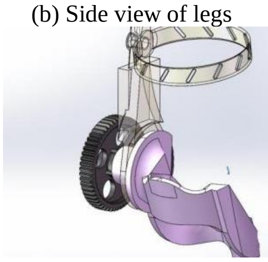
For the lower limbs, each leg is equipped with two degrees of freedom motor driveable type, and one degree of freedom is adaptive foot type, in which the two degrees of freedom motor can drive can meet the normal walking and running two states[6]. The adaptive degree of freedom of feet is to relieve the pressure and discomfort caused by mechanical devices on feet. The driving mode of the drivable degree of freedom mainly adopts the motor driving type, that is, the motor is fixed with a gear device, and the corresponding joint is driven by the transmission of the gear device. Here take the lower leg joint as an example. There are fixed gears on the lower leg joint. Assume that when the thigh joint is stationary, when the gear of the lower leg joint rotates the corresponding Angle, the lower leg joint also rotates the corresponding Angle, so the function of the lower leg joint is achieved[7]. The motor controlling the rotation of the lower leg joint is generally located on the thigh joint. The rotation of the lower leg joint is relative to the rotation of the thigh joint. Similarly, the gear controlling the rotation of the thigh joint is located in the crotch, and the rotation of the thigh joint is relative to the crotch joint, as shown in Figure 3.



(a)Front view of legs

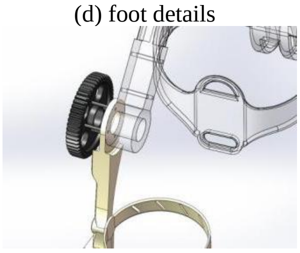


(b) Side view of legs

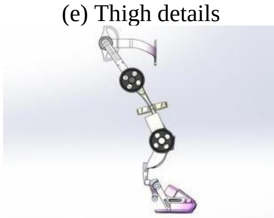


(c) Calf details

(d) foot details



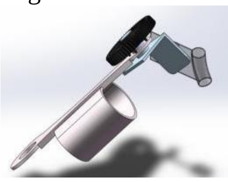
(e) Thigh details



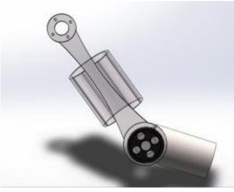
(f) Profile of a single leg

Fig.3 Exoskeleton leg structure

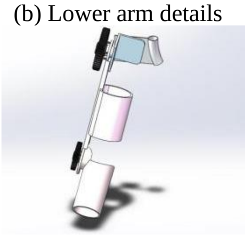
In the upper limb, the rotation of each joint is the same, and the rotation of the joint is relative to the upper joint. Therefore, by controlling the rotation Angle of each joint, the flexible operation of each joint can be achieved, that is, to achieve the rehabilitation and power effect of the exoskeleton. See Figure 4.



（a） Upper arm details

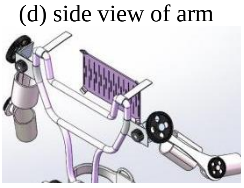


(b) Lower arm details

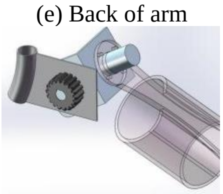


(c)Front view of arm

(d) side view of arm



(e) Back of arm



(f) Details of arm extension on both sides

Fig.4 Upper limb structure of exoskeleton

5. Exoskeleton Hardware Design Scheme

5.1. *Power management*

The exoskeleton's power core is powered by a series of lithium batteries[8]. In the process of using the battery charge, discharge constraint protection, so the design can meet the requirements of charge and discharge protection power management module. Power supply voltage is 36V. In the process of charging and discharging, the voltage of each battery is equalized by the internal self-balance and external balance of BQ76930 chip to prolong the battery life and increase the protection of battery pack overheating. In order to save MCU resources and facilitate wiring, BQ76930 is controlled by STM32 MCU.

5.2. *Control module*

The control core of the exoskeleton processes all kinds of data sampled by the exoskeleton through single chip microcomputer, and provides appropriate control strategies to guide the movement of the exoskeleton after calculation.

SCM control brushless DC motor joint to follow the human movement, considering the hardware in the process of signal acquisition, processing, control, as well as in the expansion, so choose the ARM core chip. We choose STM32 F1 series chips from the balance of cost, performance and range of use.

From the consideration of product cost and maintainability, the number of PCB boards is less. The control of each part of the joint also needs to be independently considered in the coordination from the perspective of cost saving. A main control board integrates a motor drive module and the sensor port required by this part of the joint. In the power supply part, in order to reduce the volume of the voltage regulator module, the Jin Sheng Yang voltage module is used to supply 36V to the drive module, and then to 24V, and then to 5V, and to supply the control board. In the board, there is a CAN bus module. The STM32 chip on the control board CAN interact with the chips on other control boards through the CAN bus. Achieve coordination and synchronization.

5.3. *Drive motor*

At the power core of the mechanical exoskeleton, each joint has a robotic joint motor, which transmits torque through cable traction to control the movement of each joint. The motor drive module needs to meet the motor position, speed and torque output requirements in the control process. In the selection of power, we choose the drive module has the advantages of light, small and stable product commercialization.

Choose electric drive mode. Since joint motion needs to be controlled and the motor needs to be blocked and frequently reversed, the robot joint motor is selected as the driving mode. The rotation position of the motor is detected by hall components, and the motor speed is accurately measured by a three-phase encoder.

5.4. *Sensor module*

The part of the exoskeleton interacting with human body undertakes the acquisition and preliminary processing of various external variables needed in the control algorithm, including EMG EMG signal acquisition sensor, angular speedometer, angular accelerometer, inertial navigation sensor and pressure sensor.

When the nerve controls the muscle movement, there is a potential difference along the nerve direction, and the EMG electromyography sensor determines whether there is a nerve signal controlling the muscle movement by detecting the pressure difference. According to the sensor line is relatively close, the amplifier needs to have a good common-mode rejection ratio, the AD8221 is used for signal processing, in the AD8221 front, there is TL084 operational amplifier for the first step of signal amplification.

The pressure sensor is used to detect the pressure between the skin and the outer fabric of the exoskeleton, and is used as a feedback signal to guide the drive of the motor. The optimal goal is zero pressure between the human body and the exoskeleton in all directions. The pressure sensor is a thin film piezoresistive sensor. The pressure is measured by the change of the pressure generated by the change of the resistance value of the compression resistance in the resistance bridge.

Angular speedometer and angular accelerometer are bound to each limb of the human body, without relative movement, and the angular change of the limb relative to the vertical plane is determined. Among them, MPU-3050 is a three-axis gyroscope, which measures the Angle. MMA8421 is a triaxial accelerometer that measures acceleration. The data they measure are the raw data of the generalized variables in the Lagrange equation. In hardware design, the angular velocity and angular accelerometer integrated in a 0.5 cm x 1 cm of small PCB, its with and to measure the body through the bind fixed, PCB by needle are connected to the main control board, the signal back to the single chip microcomputer, using inertial navigation position sensor, the relative position of the body and ensure the accuracy of the motion.

6. Conclusion

The project in all respects have carried on the thorough research and exploration, including mechanical structure, hardware, software, etc., this one in sports training exoskeleton is a new type of intelligent training system based on AI control technology, through the use of humanoid robot technology combined with the content of the movement technology subject, provide short-term effective mass sports training methods, It can help people to exercise correctly, so that people can reduce the time of ineffective exercise while avoiding injury, quickly achieve the desired effect, and then carry out the follow-up training, and can be connected to the mobile phone client. Correct movement mistakes in time, convenient and correct wearing, intelligent, lightweight and powerful.

Acknowledgements

This research was supported by Student’s Platform for Innovation and Entrepreneur Training Program, Tianjin University of Science and Technology(202110057076).

References

1. Wan SL, Yang MX, Xi RR, et al. Design and control strategy of humanoid lower limb exoskeleton driven by pneumatic artificial muscles. Proceedings of 2016 23rd International Conference Mechanic Machine Vision Practice. 2016
2. Han Yali, Wang Xingsong. Dynamics analysis and simulation of lower limb assisted exoskeleton [J]. Journal of system simulation,2013,25(01):61-67+73.
3. Li Longfei, Zhu Lingyun, GOU Xiangfeng. Design of human-machine closed chain lower limb rehabilitation exoskeleton for Human Movement requirements [J]. Machinery Design,2021, V. 38; No. 382 (8) : 53-60, DOI: 10.13841 / j.carol carroll nki JXSJ. 2021.08.008
4. Alan T. Asbeck, Stefano M.M. De Rossi, Kenneth G. Holt,Conor J. Walsh. A biologically inspired soft exosuit for walking assistance[J]. The International Journal of Robotics Research . 2015 (6)

5. Gao Jianshe, Zuo Weilong, Yu Qianyuan. An upper limb rehabilitation robot motion simulation and experimental research [J]. Journal of mechanical design and manufacturing, 2020, No. 349 (3) : 155-158. The DOI: 10.19356 / j.carol carroll nki. 1001-3997.2020.
6. Shi Xiaohua, Wang Hongbo, Sun Li, Gao Feng, Xu Zhen. Structural design and dynamics analysis of exoskeleton-type lower limb rehabilitation robot [J]. Journal of mechanical engineering,2014,50(03):41-48.
7. Xiang Zhongxia, Zhao Ming, Gao Fei, Jin Teng, Hu Zhigang, Zhang Jian. Simulation of training Effect of exoskeleton Rehabilitation Robot [J]. Journal of Tianjin University (Natural Science and Engineering Technology),2016, V. 49; No.302(07):695-701.
8. Fan Boqian. Research on Key Technologies of Hydraulically Driven Lower Limb Exoskeleton Robot [D]. Zhejiang University,2017.
9. Variable impedance actuators: A review[J]. B. Vanderborght,A. Albu-Schaeffer,A. Bicchi,E. Burdet,D.G. Caldwell,R. Carloni,M. Catalano,O. Eiberger,W. Friedl,G. Ganesh,M. Garabini,M. Grebenstein,G. Grioli,S. Haddadin,H. Hoppner,A. Jafari,M. Laffranchi,D. Le

Authors Introduction

Mr. Qingliang Liu

He is currently pursuing a bachelor's degree at Tianjin University of Science and Technology.

Mr. Yu Cheng

He is currently pursuing a bachelor's degree at Tianjin University of Science and Technology.

Ms. Pengyu Yao

She is currently pursuing a bachelor's degree at Tianjin University of Science and Technology.

Mr. Dechao Wang

He is currently pursuing a bachelor's degree at Tianjin University of Science and Technology.

Dr. Yizhun Peng

He is an Associate Professor in Tianjin University of Science & Technology. He received a doctor's degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Sciences, in 2006. His research field is intelligent robot and intelligent control.

Autonomous Navigation Building Climbing And Handing Robot Based on SLAM

Linhui Chen, Junjie Ta, Yizhun Peng*
*College of Electronic Information and Automation, Tianjin University of Science and Technology, Tianjin, State 300222, China**

*E-mail: * pengyizhun@126.com*

www.tust.edu.cn

Abstract

When there is no elevator at the station and airport or the elevator is crowded, the ramp of baggage checking is steep, the weight of luggage package is too heavy, it takes time and effort, and it is very easy to knock against and cause damage to the contents of the package.Our climbing robot carries out autonomous navigation based on visual SLAM. Equipped with ROS robot operating system, it adopts laser SLAM mapping navigation technology and tracks as running parts, which is suitable for various terrains. The adoption of binocular vision, can be more accurate analysis of the surrounding environment. The robot is equipped with MPU9250 attitude sensor, which is convenient for solving robot attitude. Jetson Nano was used as the upper computer and STM32F429 as the lower computer.

Keywords: SLAM build figure, Arduino, Gmaooing algorithm, Machine learning, Artificial intelligence (AI)

1. Introduction

Manually moving heavy objects upstairs is time-consuming and laborious, and it is extremely easy to bump and cause damage to the objects. Places with a large flow of people are also prone to road congestion, and people with disabilities and the elderly also have difficulties in getting upstairs. Nowadays, with the rapid development and progress of artificial intelligence technology, robotics has attracted much attention, and the requirements for robots' mobility are getting higher and higher. Mobile robots that can climb obstacles autonomously have become a hot research topic, such as: autonomous stair-climbing robots based on vision guidance, and wheel-track composite stair-climbing robots[1,2].Therefore, the scientific and practical significance of the climbing robot is very high.

A stair climbing robot based on SLAM mapping has been developed to realize autonomous positioning and navigation of the robot. Its vision system can accurately analyze the surrounding environment and autonomously avoid obstacles. The crawler type of transportation makes it adapt to a variety of environments. The robot can memorize the walking route through machine learning, and quickly plan the shortest path according to the use situation, which provides great convenience for the user.

2. The Hardware Structure

The robot is equipped with MPU9250 attitude sensor, which is convenient for solving robot attitude. Jetson Nano was used as the upper computer and STM32F429 as the lower computer.

2.1. *Sensor*

2.1.1. Nine axis attitude sensor

The MPU9250 module consists of two parts, one part is the MPU6500 three-axis gyroscope and three-axis accelerometer, and the other part is the AK8963 three-axis magnetometer. It has a 3-axis gyroscope, a 3-axis accelerometer and a 3-axis magnetometer. The output is 16-bit digital; the robot is equipped with such a sensor to exchange data with the microcontroller through the integrated circuit bus IIC interface, and it detects four The elements enter the DMP digital motion processor on the chip to calculate Euler angles and calculate the robot's posture in real time. Adjust travel speed through

Fig. 3. Jetson Nano chip

traditional PID algorithm.

2.1.2. *Image Sensor*

OV7670 image sensor, small size, low working voltage, provides all the functions of a single-chip VGA camera and image processor. Through SCCB bus control, it can output 8-bit impact data with various resolutions such as full frame, sub-sampling, and window fetching.

Fig. 2. OV7670 camera module

2.2. ***Main control chip***

The Jetson Nano hardware is a quad-core Cortex-A57 CPU, and the GPU is the smallest Maxwell architecture graphics card, with only 128 CUDA units, equipped with 4GB LPDDR4 memory and 16GB storage space. Supports high-resolution sensors, can process multiple sensors in parallel, and can run multiple modern neural

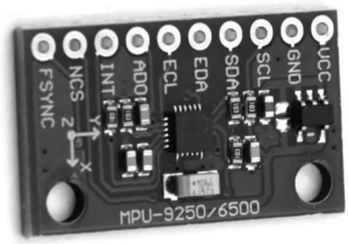


Fig. 1. Nine-axis attitude sensor modu

networks on each sensor stream. It also supports many common artificial intelligence frameworks.
Just insert the microSD card with the system image to start, built-in SOC system-level chip, can parallel processing such as TensorFlow, PyTorch, Caffe/Caffe2, Keras, MXNet and other neural networks, these neural networks can be used to achieve image classification, target detection , Voice segmentation and intelligent analysis and other functions, performance is better than Raspberry Pi, more practical, can meet the requirements.

2.3. ***Lower computer***

STM32F429 uses Crotex M4 core, the highest frequency is 180MHz, there are up to 8 UART serial ports, Flash is 1M, SRAM is 256KB, pin is 176pin, high performance, fully meets the pin and memory requirements required by the design.
It has the following characteristics:

- Operating temperature range: -40 ~+105℃.
- Operating Voltage: 1.7~3.6 V.
- Integrated high-speed embedded memory and up to 4K bytes of backup SRAM, as well as a large number of enhanced I/O and peripherals connected to 2 APB buses, 2 AHB buses and a 32-bit multi-AHB bus matrix.

STM32F429 works with OV7670 to collect images, and upload the images to the host computer through the serial port, the host computer receives and processes the data,

Fig. 6. track

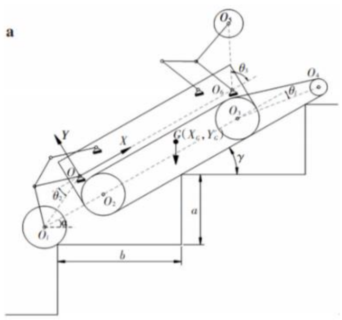


Fig. 7. supporting structure

and finally displays the color image.

2.4. *GPRS wireless communication module*

This design uses the E29V GPRS module to realize the communication between the small program and the main control part. It is small in size and supports TCP/UDP transparent transmission, MQTT transparent transmission, HTP transparent transmission, AT command function SMS transparent transmission, data transparent transmission and other functions to meet the requirements.

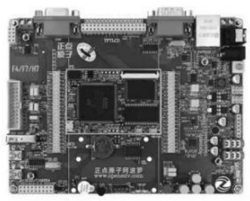


Fig. 4. STM32F429 chip

3. Mechanical structur

The movement adopts plastic crawlers to minimize damage to the ground and adapt to many kinds of environments. The measured maximum load-bearing capacity of the supporting structure is 137.87KG, which can carry most of the luggage in life.

Fig. 5. GPRS wireless communication module

4. Design functional modules

With these functional modules, the robot can have "eyes" and "logical thinking" and become more intelligent through machine vision and machine learning.

4.1. *Digital image processing*

The application of digital image processing technology makes the robot's perception of the environment more clear.

For the collected image, according to the spatial distribution of its light intensity, it can be expressed in the form of the following function:

$$I=f(x,y,z,\lambda,\ t) \tag{1}$$

4.2. *Machine learning*

Machine learning is a branch of AI, which explores ways to make computers improve efficiency based on experience[3]. Machine learning can enable robots to memorize long-term routes, and quickly plan the shortest path based on future use, and will follow the user’s Use records to push related endpoints.

4.2.1.*Dijkstra algorithm*

Dijkstra algorithm uses breadth-first search to solve the single-source shortest path problem of weighted directed graphs or undirected graphs, and the algorithm finally obtains a shortest path tree[4]. It can realize global path planning, so that the robot can quickly calculate the optimal path that needs to be taken

4.2.2.*Naive Bayes Algorithm*

Naive Bayes makes the assumption of independence, assuming that each feature is independent and uncorrelated, and calculates the probability:

$$P(y \mid x_1,\cdots,x_n)=\frac{P(y)P(x_1,\cdots,x_n \mid y)}{P(x_1,\cdots,x_n)} \tag{2}$$

The Naive Bayes algorithm analyzes the data, predicts where the user may go, and at the same time calculates the optimal path.

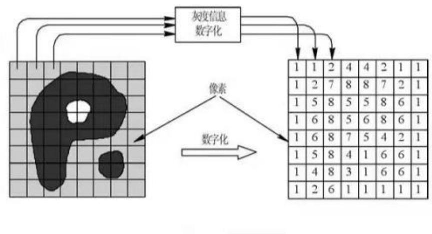


Fig.8 . Digital image processing

4.2.3 *A**

In a * algorithm, heuristic information is represented by function f *: f * (x) = g * (x) + h * (x)

Where: g * (x) is the cost of the best path from the initial node to node X;

H * (x) is the cost of the best path from X to the target node;

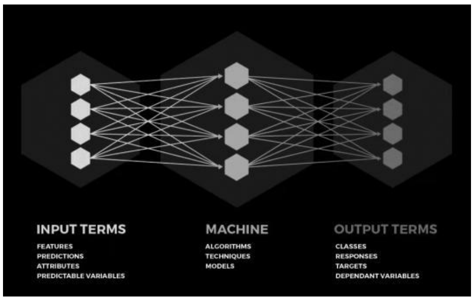


Fig. 9. Machine learning

F * (x) is the total cost of the best path from the initial node to the target node through node X.

Based on the above definitions of G * (x) and h * (x), G (x) and H (x) in the heuristic search algorithm are limited as follows:

① G (x) is an estimate of G * (x), and G (x) > 0;

② H (x) is the lower bound of h * (x), that is, H (x) ≤ h * (x) for any node X.

The ordered search algorithm when the above conditions are met is called a * algorithm.

For a search algorithm, when the best path exists, it can be found, then the algorithm is called admissible. It can be proved that a * algorithm is a Connor algorithm. In other words, for the ordered search algorithm, when the condition of H (x) ⩽ h * (x) is satisfied, as long as the best path exists, it can be found.

The heuristic search of a * algorithm is to search in the state space. First, evaluate the location of each search to get the best location, and then search from this location to the target. In this way, compared with Dijkstra algorithm, it can omit a large number of unnecessary search paths and improve the efficiency.

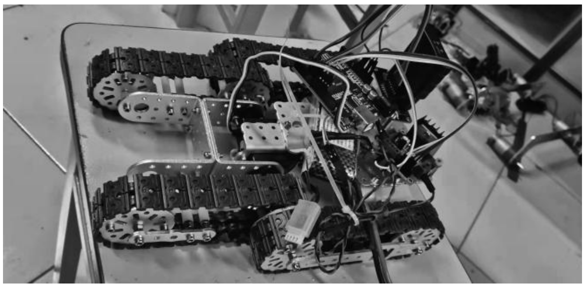


Fig. 10. Physical prototype

5. Conclusion

The above is the analysis and theoretical support for the Autonomous Navigation Building Climbing And Handing Robot Based on SLAM after reading a lot of literature. We have made a theoretical prototype to prove the feasibility of the idea.

Acknowledgements

This research was supported by Student's Platform for Innovation and Entrepreneur Training Program, the Ministry of education of China(Tianjin University of Science and Technology, the project number is 202110057137.

References

1. Zhu Fenyong, Chen Hucheng, Nie Weijun. A wheel-track composite staircase climbing robot. Equipment Manufacturing Technology, 2020(05): 35-39+43.
2. Zhang Xiaobing, Song Aiguo. Robot autonomous climbing stairs based on vision guidance. Journal of Anhui University of Technology (Natural Science Edition), 2006(02):186-188.
3. Hemachandran K, Shubham Tayal,Preetha Mary George: Bayesian Reasoning and Gaussian Processes for Machine Learning Applications,CRC Press:2021-11-01.
4. Yan Biyu, Tong Youyou, Ma Jingyu. "Research on Optimization of Takeaway Delivery Time Based on Shortest Path." Modeling and simulation 10.02(2021): doi:10.12677/MOS.2021.102024.

Authors Introduction

Ms. Chen Linhui

She is currently a student of Robotics Engineering at Tianjin University of Science and Technology.

Mr. Tai Junjie

He is currently a student of Robotics Engineering at Tianjin University of Science and Technology.

Dr. Yizhun Peng

He is an Associate Professor in Tianjin University of Science & Technology. He received a doctor's degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Sciences, in 2006. His research field is intelligent robot and intelligent control.

Simulation Research on Automatic Navigation of Indoor Wheelchair

Peng Shi[1,2], Yizhun Peng[1,2,*]
[1] College of Electronic Information and Automation, Tianjin University of Science and Technology, Tianjin, 300222, China,
[2] Advanced Structural Integrity International Joint Research Centre, Tianjin University of Science and Technology, Tianjin, 300222, China
E-mail: pengyizhun@tust.edu.cn*
www.tust.edu.cn

Abstract

A new auto-navigation wheelchair based on ROS system is proposed to deal with the global aging and the low behavioral ability of the elderly. Lidar is used to locate and map the active area using gmapping algorithm. Real-time map information is transmitted to the processor by the camera and lidar working together. Automatic navigation is completed by A* algorithm calculation. Important information points are marked by QR code and precisely positioned by camera recognition, which enables wheelchair to have automatic navigation function. It can help older people move safer and more freely at home; It can also be applied to nursing homes to reduce the pressure of nurses and centralize management of the elder.

Keywords: wheelchair, ROS, Gmapping, A algorithm*

1. Introduction

According to the World Bank, by the end of 2019, nearly 654.6 million elderly people were aged worldwide, accounting for 9% of the total population. With the increase of the total number of the elderly, it has become more important to pay attention to the quality of life of the elderly, among which, the travel convenience of the elderly is the premise of their high quality of life. As a frequently used tool for the elderly, its convenience is particularly important. Long-term use of traditional manual wheelchairs will have a huge burden on people's wrists and arms; electric wheelchairs need operators to pay attention to the surrounding environment for a long time, pay attention to their own control, easy to produce fatigue and then cause accidents.

For the above problems, we will design an automatic navigation wheelchair robot built around the ROS[1] platform. In the indoor environment, the surrounding environment can be perceived through lidar and multiple sensors. Establish a model in gazebo for simulation, and use the Gmapping[2] algorithm in the SLAM[3] algorithm to create a graph.[4] Then display in rviz simulation. Navigation using the A* algorithm[5] allows the wheelchair for stable autonomous driving in a simulated indoor environment.[6]

2. Model Building

For the authenticity of the simulation experiments, wheelchair models need to be imported into the Gazebo simulation environment. The SolidWorks software provides plugins to automatically export URDF files with physical shape, collision parameters, and inertia parameters, which can be used directly by the Gazebo simulation environment. Therefore, SolidWorks software was selected to build the physical model of the wheelchair. To enable the wheelchair to complete SLAM and navigation, add lidar and camera to its pedal position and upper backrest. The model established by SolidWorks is shown in Fig. 1.

Fig. 1. The model established by SolidWorks

In SolidWorks only has physical models, and in order to complete simulating the real simulation environment in Gazebo, it is also necessary to add wheel drive plugin, differential drive control plugin, and sensor plugin to the established physical model. The simulations of the full model in Gazebo are shown in Fig. 2.

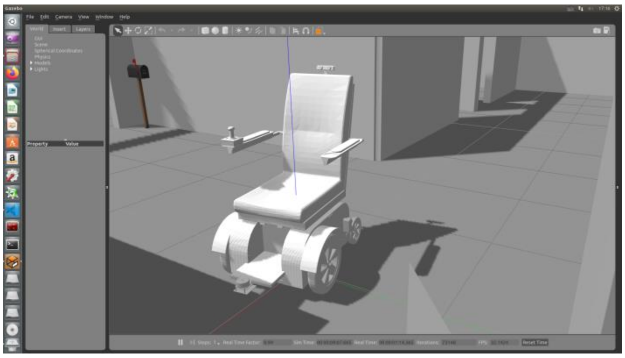


Fig. 2. The simulations of the full model in Gazebo

2.1. *Gazebo simulation environment construction*

The robot model is shown in the Gazebo. But currently by default, the robot model in Gazebo is in empty world, and there are no emulations similar to rooms, furniture, roads, trees.... There are three ways to create a simulation implementation in Gazebo:

- Add built-in components directly to create a simulation environment,
- Manual drawing of the simulation environment,
- Download using the official or third-party improved simulation environment plugins directly.

The simulation selection manually drew the simulation environment. A resthome environment with an area of 200m^2 was drawn with parts of furniture added to the environment to simulate the real nursing home environment. The established Gazebo simulation environment is shown in Fig. 3.

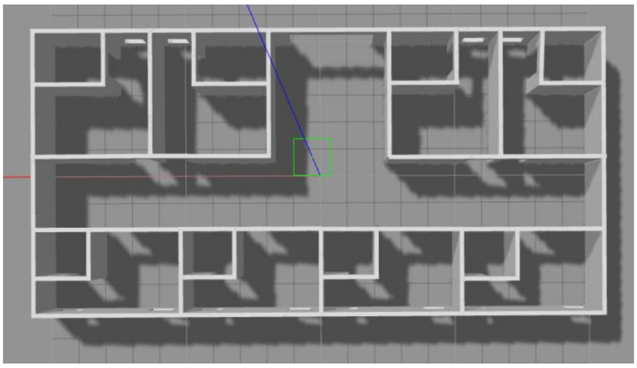


Fig. 3. The established Gazebo simulation environment

3. Navigation Preparation

3.1. *Slam*

SLAM, also known as CML, is Concurrent Mapping and Localization, or simultaneous localization and mapping. The question can be described as whether putting a robot in an unknown position in an unknown environment, is there a way to let the robot move and gradually draw a complete map of the environment. The so-called a complete map refers to every corner of the room without obstacles.

Gmapping is one of the more commonly used and relatively mature SLAM algorithms in the ROS open-source community. Gmapping can draw a two-dimensional grid map according to the mobile robot odometry data and laser data. Correspondingly, Gmapping also has certain requirements for hardware:

- The mobile robot can post an odometry message,
- Robots need to release radar messages.

Frist, write the launch file related to the Gmapping node and start the Gazebo simulation environment. Then start the mapping launch file and start the keyboard control node to control the robot movement. Add components to display the grid map in Rviz. Finally, the robot movement can be controlled in the Gazebo through the keyboard, and the grid map data released by the Gmapping can be displayed in Rviz. The display of the grid map data in Rviz is shown in Fig. 4.

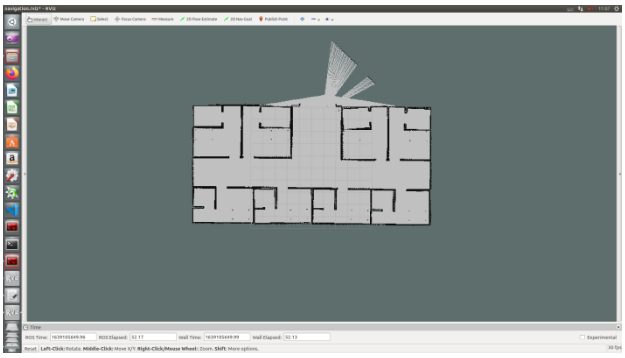


Fig. 4. The display of the grid map data in Rviz

3.2. *Robot localization*

localization is to calculate the location of the robot itself in the global map. Of course, SLAM also includes localization algorithm implementation, but SLAM localization is used to build a global map, belongs to the stage before the beginning of navigation, and the current localization is used for navigation. In the navigation, the robot needs to move according to the set route. Through localization, it can be judged whether the actual trajectory of the robot meets expectations. The AMCL feature package is provided in the ROS navigation feature package for enabling robot localization in navigation. AMCL is a probabilistic localization system for 2D mobile robots that implements an Adaptive Monte Carlo localization method to calculates the robot position using particle filters based on existing maps.

3.3. *Coordinate transformation*

The odometry itself can also assist the robot in localization, but there are odometry cumulative errors and some special cases. AMCL can improve thelocalization accuracy by estimating the robot's posture in the map coordinate system, and then combining with the odometry. The odometry itself can also assist the robot in localization, but the odometry has accumulation errors and localization errors may occur in some special circumstances. AMCL can improve the localization accuracy by estimating the robot's posture in the map coordinate system, combined with the odometry. Odometry localization and AMCL map localization is shown in Fig. 5. The simulation of the AMCL algorithm in Rviz is shown in Fig. 6.

- Odometry localization: Only coordinate transformation between /odom_frame and /base_frame via odometry data,
- AMCL map localization: You can provide coordinate transformations between /map_frame, /odom_frame, and /base_frame.

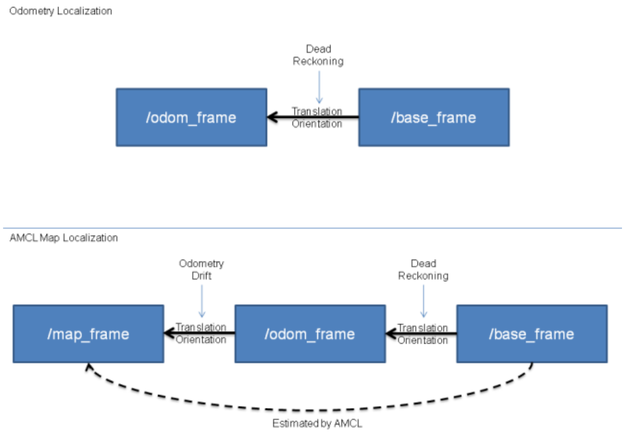


Fig. 5. Odometry localization and AMCL map localization

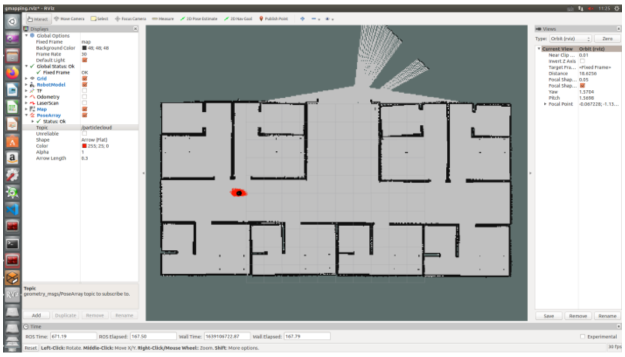


Fig. 6. The simulation of the AMCL algorithm in Rviz

4. Navigation Path Planning

4.1. *Introduction to the move_base*

Undoubtedly, path planning is one of the core functions in navigation. The move_base function package is provided in the navigation function package set of ROS to realize this function. The move_base function package provides an action-based path planning implementation.

move_base can control the robot chassis to move to the target position according to the given target point, and continuously feedback the robot's own posture and the status information of the target point during the movement. move_base is mainly composed of global planner[7]and local planner.[8]

4.2. *Costmap*

Robot navigation relies on a map. The map in ROS is actually a picture. This picture has metadata such as width, height, and resolution. The gray value is used in the picture to indicate the probability of obstacles. However, the map constructed by SLAM cannot be used directly in navigation, because the map constructed by SLAM is a static map. In the navigation process, the obstacle information is changeable. The obstacle may be removed or new obstacles may be added. Obtain the obstacle information from time to time during navigation. It is best to set a warning on the edge of the obstacle on the map. In the area, try to prohibit robots from entering. Therefore, static maps cannot be directly applied to navigation. On top of it, some auxiliary information needs to be added to the map, such as obstacle data obtained from time to time, and data such as inflation layer added based on static maps.

There are two cost maps: global_costmap and local_costmap. The former is used for global planner, and the latter is used for local planner. Both cost maps can be stacked in multiple layers, and generally have the following layer:

- Static map layer: Static map built by the SLAM,
- Obstacle map layer: The obstacle layer tracks the obstacles as read by the sensor data,
- Inflation layer: Inflate on the above two layers to avoid the robot from hitting obstacles,
- Other layers: Other layers can be implemented and used in the costmap via pluginlib.

4.3. *Collision algorithm*

Inflation is the process of propagating cost values out from occupied cells that decrease with distance. For this purpose, we define 5 specific symbols for costmap values as they relate to a robot. Collision algorithm is show Fig. 7.

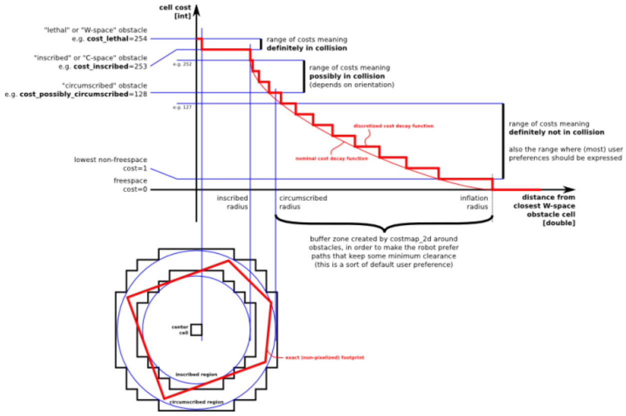


Fig. 7. Collision algorithm

- "Lethal" cost means that there is an actual (workspace) obstacle in a cell. So if the robot's center were in that cell, the robot would obviously be in collision,
- "Inscribed" cost means that a cell is less than the robot's inscribed radius away from an actual obstacle. So the robot is certainly in collision with some obstacle if the robot center is in a cell that is at or above the inscribed cost,
- "Possibly circumscribed" cost is similar to "inscribed", but using the robot's circumscribed radius as cutoff distance. Thus, if the robot center lies in a cell at or above this value, then it depends on the orientation of the robot whether it collides with an obstacle or not. We use the term "possibly" because it might be that it is not really an obstacle cell, but some user-preference, that put that particular cost value into the map. For example, if a user wants to express that a robot should attempt to avoid a particular area of a building, they may inset their own costs into the costmap for that region independent of any obstacles. Note, that although the value is 128 is used as an example in the diagram above, the true value is influenced by both the inscribed_radius and inflation_radius parameters as defined in the code,
- "Freespace" cost is assumed to be zero, and it means that there is nothing that should keep the robot from going there,
- "Unknown" cost means there is no information about a given cell. The user of the costmap can interpret this as they see fit,
- All other costs are assigned a value between "Freespace" and "Possibly circumscribed" depending on their distance from a "Lethal" cell and the decay function provided by the user.

4.4. *parameter setting*

In the simulation, it may happen that the robot enters the expansion area when the local planner does not conform to the global planner and appears to "feign death." In order to avoid this situation as much as possible. Through the physical model of the robot itself, the parameters set by the global planner and the local planner can be changed. In this way, in the global planner, the planner will be as far away from the obstacles as possible, and in the local planner, even if the robot deviates from the global planner, it will retain more free space between the obstacles, thereby avoiding the "feign death" situation.

Set the maximum speed in the x direction in the basic local planner parameters to 0.3 m/s, configure the expansion radius of the global costmap to 0.3m, and the expansion radius of the local costmap to 0.1m. Set the size of the local cost map to 3m * 3m * 3m. Set the obstacle perception range to 2m, and eliminate the obstacle range after it is greater than 2.5m.

5. Conclusion

First start the Gazebo simulation environment; load the launch file related to the startup navigation; load the Rviz component with the added configuration data. Navigate by setting the destination through the 2D Nav Goal on the Rviz toolbar. The action trajectory is shown in Fig. 8. The green line represents the global planner, and the red line represents the local planner. For the obstacles that appear on the global planner route, the simulation will avoid the obstacles through local planner and guide the robot to reach the target point. Obstacle placement is shown in Fig. 9. Local planner avoid obstacles is shown Fig. 10. Reaching the final target point is shown Fig. 11.

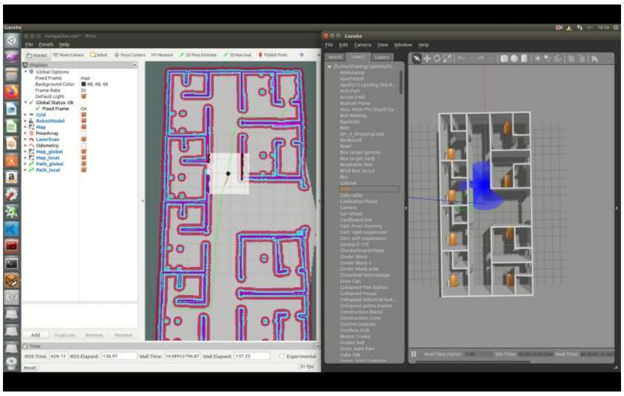


Fig. 8. The action trajectory

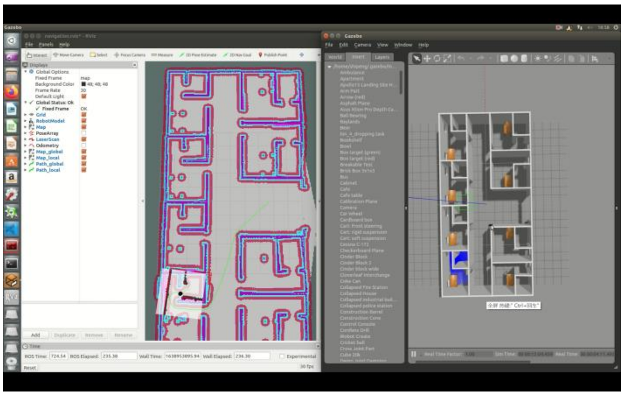


Fig. 9. Obstacle placement

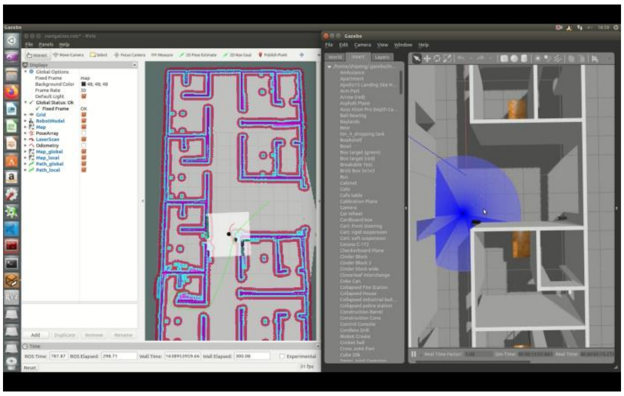


Fig. 10. Local planner avoid obstacles

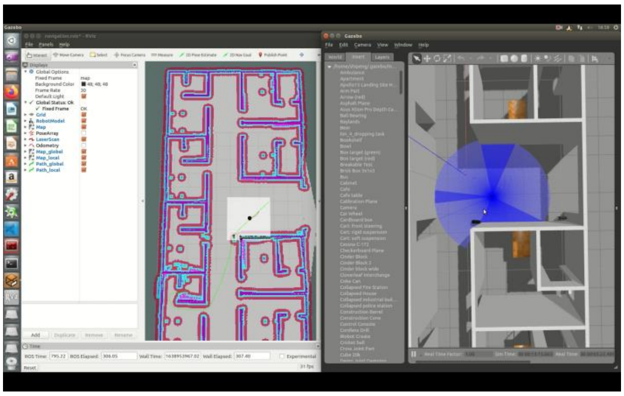


Fig. 11. Reaching the final target point

References

1. Koubâa, Anis, ed. *Robot Operating System (ROS)*. Vol. 1. Cham: Springer, 2017.

2. G. Grisetti, C. Stachniss and W. Burgard, "Improving Grid-based SLAM with Rao-Blackwellized Particle Filters by Adaptive Proposals and Selective Resampling," *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, 2005, pp. 2432-2
3. Y. Abdelrasoul, A. B. S. H. Saman and P. Sebastian, "A quantitative study of tuning ROS gmapping parameters and their effect on performing indoor 2D SLAM," *2016 2nd IEEE International Symposium on Robotics and Manufacturing Automation (ROMA)*, 2016, pp. 1-6
4. W. Xiaoyu, L. Caihong, S. Li, et al. "On Adaptive Monte Carlo Localization Algorithm for the Mobile Robot Based on ROS," *2018 37th Chinese Control Conference (CCC), 2018*, pp. 5207-5212.
5. J. Röwekämper, C. Sprunk, G. D. Tipaldi, et al.On the position accuracy of mobile robot localization based on particle filters combined with scan matching, *Intelligent robots and systems*,2012:3158-3164.
6. Gilli é ron, Pierre-Yves, et al. "Indoor navigation performance analysis." *Proceedings of the 8th European Navigation Conference GNSS*. No. CONF. 2004.
7. Huijuan Wang, Yuan Yu and Quanbo Yuan, "Application of Dijkstra algorithm in robot path-planning," *2011 Second International Conference on Mechanic Automation and Control Engineering, 2011*, pp. 1067-1069.
8. Eduardo J. Molinos, Ángel Llamazares, Manuel Ocaña, Dynamic window based approaches for avoiding obstacles in moving, Robotics and Autonomous Systems, Volume 118, 2019, Pages 112-130, ISSN 0921-8890.

Authors Introduction

Mr. Peng Shi

He received a bachelor's degree in automation from Tianjin University of Science and Technology in 2019. He currently holds a master's degree in electronic information from Tianjin University of Science and Technology.

Dr. Yizhun Peng

He is an Associate Professor in Tianjin University of Science & Technology. He received a doctor's degree in control theory and control engineering from the Institute of Automation,Chinese Academy of Sciences, in 2006. His research field is intelligent robot and intelligent control.

Design of a Humanoid Dance Robot for Dancing Baduanjin

Zhuofan Xu, Ruitao Li, Binfu Zhong, Yizhun Peng*

College of Electronic Information and Automation, Tianjin University of Science and Technology, 300222, China;

*E-mail: * pengyizhun@tust.edu.cn*

www.tust.edu.cn

Abstract

This paper introduces a humanoid robot which consists of 16 servo motors driving joint and adopts STM32 as the main control chip and dual mode Bluetooth4.0 BLE+EDR wireless control.by programming, the motor of humanoid robot can reach the specified space position and attitude at the specified time. At the same time, digital closed-loop control PID algorithm and synchronous compensation algorithm are used to modify and optimize the Angle of the servo motor in real time, and the designed action is vividly displayed.

Keywords: Humanoid robot, STM32, PID, Baduanjin

1. Introduction

Baduanjin is a traditional fitness qigong in China. Whether from the perspective of traditional Chinese medicine, or from the successful experience of many makeshift hospitals in Wuhan in the rehabilitation of COVID-19 patients, it has been shown that regular practice of Baduanjin can enhance the lung and respiratory function and enhance the body's resistance. To a certain extent, help COVID-19 patients restore their lung breathing function.

At present, the COVID-19 prevention situation is still very serious around the world. Using robots instead of medical staff to guide patients to practice Baduanjin on a regular basis can effectively reduce the risk of infection of medical staff and help patients recover their lung breathing function[1].

This paper introduces a humanoid robot which is composed of 16 servo motors, and uses STM32 as the main control chip and dual mode Bluetooth4.0BLE+EDR wireless control. [1]Through the coordination of the upper computer and the lower computer, the motor of the humanoid dancing robot can reach the specified space position and posture at the specified time. At the same time, digital PID algorithm and synchronous compensation algorithm are used to correct the servo motor Angle in real time, so that baduanjin can be vividly displayed and make an important contribution to the world's fight against COVID-19.

2. System Overview

The device is a humanoid dancing robot combined with artificial intelligence technology. The robot can dance independently according to the choreography. Humanoid dance

The robot's voice recognition and processing and big data processing technology are integrated. Through self-designed PCB, its distributed framework can build Bluetooth communication module, motor drive module and so on.It can realize three characteristic functions: first, it can imitate human dances, and realize the choreography of various dances based on different kinds of dances. At the same time, it can set the duration, so that people of all ages can enjoy the performance better;Second, the independent dance, cooperate with each other based on the upper machine and lower machine can make better humanoid robot to dance dance choreography and start and stop, without human intervention in the process of dancing can complete set a good time dancing, at the same time we have already set some dance as human-computer interaction reference is more intelligent and strong entertaining;Thirdly, Bluetooth interaction. Based on the existence of Bluetooth communication module, our mobile phones can also control the start and stop of the robot, and see the status of the robot as well as the choreography of the dance, which can realize edutainment and fun, and better and more convenient to learn dance movements.

3. The hardware structure design

The core of the robot is the combination of upper and lower machines, and the PID algorithm makes the Angle and speed of the servo motor more accurate.The hardware is composed of self-designed PCB, Bluetooth communication module, servo motor, 2200mAh lithium battery, encoder and so on.The encoder is installed on the servo motor, used to calculate the rotation Angle of the servo motor, and determine the rotation Angle from the mechanical structure, software program, upper and lower machine control;The driving part is combined with the main control on a PCB, which can make the wiring simple, not messy, and better observe the action of each joint of the robot, to achieve independent movement;The Bluetooth module can display the robot's status on the phone, and the phone can also control the robot's start and stop.

3.1. *Shell modeling part*

The structure of humanoid dancing robot is mainly divided into three parts: upper body, waist and lower body.The upper body is mainly composed of the head, shoulders and arms. The head is made of a white ABS material, and the shoulders and arms are composed of 6 servo motors. The shoulders and arms can be turned 180 degrees to complete relatively difficult movements. The waist is mainly composed of two servo motors, which can make the waist stronger and have stronger flexibility. The lower body consists mainly of legs and feet, connected by eight servomotors, which can support the weight of the whole body.Can complete the splits, high legs and other difficult movements, better personification.The robot shell material is mainly made of ABS plastic material and aluminum alloy. Because the dancing robot needs to have good anti-fall ability, the combination of ABS plastic material and aluminum alloy is not easy to deform after heavy fall.The battery and PCB are fixed on ABS plastic material so that the PCB and battery are not easy to fall off and play a protective role.Fig.1.shows the full picture of the humanoid dancing robot.

Fig.1. Full view of dancing robot

3.2. *Design of PCB*

It includes the main control part, power module, motor drive voltage regulator module, Bluetooth module.The core chip of dancing robot is single chip microcomputer STM32f103.This microcontroller is used as the main control. Compared with other microcontrollers, STM32 series has high performance, low cost, low power consumption of embedded applications, specially designed ARM Cortex-M3 kernel. According to these advantages, STM32 is used as the core of the robot to design the schematic diagram and printed circuit board. Fig.2. shows the schematic diagram of the main control part.

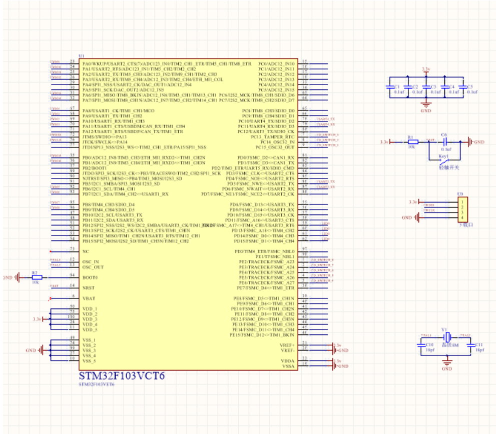


Fig.2. Schematic diagram of main control circuit

The power module adopts 2200MAh lithium battery, which has small internal resistance, large capacity, more cycles and no memory effect. Internal protection design, can prevent charging time is too long, excessive discharge, instantaneous current is too large damage, because the protection of the battery shell will not be more serious impact, can increase the battery life, so that the dancing robot can continue to dance for 2 hours.Fig.3.shows the schematic diagram of the power module.

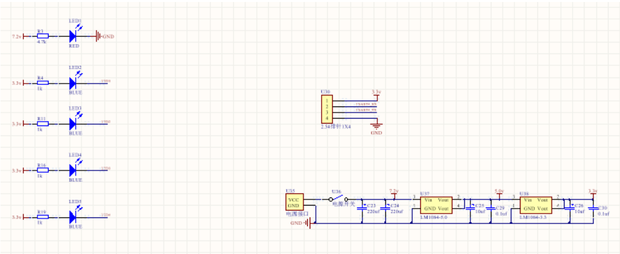


Fig.3. Power supply The power supply circuit

Motor drive voltage regulator module is mainly composed of voltage regulator chip, pin row and bus driver chip, bus driver chip using 74LVC245.Fig.4.shows the function block diagram of 74LVC245.

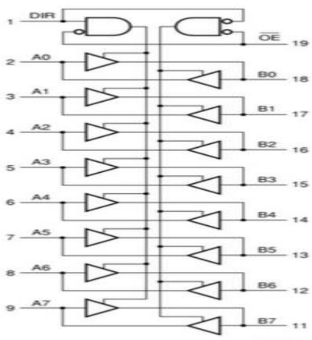


Fig.4. Block diagram of display functions for 74LVC245

Voltage regulator chip we use MIC29302A, stable output, large input voltage range, can better adapt to our servo motor.Fig.5.shows the principle diagram of the voltage regulator driven by the motor.

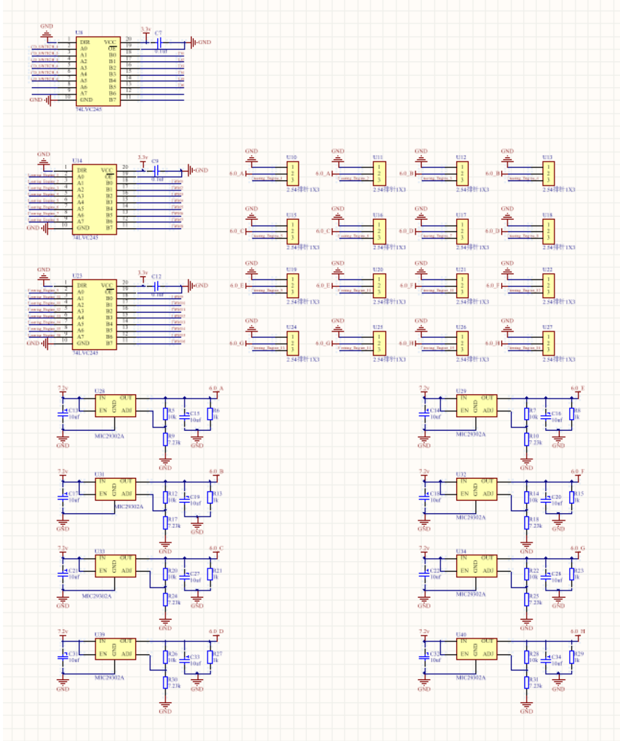


Fig.5. Motor driving voltage regulator module

Bluetooth module adopts bluetooth 4.0BLE +EDR dual-mode Bluetooth control this Bluetooth has many advantages:

- Fast transmission speed
- Low frequency modulation, strong anti-interference
- It is only woken up when execution requires it Low latency
- Wide connection range
- High security

Not only can the robot dance autonomously, it can also start and stop the robot via Bluetooth, which can be both entertaining and addictive. This is the schematic diagram of our Bluetooth module as shown in Fig.6.

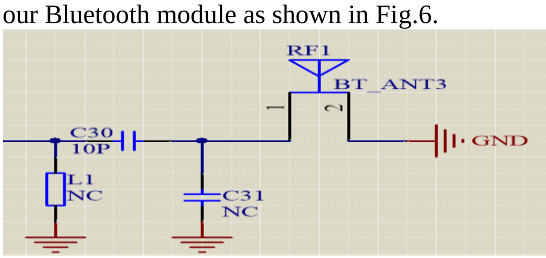


Fig.6.Schematic diagram of the Bluetooth module

3.3. *Cartesian spatial position control of humanoid dancing robot*

We independently design of humanoid robot dance by PCB design control motor rotation, for robots, position control is the precondition of motion control, and need in a certain period of time make the joint to the specified location, position control includes two aspects, respectively is two aspects of the cartesian space and joint space position control. Cartesian open loop control is shown in Fig.7.

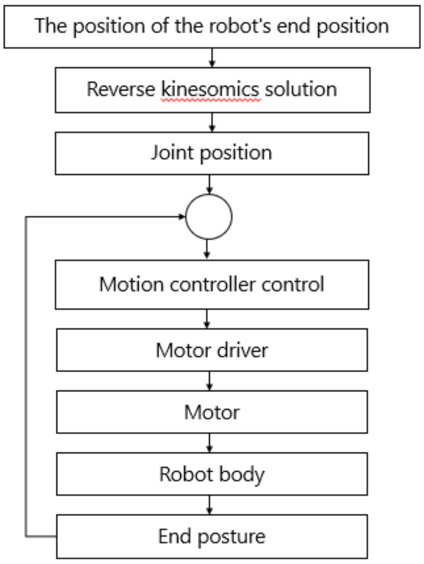


Fig.7.Cartesian open loop control

The problem of the robot's end pose was solved by inverse kinematics, and the coordinate of the end joint in cartesian space was obtained, and the point value of the position on a single joint controller was formed according to the modified coordinate.The joints of the robot are controlled by the speed closed-loop and position closed-loop systems to control the robot to move according to the specific attitude and position.The speed loop is located in the inner loop, while the position loop is located in the outer loop.

Cartesian coordinate system has the following characteristics: two number lines intersect at the origin and have equal units of measure, which can form a plane radial coordinate system.When these two numbers are perpendicular, we call them cartesian coordinates, and when they are not perpendicular, we call them Cartesian bevel coordinates.Cartesian coordinate system and cartesian coordinate system have similarities and differences, the similarity is that cartesian coordinate system and cartesian coordinate system are the description of the position of a point in space, under certain conditions, cartesian coordinate system can be transformed into cartesian coordinate system, cartesian coordinate system can also be transformed into cartesian coordinate system.And the difference is that cartesian coordinates not only have right angles but they can be non-right angles, and they can describe positions in space in a variety of ways including three dimensional coordinate matrices and position vectors.Fig.8.shows the Cartesian coordinate system.

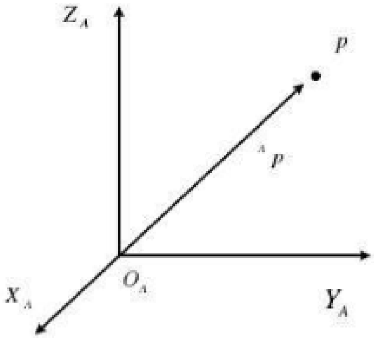


Fig.8.Cartesian coordinate system

3.4. *The PID algorithm*

In order to make the humanoid robot dance better, more accurate completion of the action, motor control usually uses PID algorithm, PID refers to the proportion, integral and differential control, after getting the system output will be output after proportion, integral and differential three operation modes, to overlay the input to make it up to the specified area and control motor position makes dance moves more accurate.Fig.9.shows the PID algorithm block diagram.

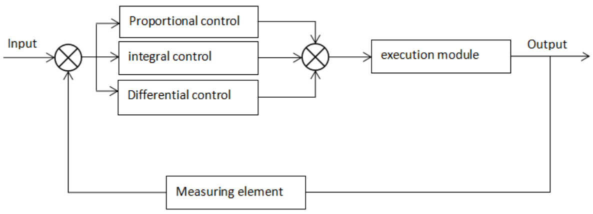


Fig.9.PID algorithm block diagram

3.5. *Motor*

Dance humanoid robot is can complete a variety of difficult moves, because each joint has equipped with motor, commonly used in present dancing robot motor mainly has three categories: dc motor, servo and stepper motors, they all have obvious advantages and disadvantages, including dc motor with low power consumption, torque, is often used in closed loop control system[5].Steering gear control precision,small torqueStepper motor control precision, large power consumption, small torque, is often used for open loop control system, which has obvious advantages and disadvantages of three kinds of motor, we adopt servo motor to dance for our robots "joint", realized the advantages of the servo motor position, speed and torque closed loop control, overcome the out-of-step problem of step motor, And strong overload resistance can withstand more than three times the rated torque of the load, and stable operation at low speed can not produce similar stepping operation phenomenon at low speed. Figure 3-10 shows the servo motor distribution of humanoid dancing robot.

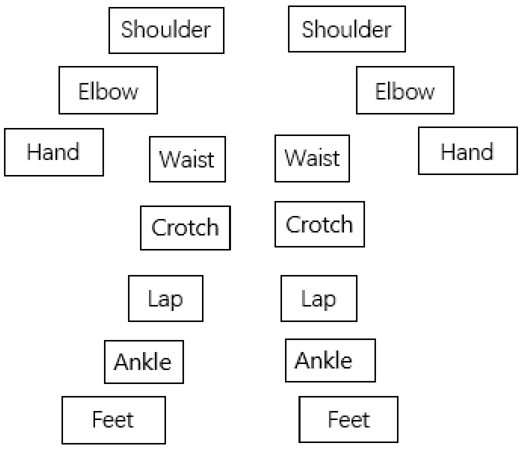


Fig.10.Distribution of servo motors

4. Software design

For humanoid dancing robot, its control system includes two aspects, namely hardware system and software system. Hardware system is the basic system of the overall control, while software system is based on the hardware system.The software system consists of two parts: the upper computer software system and the lower computer software system.In the dynamic debugging of upper computer software, edit and debug the robot dance movements, and record the coordinate data information of each joint position to form a data information table about dance movements, and then send the position data information in the table to the lower computer system.According to the sent data information, the robot movement control to form a dance.Fig.11 shows the flowchart

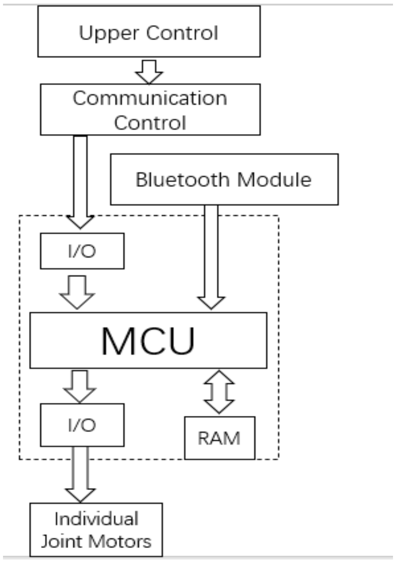


Fig.11.Control system block diagram

PC control dynamic debugging software dance robot movement there are two main ways: one is the online control, dynamic debugging by dragging the upper machine software on a joint corresponding sliding block can be directly observed the rotation to the Angle of the corresponding position, the real-time online watch dance robot joint position, easier to detailed debugging dance moves. The other is offline control, which is actually based on online control. That is to say, a dance posture should be adjusted by online control first, and the control data of each joint of the dance posture should be saved in the dance movement data sheet as a record of the movement table. Then a complete set of dance movements is an action table composed of multiple records. The data in the action table is downloaded to the lower computer, and the single chip microcomputer controls the dance robot according to the data in the action table to complete a complete set of dance movements. When using this control method, after the data in the action table is downloaded to the single chip microcomputer through the serial port, the serial port line can be unplugged to enable the dancing robot to complete the dance performance independently under the control of the single chip microcomputer, so it is called offline operation.

5. Other parts of the humanoid dancing robot

5.1. *Costume design of dancing robot*

The costume design of the humanoid dancing robot is mainly for the following aspects:

- Deepen the theme so that the audience can remember the performance.
- Beautiful, enhance the beauty of dancing.
- Increase the adaptation and story with music.

The costumes of our humanoid dancing robot are independently designed, including lotus, auspicious clouds, wild cranes, tai Chi and many other elements, which are closer to the theme of the dance and deepen the meaning and connotation of the dance.Fig.12.shows the clothes.

Fig.12. Self-designed clothes

5.2. *Music and background design*

The AU software is used to gradually decompose each action frame of Baduan Brocade for music design, so that its actions can better fit each frame to achieve better dance performance and stage effect.

The background uses Chinese ink style tai Chi background video to make the dance more coordinated and story-telling. Resonates with the dress and dance itself, making it more vivid.Fig.13.shows the background design.

Fig.13. Background design

5.3. *Dance background and artistic analysis*

Baduanjin is a traditional Chinese health qigong. Regular practice of baduanjin can strengthen tendons and bones, dredge meridians and collaterals, improve gastrointestinal and cardiovascular functions, and improve human flexibility[3].As a product of high-tech development, robots can reflect a country's scientific and technological level to a certain extent. Eight jin as a basis, we believe that the eight period of jin is the treasure of the Chinese traditional culture, has a strong body, enhance the effect of cardiopulmonary function, especially in the outbreak of the new champions league now spread so quickly under the condition of constant, eight period of jin can not restricted by space, eight jin long do improve cardiopulmonary function well.

After the adaptation of the action to make it more ornamental and interesting, through the design, production, programming, debugging to make the baduan brocade vivid display, so as to achieve the combination of Chinese culture and modern science and technology [4].

6. Test result

After our test, the humanoid dance robot can complete independent dance, and can complete the baduanjin performance according to the dynamic debugging software of the upper computer, the dynamic debugging software of the lower computer and the PID algorithm. Fig.14. shows the static figure of the dance.


Fig.14. Static display of dance

7. Conclusion

Humanoid dancing robot system is relatively perfect, in the design aspect of machinery, control, music, choreography, etc., need to integrate and deepen each other, in the core design of the robot, the control system is the most critical, its design will have a certain impact on the robot itself. In the robot dance movement, the corresponding digital algorithm is used and the Cartesian coordinate system is established. Closed-loop feedback control can be used when positioning the robot walking path. In addition, synchronous compensation algorithm and PID algorithm can be used for walking and other movement problems.

The software and hardware structure of humanoid dancing robot is established, the coordination control system of robot music and dance movement is established, and the related strategies of music feature recognition are explained to realize the related control system of dance robot.

After the robot shape and system are assembled, some basic movements can be carried out according to the input system. However, dance robots are not human beings, so the center of gravity needs to be considered. Some movements that robots cannot complete cannot be forced to be completed.Due to the lack of innovation in related technologies, robots still have the following problems:The system computing capability is insufficient.The procedure of the whole robot adopts the method of single chip computer operation.Although the single chip microcomputer has a qualitative leap over the past, the internal computing system of humanoid dancing robot is relatively large, and the existing single chip microcomputer is difficult to perfect the real-time information of the robot. Based on the above problems, we believe that the main research directions in the future are as follows. Use the controller applicable to the RTOS. The full name of RTOS controller is real-time Operating System, which is used to deal with real-time situations. It has high practicability, timeliness and flexibility while processing data. With the rapid development of technology and technological capabilities, microcontrollers have made continuous progress. It is believed that dance robots with RTOS will have higher flexibility and adaptability. We can install infrared photoelectric sensors in the chassis of the robot to detect the occurrence of the actual situation. The sensors can be divided into two groups, and the signals of each group of sensors are connected with the MCU by input. When the robot performs, no matter which area enters into the effective range of infrared ray, high level can be transmitted in time. And, or and NOT gate will be used to convert high level to low level and finally external interruption will be made to avoid rollover phenomenon.

References

1. Mai Weizha, Liang Huiyuan, TAN Yuxin, Li Zhiheng, Chen Daorui. A case study of the influence of traditional culture on mental health . Electronic technology,2020,49(10):90-91.
2. Dong Xin. Design and development of dance robot in cultural communication. Mechanical design, 2021, 38 (4) : 153. DOI: 10.13841 / j.carol carroll nki JXSJ. 2021.04.032
3. Yu Xiaomeng, Gong Xianghui, Song Fengxiang, Liang Yuanfeng, Zhang Penghui. In the new era, the scientific nature of baduanjin is discussed based on indoor fitness regimen. Research on Innovation of Ice and Snow Sports,2021(08):181-182
4. Luan Minghui, Li Songsong, He Huimin, Zhang Qi, Yang Ying, Li Chen. Design of a new dancing robot. Electronic and information technology, 2021, 5 (6) : 223-224 + 227. DOI: 10.19772 / j.carol carroll nki. 2096-4455.2021.6.098
5. TANG Jiawen. Design and implementation of small dancing biped robot. Horizon of science and technology, 2020 (9) : 177-179. The DOI: 10.19694 / j.carol carroll nki issn2095-2457.2020.09.069
6. Zhang Peiyan, LUO Shuai. Design of a dance robot control scheme. Henan Science and Technology,2019(31):26-28.
7. HUANG Bing. Principle and application of microcomputer. Chongqing: Chongqing University Press,2001.
8. CAI Zixing. Robotics. Beijing: Tsinghua University Press,2000.9.1.

9. Zhou Yuanqing, Zhang Zixing, Xu Wanyong, Jia Peifa, et al. Intelligent robot system. Beijing: Tsinghua University Press,1988.
10. Encyclopedia of Computer Science and Technology (Selected Edition). Tsinghua University Press, 2002.
11. Yu Yongquan.Power interface technology of single chip microcomputer application system. Beijing: Beihang Press, 1991.
12. Luo Yijing.Multimedia computer software. Chengdu: University of Electronic Science and Technology of China, 1991.
13. Gao Haisheng et al. Monolithic integrated circuit application technology.Chengdu: Southwest Jiaotong University Press, 1996.
14. He Limin.Single-chip microcomputer application technology selected 1/2/3.Beijing: Beihang Press, 1992.
15. Li Bingcao.MCU interface technology and its application in MCU.Shaanxi: Shaanxi Electronic Editorial Department, 1993.

Authors Introduction

Mr. Zhuofan Xu

He is an undergraduate majoring in robotics engineering at Tianjin University of Science and Technology. The research direction is intelligent robot, proficient in AD design and circuit simulation

Mr. Ruitao Li

He is an undergraduate majoring in robot engineering in Tianjin University of science and technology.His research direction is robot simulation technology, proficient MCU development

Mr. Binfu Zhong

He is an undergraduate majoring in robot engineering in Tianjin University of science and technology.His research direction is robot simulation design, proficient in the development and use of upper computer software and lower computer software

Dr. Yizhun Peng

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 Shading System Suitable for Parenting Products

Yihan Yan, Songyun Shi, Yanxin Shen, Yizhun Peng*

College of Electronic Information and Automation,
Tianjin University of Science and Technology, 300222， China;

*E-mail: * pengyizhun@tust.edu.cn*

www.tust.edu.cn

Abstract

With the development of science and technology, if you want to provide a more safe and healthy growth environment for babies, simple human protection is far from enough. And as smart home is gradually accepted by the public, smart home does bring people a more convenient life In daily life, proper exposure to the sun can help nutrients be quickly absorbed and used, so that babies can better explore this new world, but direct sunlight will bring great harm to babies' eyes This product is an intelligent shading system, which can be connected with the smart home system and the Internet. Based on specific seasons, time periods and different geographical locations, the position of the sun and light intensity can be determined, so as to adjust the shading Angle in time, so that babies can bask more scientifically.

1. Introduction

Since ancient times, people always yearn for light, hate the dark, from primitive man to the invention of Edison electric lamp, people always use a variety of ways to get light. If you want to give babies a really safe growth environment, it is difficult to achieve by human care alone, but the rational use of science and technology will create a more scientific and safe environment for babies. However, with the development of modern society and economy and the progress of science and technology, the unreasonable use of light has brought a terrible disaster -- light pollution, which will not only harm adults, but also affect the healthy growth of babiesp[1].

Numerous studies have linked light pollution to high rates of myopia and cataracts in humans. The human eye produces visual scene just like the camera takes a photo. The structure of the eyeball from the cornea to the crystal is like the lens, and the retina is like the film. The error of any structure will lead to visual disorder, such as the photo will not be clear, the photo will not be real or virtual, or the film will not be exposed and displayed. The person works and lives in the white bright environment for a long time, each structure of the eye will be affected, the function of retinal photoreceptor cells will be inhibited, causing visual fatigue and sharp decline in vision, and the incidence of cataract increases greatly.

Additional, a few parents to build showily, sweet domestic atmosphere, the choice is colourful and very the lamp that grabs an eye is acted the role of and bright dazzling pink wall. But do not know, it is these "dazzling light" harm their own or even children's vision health. The strong light that alternates light and shade stimulates eyeground in turn to make optic nerve fatigue, cause vision to drop. Someone has done an interesting survey: before the age of 2, children who have the habit of turning on the light to sleep, the myopia rate is 55%, while children who have the habit of turning off the light to sleep, the myopia rate is only 10%. Therefore, parents should pay more attention to the protection of babies' eyes, and the intelligent shading system can help parents pay attention to the problem of light exposure to babies.

2.The Hardware Design

And the invention is a smart baby bassinet adaptive intelligent lighting system, can be connected with the smart home system, and links to the Internet, more specific season, time, different geographical position, to determine the position of the sun, and the light intensity, so as to timely adjust the Angle of window, let the baby more scientific in the sun[2].

The invention consists of STM32F767IGT6 minimum system, AP3216C light environment sensor, MG995 steering gear, 7-inch RGB color touch screen, 4-pin button and cradle curtain. The COMMON IO port of STM32F767 simulates the IIC timing sequence and drives the AP3216C optical environment sensor. Thus detecting ambient light intensity (ALS).

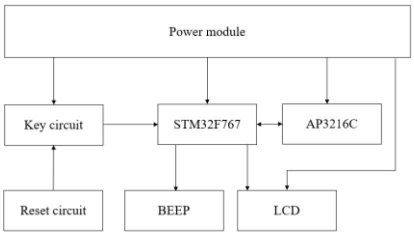


Fig.1.Hardware design drawing

2.1.*Embedded main control chip*

The STMicroelectronics STM32F7 32-bit MCU+FPU uses a high-performance ARM® Cortex® -m7 32-bit RISC core, operating at up to 216MHz. The Cortex-M7 kernel has single-floating-point unit (SFPU) accuracy, supports all ARM single-precision data processing instructions and data types, implements a complete DSP instruction set and memory protection unit (MPU) for enhanced application security[3].

Features, kernel: ARM 32-bit Cortex-M7 CPU with FPU, Adaptive Real-time accelerator (ART Accelerator ™), and L1 cache (4KB data cache and 4KB instruction cache), So there is no need to wait when executing code from embedded flash and external memory), operating frequency up to 216MHz, MPU, 462 DMIPS/ 2.14DMIps /MHz (Dhrystone 2.1) DSP instruction and other features[4].

It has up to 1MB of flash memory, and has 1024 bytes of OTP memory, SRAM: Flexible external memory controller with up to 32 bit data bus flexible external memory controller with up to 32 bit data bus SRAM, PSRAM, SDRAM/LPSDR SDRAM and NOR/NAND memory. And dual-mode Quad SPILCD parallel interface, compatible with 8080/6800 mode. [5]LCD-TFT controller (resolution up to XGA) with dedicated Chrom-Art Accelerator ™ enables enhanced Graphic Content Creation (DMA2D) clock, reset and power management for 1.7V to 3.6V application power and I/O, POR, PDR, PVD and BOR.

Fig.2. STM32F767IGT6 main control chip

2.2.*Host computer design*

AP3216C is a three-in-one environmental sensor that integrates: Ambilent Light Aensors (ALS), Proximity Sensor (PS), and an Infrared Radiation LED(IR LED) The chip is connected to FOGA through IIC interface.

The IIC interface of AP3216C supports up to 400KHz communication rate and supports multiple working modes (ALS, PS+IR, ALS+PS+IR, etc.)[3]. Built-in temperature compensation circuit, operating temperature support -30~80℃. The ambient light sensor has a 16-bit resolution and the proximity sensor has a 10-bit resolution. The infrared sensor has a 10-bit resolution in an ultra-small package (4.12.41.35mm).

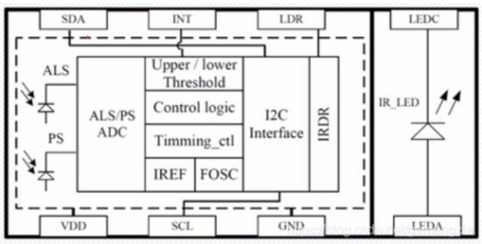


Fig.3.AP3216C diagram

2.3. *The detection system and feedback*

The detection system and feedback system are composed of circulating pump, encoder, 4-pin button, AP3216C optical environment sensor, MG995 steering gear and 7-inch RGB color touch screen. This is also the most important part of the robot[6]. It works as follows：

(1). Sensor initialization: initialize IIC; Reset AP3216C(at least 10ms); Write a byte 0X03.
(2). Read data: loop (6) reads all the sensor data, each cycle of six data in buf [I] in the array, by manipulating the: * als = ((under-16 buf) [3] < < 8) | buff [2]
The ALS value can be obtained (by moving the fourth element of the loop array 8 bits to the left or the third element as the address of the ALS value)
4 pin button and STM32F767 minimum system common IO link, used to set trigger shading threshold, control menu options, etc..
7 inch RGB color touch screen for displaying light intensity, cradle curtain position and other information.
The steering wheel of MG995 steering gear is connected with the cradle curtain, and the curtain can be raised or lowered by rotating.
STM32F767IGT6 minimum system through IO output PWM control steering gear movement, according to whether the ALS value is greater than the set threshold, control steering gear positive and negative rotation or stop.

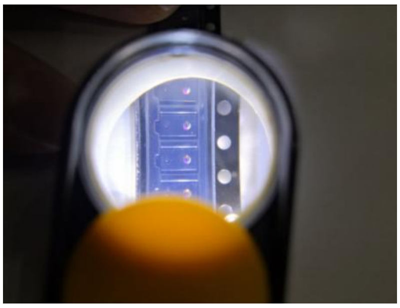


Fig.4.AP3216C

3.Software Design

The decision-making of the robot is completed by Jetson nano sending instructions to the lower computer after data processing and analysis. The lower computer preprocesses the data according to the feedback value given by each sensor module, and then returns it to the upper computer. The software design block shows in Fig.5.

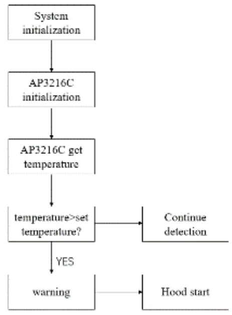


Fig.5.The software design

4.Test

After our test, the robot can simulate the shading of sunlight under normal conditions; And it can simulate the shading situation of different time, different seasons and other environmental factors according to the setting of wireless AD hoc network design.

5.Conclusion

Through the solution of artificial intelligence, study the light induction of shading shed, and combine with the simulation of Internet intelligent learning. The intelligent shading system is developed. It is of great significance in the field of Internet of things and smart home. The invention achieves intelligent shading simulation through light induction, which not only provides a safer guarantee for the growing environment of babies, but also expands the field and provides a new idea for the Internet of things.

References

1.ling-ling xing. The intelligent greenhouse monitoring system based on iot research [J]. Journal of electronics, 2021 (16) : 26-29. DOI: 10.16589 / j.carol carroll nki cn11-3571 / tn. 2021.16.010.
2.DU Kai. Research on pop-up visor for optical System [D]. Graduate University of Chinese Academy of Sciences (Institute of Opto-Electronics Technology),2014.
3.Zheng Zhongyang, Ye Haijun, DING Chunlei. Science and Technology Innovation,2021(31):172-174.]

4. shuai jianping. Application of light sensor in remote control [J]. Industry & technology forum,2014,13(11):53-54.
5.Chen Qing-yuan. Multi-optical induction measuring instrument [J]. Physics Experiment,2002(02):23-25.
6.VEML6035 ambient light sensor [J]. Journal of sensor world, 2019, 25 (08) : 43. DOI: 10.16204 / j.carol carroll nki sw. 2019.08.011.

Authors Introduction

Ms. Yinhan Ya

She is an undergraduate majoring in automation specialty in Tianjin University of science and technology. Her research direction is intelligent robot.

Mr. Songyun Shi

He is an undergraduate majoring in robot engineering in Tianjin University of science and technology. His research direction is intelligent robot, mastering MATLB simulation and robot operating system design.

Ms. Yaxin Shen

She is an undergraduate majoring in robot engineering in Tianjin University of science and technology. Her research direction is robot simulation technology..

Dr. Yizhun Peng

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

Design of Intelligent Personalized Nutrition Supplement Machine

Xinpeng Yang[2],Sidan Liu[1], Lirui[1], Junhui Yin[1,*],Yizhun Peng[2,*]

[1]*School of Economics and Management,Tianjin University of Science and Technology, Tianjin, China*
[2]*School of Electronic Information and Automation,Tianjin University of Science and Technology, Tianjin, China*
*E-mail: *junhuiyin@126.com,*pengyizhun@tust.edu.cn*
www.tust.edu.cn

Abstract

Reasonable nutritional supplement is the necessity for growth and development of living organisms. Residents need various nutrients daily to maintain bodies. However, in the post-epidemic era, people's physical and mental pressure has increased due to the fast life pace, and most Chinese residents are sub-healthy. To improve the current incomplete and untimely citizen nutrition supplement mode, this work uses big data, artificial intelligence and Internet of things to design an unmanned intelligent personalized nutrition supplement machine. Combining individual body data automatically, the machine can prepare exclusive nutritional supplement drinks timely, avoid crowd contact and support the citizen health quality improvement with effective software and hardware.

Keywords: Nutritional supplement; The Internet of things; Embedded system; Mini-program

1. Introduction

National health is related to the development of national quality and the promotion of economic level. Under the impact of COVID-19, society has a deeper understanding of health and healthcare, and people's awareness of their own physical fitness has been significantly improved[1]. More and more people have realized the importance of reasonable nutritional supplements. In the field of health care, the health care industry combining big data analysis[2], Internet of things[3], human-computer interaction and other new technologies[4] has become a research hotspot. This research has greatly improved the competitiveness of products, opened up a new mode of reasonable nutrition supplement, and helped epidemic prevention and health security. It has laid a practical foundation for the national and local governments in China to improve the health quality of residents.

2. Nutritional design

We aim to customize recipes for each user scientifically and reasonably according to the Dietary Guidelines for Chinese Residents (2016) from the original unilateral nutrition supplement to the dietitian at home!

2.1. *select (suitable) material*

Kiwi :(can be used as A supplement for vitamins A, C and E, it is rich in vitamins C, A and E. The contents of Vc, Mg and trace elements in kiwi fruit were the highest. At the same time, Ve and Vk contents in Kiwifruit were determined as excellent) [5].Mulberry: can be used as iron and calcium supplement, containing calcium, iron, zinc, selenium and other minerals lacking in human body, which can enhance immunity, promote the growth of hematopoietic cells and promote metabolism, etc.)[6]; Tomato :(can be used as vitamin B supplement, tomato is rich in vitamin A, vitamin C, vitamin B1, vitamin B2)[7]; Apples :(as a zinc supplement) apples are not only rich in essential nutrients for the brain such as sugar, vitamins and minerals, but more importantly, they are rich in zinc.

2.2. *Principle of nutrition scheme ratio*

First, calculate the user's body mass index (BMI) : BMI= weight (kg)÷ (height * height) (meter)

Second, according to the relevant requirements of users to complete the plan

Basis for plan making According to Mifflin St. Jeor Calculator, the formula included height, weight and age, and the calculated results were more suitable for Chinese people's physique. The formula is as follows: Static energy consumption (REE) calculation formula (unit: Kcal) : Male: 10* weight (kg)+6.25* height (cm) - 5* age (Y)+5; Female: 10* weight (kg) + 6.25* height (cm) - 5* Age (Y) - 161

Third, daily total energy consumption (TDEE) calculation formula (unit: Kcal) : REE*1.2: sedentary; Little or no exercise throughout the day; REE*1.375: Light exercise (1-3 days per week); REE*1.55: Moderate physical activity (3-5 days per week); REE*1.725: Strenuous exercise (6-7 days per week); REE*1.9: Overweight activities (extremely heavy activities/sports, heavy physical work or training twice a day such as marathons, races, etc.).

Program development (e.g., fat reduction)

The formula is as follows: Your current weight is: __ kg; Your target weight is: __ kg: The weight you want to lose: Current weight - Target weight = __ kg; Total calories that must be reduced during weight loss: (current weight - target weight) *7700 = __kcal Daily calories that must be reduced: (current weight - target weight) *7700 ÷ weight loss period (days) = __kcal Daily calories that must be absorbed during weight loss: TDEE- [(current weight - target weight) *7700÷ weight loss period (days)] = __kcal[8].

Formulated recipes

According to the Dietary pagoda of Chinese Residents (2016) combined with the food exchange method (taking the diet arrangement of healthy elderly people as an example)

Table 1 Daily diet for adult women (diet provides energy of 2000kcall, suitable for light physical activity level over 18 years old)

Name	Recipe
Breakfast	One bowl of purple rice porridge(35g), egg (50g), milk(300g), stir-fried cabbage with garlic (100g), cashew nuts (10g)
Lunch	Steamed rice (30g), steamed chicken (50g chicken, 50g pumpkin, 20g sweet potato), shredded potato (100g) mapo Tofu(100g), tomato and shrimp soup (50g tomato, 5g shrimp), banana (100g)
Supper	Rice (90g rice, 30g sweet potato), braised crucian carp (50g crucian carp), stir-fried asparagus (100g bamboo shoot), stir-fried spinach (100g spinach) laver soup (5g laver), apple (100g)
Breakfast	Drink plenty of water, 1500-1700ml a day, For added sugar, less than 25g is best . If drinking alcohol, do not exceed 15g. Eat and exercise balance, at least 6000steps per day or 30 minutes of moderate intensity exercise, exercise consumption at least 270kcal.

3. Integrated system design

3.1. *Overall functional framework*

The system consists of embedded equipment, mobile phone client and background server, which integrates TWO-DIMENSIONAL code scanning, network communication technology and asynchronous parallel communication technology, as shown in Figure 1. In the three components, the client can realize two-way data exchange and two-dimensional code scanning one-way data transmission with the embedded system through Bluetooth, two-way data exchange between the client and the background through wireless data network, and embedded devices can send information to the background through GPRS.

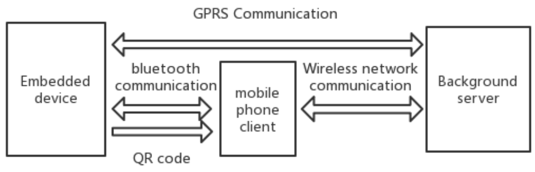


Fig.1 The overall functional framework diagram

When a user queries using machines when the ratio of health, through the mobile phone client posted some qr code scanning machine, machine number, cell phone a client and its synchronization, and then by the user on the embedded system provide the physical data acquisition device, at the same time embedded systems provide data to mobile client hint whether it is right. If the data is not provided correctly, the mobile client will continuously prompt the user until the user provides the data correctly. Then the mobile phone client sends information to the background server for the background server to calculate and process the user's information, calculate the nutrition ratio suitable for the user, and update the user's body data. Finally, the background server will send the user's body data and recommended ratio data to the mobile client for customers to check.

When the user receives the configuration supplement drink, click the "Query" button on the mobile client, the mobile client sends information to the background server, and the background server determines and sends the client identity, body data and proportion information to the client according to the received information. User to confirm the information you have supplied before clicking the "confirm" button, the client ratio for sending information to the embedded devices, embedded devices by accepting information ratio of drinks, wait for after the ratio, the user can be gotten in retrieving the drinks, at the same time, the mobile client will send information to the backend server to update the user insist on drinking nutrition matching the number of drinks and other information.

3.2. *Embedded device design scheme*

Bluetooth technology is a wireless technology standard with radio frequency characteristics. It applies frequency hopping technology and wireless technology, etc., with advantages of high transmission efficiency and high security, it has been applied by all walks of life [9].

Photoelectric sensor is generally composed of two parts: processing path and processing element. Its basic principle is based on the photoelectric effect, the change of the measured into the change of the optical signal, and then with the help of photoelectric elements to further convert non-electrical signals into electrical signals. The Beidou Navigation Satellite System is a global navigation satellite system independently developed and operated by China. Beidou navigation system is an active bidirectional ranging two - dimensional navigation system. The ground central control system is calculated to provide users with 3D positioning data. Beidou is the world's first satellite navigation system to provide three-frequency signal service. Three-frequency signals can better eliminate the influence of high-order ionospheric delay and improve positioning reliability. In addition, Beidou also has its own unique short message communication function[10].

GPRS is the abbreviation of general packet wireless service technology, it is a mobile data service available to GSM mobile phone users, GPRS and the previous continuous channel transmission way is different, it is packet type to transmit, using packet switching technology, compatible with GSM network and more effective transmission of high-speed data and signaling on the network. Users are allowed to send and receive data in end-to-end packet transfer mode without the need for network resources in circuit switch mode[11].

Drive module adopts L298N, L298N is a motor driver that accepts high voltage, dc motor and stepper motor can drive.

Stepper motor controller can send uniform pulse signal of electronic products, it sends the signal into the stepper motor driver, will be converted by the driver into the stepper motor required by the strong current signal, driving the stepper motor operation. The stepper motor controller can accurately control each Angle of the stepper motor.

OpenMV is an open source, low cost, powerful machine vision module. Machine vision algorithms on OpenMV include color block search, face detection, eye tracking, edge detection, marker tracking, etc.

Relay, also known as relay, is an electronic component that controls the circuit. It is composed of two subsystems, the control system and the controlled system. It automatically controls the circuit on and off by induction of the size of the access current. Since relays can control large current through small current, they have been well applied in electrical engineering and automation of low-voltage appliances[12].

LCD touch screen, also known as touch panel, is an inductive LCD display device that can receive the input signal of touch head.

Solar cell, is a kind of photoelectric semiconductor chip that uses sunlight to generate electricity directly, also known as "solar chip" or "photocell", it can output voltage and produce current in a loop as long as it is satisfied with a certain illuminance condition. Solar cells reduce energy consumption.

The embedded end uses stm32F103 series MCU produced by STMICROELECTRONICS as the microcontroller. In addition, Outside the microcontroller, bluetooth communication module, photoelectric detection module, Beidou positioning module, OpenMV module, GPRS module, drive module, stepper motor, LCD touch screen and relay used to control the switch of storage unit are also integrated. The power supply uses solar cells, as shown in Figure 2. Among them, Bluetooth is used to communicate with the mobile phone client, photoelectric detection is used to detect the status of cups, GPRS module and Beidou module are combined for machine positioning, OpenMV is used to identify the TWO-DIMENSIONAL code of the mobile phone client, the drive module is used to drive the stepper motor, and the LCD touch screen is used for better human-machine interaction.

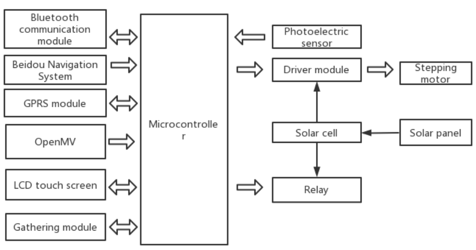


Fig.2 The overall design scheme of embedded equipment

3.3. *Mobile phone client design scheme*

Mobile phone client uses wechat small program, integrated with scanning code, Bluetooth communication, network communication and so on. Among them, Bluetooth is used to communicate with embedded devices, and network communication is used to communicate with the backend server. Users can check their identity information, body data, nutritional ratio recommended by the background server and the number of days they adhere to through the mobile phone client. At the same time, they can also scan the code to confirm their identity information, upload body data and receive drinks.

3.4. *Background server design scheme*

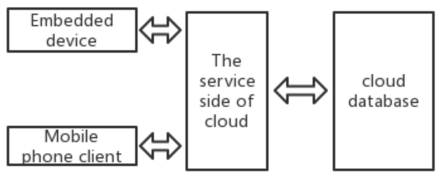


Fig.3 The overall framework of background server

The data processing part of the background server takes Ali Cloud server as the carrier, and the data storage part takes Ali Cloud database as the carrier, as shown in Figure 3. After the user finishes providing body data and receiving drinks, the mobile phone client sends information to the background server through the network, and the background server analyzes the received information and updates the database. At the same time, the background server can also provide the mobile client with the current user identity data, recommended ratio and the number of days to adhere to the ratio plan.

4. System software design

4.1. *Embedded device software design*

The software design of the embedded device (see Figure 4) includes three parts: communication between the embedded device and the mobile phone client, body data collection and detection by the embedded device, and configuration of the embedded device with drinks.

The working process can be divided into two steps: data detection and beverage configuration.

4.1.1. *data detection*

When the machine performs data detection, the mobile phone client sends instructions to the embedded device through Bluetooth in the format of frame header + command. Wherein, the frame header is 0xFF 0xFF two consecutive data, which is used to prompt the embedded processor to enter the data detection mode. The command is an 8-bit data, which represents the specific steps of data detection. There are two steps of data detection, namely, measuring body weight and measuring human resistivity. For example, 0xFF 0xFF 0x01 indicates to remind the embedded device to carry out weight measurement. When the embedded processor receives the instruction, it sends 0x21 to the mobile phone client and 0x22 when it is finished. When the mobile client receives 0x22, it sends 0xFF 0x02, and the embedded device enters the human body resistivity measurement. Similarly, the embedded device sends 0x03 when receiving, and sends 0x04 when completing, and updates various body data of the user to the background server. After receiving 0x04, the mobile client stops sending instructions to the embedded system.

4.1.2.*Configuration drinks*

When the mobile client receives the nutrition ratio data from the background server and the user agrees to transmit it to the embedded device, the mobile client sends instructions to the embedded device in the form of frame header + command. The frame header is 0x05 0x05, which is used to prompt the embedded processor to enter the beverage configuration mode. The command is 4 8-bit data, which is used to represent the amount of dropping of 4 nutritional supplements, such as 0x05 0x05 0x22 0x5F 0xFF 0x26. When the embedded device receives the data, it sends 0x01 to the mobile phone client, which is used to prompt the mobile phone client to receive the current status of the embedded device. When the configuration is complete, send 0x02 to remind the user that the beverage configuration is complete through the mobile client. At this time to detect whether the drink is removed, if not removed every 5s to the mobile phone customer user sends 0x03 once, sends 0x04 when it detects that the drink is taken away, and sends the number of times that the background server and the new user insist on drinking the drink with nutritional ratio, and then restores the initial setting.

4.2. *Mobile phone client design*

WeChat small program referred to as "small program, is a kind of don't need to download and install the use of application, it implements the application" within reach "of the dream, the user through a sweep or search can open the application, it embodies the concept of" used up ", users no longer care about too much installation application will lead to problems such as mobile phone

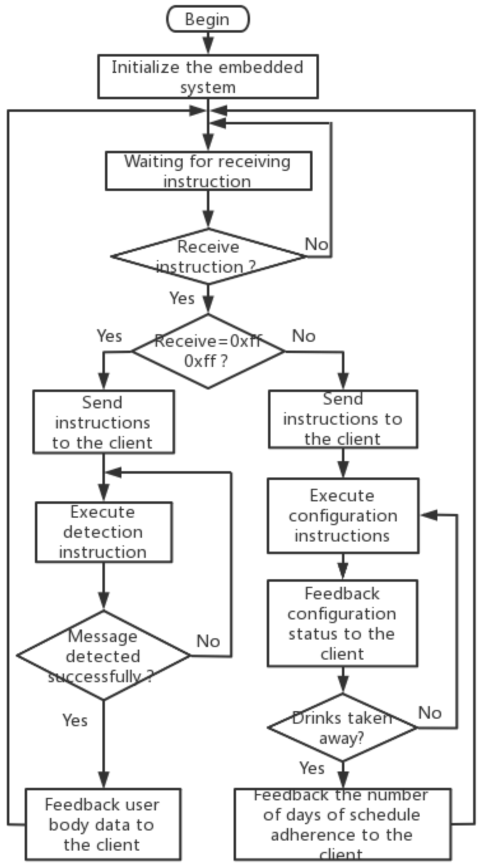


Fig.4 The software block diagram of embedded system

memory, can use at any time, do not need to install the program. For the developers of wechat, the development difficulty of wechat applets is far less than that of APPS. Ordinary users can develop simple and practical applets through simple learning[13].

The mobile phone client mainly provides users with operations such as body data collection, nutrition data query, and self-fetching of drinks. The working process can be divided into two modules of nutrition data query and self-fetching of drinks, as shown in Figure 5.

Before using drinks, users must scan the two-dimensional code on the box of the machine to confirm the current machine number. The two-dimensional code contains the 48-bit MAC address of the Bluetooth module of the machine, which is globally unique and can be used as the number of different machines.

4.2.1. *Nutrition Data Query*

First click on the "nutrition data query" button, enter the authentication interface, the interface has two

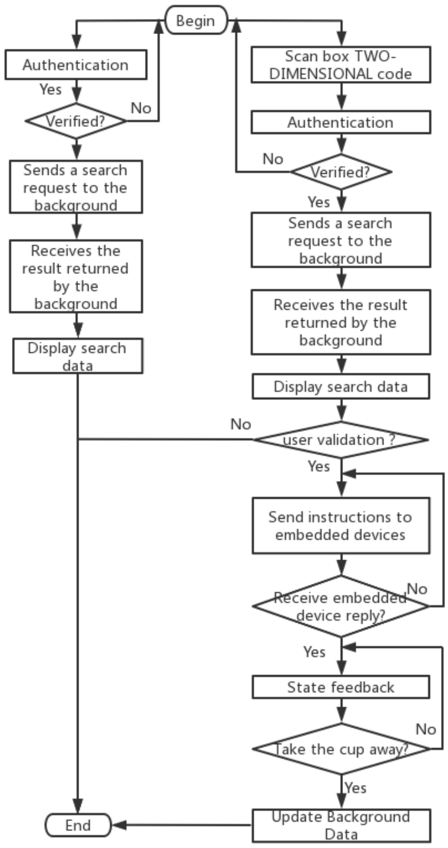


Fig.5 Client program flow chart

input boxes, require the user to enter the name and id number, click on the "validate" button, mobile phone client sends the information to the backend server, if the verification through the access backend server again, for users' personal data body, nutrition matching results and adhere to the matching results the number of drinking, And jump to the next interface, the interface will display the requested information; If the authentication fails, the system returns to the page and prompts the user that the user name or ID number is incorrect.

4.2.2. *Drinks come undone*

Click drinks come undone, just like the above identification process, only when the authentication is successful to back-end server for the number of users insist on drinking drinks and remnant data such as number, and into the drink must display interface, at the bottom of the interface through the "+" and "-" select number of thyself, confirmed, click on the "confirm" button, The mobile client sends configuration instructions to the embedded device. When the configuration is complete, the user is reminded that the configuration is complete and the back-end server data is updated.

4.3. *Background server design*

In the process of Internet application, network server is a very key tool to provide services through computer network system. In the application process, the network server can effectively classify, sort out and save the network information, and also process it according to the needs of users[14].

The background server includes two parts, server and database, which are respectively used to realize data interaction with the client and record the user's body data, recommended ratio data and drinking days. Its program flow chart is shown in Figure 6, which can be divided into two parts: data query and maintenance days query.

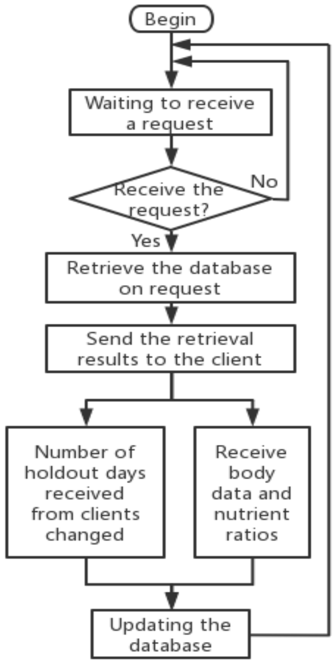


Fig.6 Background server process flow chart

4.3.1.*query and pivot*

When data query, the background first receives the identity information sent from the mobile terminal to start the request, and the server retrives the corresponding user's body data and provides the relative nutritional ratio in the database according to the identity information, and sends it to the mobile terminal. When the user re-provides the physical data, the background will update the database data in real time according to the information sent by the mobile terminal.

4.3.2.*Maintenance days Query*

Maintenance days are proposed as a way to promote enhanced efficacy, and the specific query process is the same as data query.

5. Conclusion

Intelligent personalized nutrition supplement machine based on big data analysis, through the Internet of things technology, software production, intelligent machine operation and other practical operations, to achieve timely and reasonable nutrition supplement, to provide exclusive health drinks for sub-health people, changed the traditional mode of production and sales, to ensure the quality and function of reasonable nutrition. The machine has a wide range of applications, which can be placed in schools, communities, gyms and other areas. With the continuous promotion of the machine, more and more people can experience the convenience of real-time nutrition supplement. The machine will be "invisible to everywhere", making a contribution to the comprehensive nutrition and health promotion.

Acknowledgements

This research was supported by Student’s Platform for Innovation and Entrepreneur Training Program, the Ministry of education of China (202110057034).

References

1. Sui Yue-jiao, JI Tian-yi, BIAN Di, WANG Ying. Exploration and application of the curriculum construction of Diet and Health conservation in Traditional Chinese Medicine under the background of "Great Health" Era [J]. Modern Vocational Education,2021(50
2. Shi Yufei. Thinking on the development and application of health care big data [J]. Wireless Internet Technology, 201,18(19):94-95.
3. Wang Y, Wang Y, Wang Y, et al. The development of the Internet of Things [J].
4. Jiang Liming, Tian Xuetao, Ren Ping, Luo Fang. Artificial intelligence aided by the mental health of new assessment [J/OL]. Psychological science progress: 1-11 [2021-12-13].
5. Heben. Eating kiwifruit benefits more [J]. China Fruit Industry Information, 2021(03):51
6. Nutritional value of Cherry tomatoes [J]. Northern Horticulture, 2019(16):176
7. Han Songlin, ZHAO Dongsheng. Nutritional value of Mulberry and fermentation wine production technology [J]. Food Safety Guide.2021(03):135-136 DOI:10.16043/j.cnki.cfs.2021.03.076.
8. Kou Wenxin. Research on Intelligent Nutrition Catering System and Its Core Algorithm [D]. Beijing University of Technology.2015(03)
9. QIN Qinglei.- Design of an ultra-low power consumption infrared obstacle detection module [J]. Electronic Manufacturing, 2017. (7): 15-16.]
10. Shen L X, PANG J Y, ZHANG Q Y, et al. Design and implementation of location service Platform based on Beidou Navigation [J]. Computer Technology and Development, 2019(1):1-4.
11. Chen K X, XIE H T, GPRS principle and application [J]. Railway Communication Signal, 2003, 39 (7): 7-8. (in Chinese)

12. BI Jialin. Application research of Relay in Electrical Engineering and Its Automatic low-voltage Apparatus [J]. Heilongjiang Science, 2016, 7 (20): 12-13.
13. Tan Yongping, Basic Characteristics and Implementation Strategies of Blended Teaching Mode []. China Vocational and Technical Education, 2018 (32): 5-9.
14. Zhao Chunxia. Research on Security Management and Maintenance Strategy of Network Server [J] computer Programming Skills and Maintenance, 2019 (4): 149-150,172.
15. " RoboCup Federation official website, " , https://www.robocup.org/
16. Y. Yasohara, H. Suzuki, " Development of Omni-directional Mobile Mechanism for RoboCup MSL " , Proceedings of 31st Fuzzy System Symposium, pp. 151-152, 2015. (In Japanese)
17. FLIR, "Flea3(USB3 Vision Camera)", https://www.flir.com/products/flea3-usb3/
18. T. Yoshida, H. Suzuki, "Real-Time Self-Localization for RoboCup Middle-Size-League", Proceedings 32nd Fuzzy System Symposium, pp. 397-398, 2016. (In Japanese)
19. D.E.Goldberg, " Genetic Algorithms in Search, Optimization and Machine Learning", Addison-Wesley, 1989.
20. H. Suzuki, M. Minami, "Visual Servoing to Catch Fish Using Global/Local GA Search " , IEEE/ASME Transactions on Mechatronics, Vol. 10, No. 3, pp. 352-357, June, 2005.

Authors Introduction

Mr. Xinpeng Yang

He is an undergraduate of Tianjin University of science and technology, his research field is embedded system.

Ms. Sidan Liu

Sidan Liu, She is an undergraduate of Tianjin University of science and technology, she research field software design.

Ms. Rui Li

She is an undergraduate of Tianjin University of science and technology, she research field is nutrition supplement program and nutrition

Ms. Junhui Yin

She is an Professor in Tianjin University of Science &Technology. Tianjin first batch of management accounting consulting experts. Her research field is financial management and management accounting theory and practice research.

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 Derivative Oriented Thresholding Approach for Feature Extraction of Mold Defects on Fine Arts Painting

Hilman Nordin
Department of Mechanical Engineering, Faculty of Engineering,
Universiti Malaya, 50603, Kuala Lumpur Malaysia

Bushroa Abdul Razak*
Department of Mechanical Engineering, Faculty of Engineering,
Universiti Malaya, 50603, Kuala Lumpur Malaysia

Norrima Mokhtar*
Department of Electrical Engineering, Faculty of Engineering,
Universiti Malaya, 50603, Kuala Lumpur Malaysia

Mohd Fadzil Jamaludin
Centre of Advanced Manufacturing and Material Processing (AMMP Centre),
Faculty of Engineering, Universiti Malaya, 50603, Kuala Lumpur Malaysia

E-mail: hilmannordin@siswa.um.edu.my, norrimamokhtar@um.edu.my, bushroa@um.edu.my, ibnjamaludin@um.edu.my

Abstract

Identification of mold defects is an important step in the restoration of damaged paintings. The process is usually lengthy and depends heavily on the qualitative visual judgement of an expert restorer. This study proposes an automatic mold defect detection technique based on derivative and image analysis to assist in the restoration process. This new method, designated as Derivative Level Thresholding (DLT), combines binarization and detection algorithms to detect mold rapidly and accurately from scanned high-resolution images of a painting. The performance of the proposed method is compared to existing binarization techniques of Otsu's Thresholding Method, Minimum Error Thresholding (MET) and Contrast Adjusted Thresholding Method. Experimental results from the analysis of 20 samples from high-resolution scans of 2 mold-stained painting have shown that the DLT method is the most robust with the highest sensitivity rate of 84.73% and 68.40% accuracy.

Keywords: image processing, defect detection, Derivative Oriented Thresholding, fine art

1. Introduction

One of the major problem for old paintings is the physical damage on the artwork caused by several factors such as mishandling, high humidity, rapid fluctuations in temperatures and interactions with pollution-dust particles in the air. These factors cause growth of molds, craquelures and other types of wear on the materials especially for improperly stored artworks. If not detected and treated early, it will cause significant irreversible damage to the original painting.

Fortunately, these defects can be evaluated and subsequently repaired by expert restorers. The start of the restoration work consists of assessing the current conditions of the paintings, identifying defects and determining the suitable method of repair. Traditionally, restoration processes are conducted manually, relying heavily on the judgment and skills of the restorers. With the advent of technology, computer-assisted scanning and image processing techniques are available to assist in the restoration process. These techniques use hi-tech detection methods and computer algorithms to objectively emulate the visual assessment and judgment of the restorers.

Imaging technologies for the scanning of 2D artworks includes mass spectrometry[1], photoluminescence spectroscopy[2], x-ray fluorescence analysis[3], and shearography[4]. Once the paintings are scanned, they can be analysed to extract relevant information using techniques such as heuristic graph searching[5], hyperspectral crack detection[6], colorimetry and

Fig. 1 Painting 1 – Ink on Paper

Fig.2 Painting 2 – Ink on Paper

watershed segmentation[7], and user intervention-based detection methods[8, 9].

Advances in imaging technologies in recent years have made it possible for the high-resolution scanning of paintings in its entirety. This technology, known as mesoscopy, enables the recording of details at very high resolutions. Images scanned at high resolution contains accurate colour information that may be discernible to the naked eye. Subsequent processing of the scanned data, such as image segmentations and thresholding, enables identification of features such as defects and inconsistencies. The combined technique of mesoscopy and image analysis have been demonstrated by Win et al [10] in the automatic detection of defects in coated metal specimens for the manufacturing industry. However, the technique has yet to be utilized in the restoration processes of artworks such as drawings and paintings.

This paper proposes a derivative oriented thresholding method for the automated detection of mold on paintings. This method combines two processes of binarization and detection. In the binarization process, the original image is converted into a binary image containing the defects while the detection process identifies the defects using a filtration approach. This newly proposed binarization process is benchmarked with existing thresholding based binarization methods of Contrast Adjusted Thresholding[10], Otsu's Method[11] and Minimum Error Thresholding (MET)[12]. The resulting binary images from these different methods are then filtered in the detection process, where the mold detection results are then compared to a ground truth image. The ground truth image is produced by manually labeling each pixel as either defects or non-defects, based on the feedback from an expert restorer. The accuracy and sensitivity of the detection results are then rated by comparison to this ground truth image.

2. Methodology and Testing

2.1. *Image Acquisition and Sampling*

Two paintings comprising of ink sketches on paper by a well-known Malaysian artist, the late Ibrahim Hussein (1936-2009), were selected as the subject of this study. A Niji-X High-Resolution Scanner (Kyoto University, Kyoto, Japan) was used to scan the paintings at a resolution of 600 dpi. 错误!未找到引用源。igure 1 and Figure 2 show the two paintings, designated as Painting 1 and Painting 2. These two paintings were chosen as they contained mold defects.

For mold detection process, the painting images are divided into smaller image samples. A 200 x 200 pixels capture area was assigned for the sample size, which corresponded to a scanned area of 8.4 mm x 8.4 mm on the actual artwork. Thus, each recorded pixel was approximately 42 μm x 42 μm in size. The sampling process has generated a total of 494 sample images from Painting 1 and 391 sample images from Painting 2. However, for the purpose of this study, only 20 image samples containing mold defects were selected, 14 from Painting 1 and 6 from Painting 2.

2.2. *Image Acquisition and Preparation*

The captured images were then converted into grayscale images for pre-processing. The grayscale image can be expressed in L gray levels $[1,2,\ldots,L]$. Each level consists of m points and the total number of points, *M*, is the sum of m_j where *j* are the individual levels by: $M = m_1 + m_2 + \ldots + m_L$. In the grayscale images, the mold defects are expected to have gray levels value, *s*, between 0 (black) and L (white). Figure 3 and Figure 4 shows the selected 20 samples of the originally scanned specimen images containing the mold defects.

Defects extraction from the grayscale images was performed by transforming them into their corresponding binary images. Binarization is carried out by determining the threshold value, *t*, which is a gray level that divides the images into two sets: C_0 (foreground) and C_1(background). The set C_0 consists of points with gray levels of $[1,2,\ldots,t]$ while C_1(have gray levels of $[t+1, t+2, \ldots, L]$.

As the mold defects are postulated to have gray levels values between 1 to L, determining the correct threshold value is essential for mold defect detection. In this paper, three established image thresholding methods were selected to determine the threshold gray value, t, and for binarization of the images. The image thresholding methods selected are Otsu's Method for Thresholding [11], Minimum Error Thresholding (MET) Method and Contrast Adjusted Otsu Thresholding. The three methods will be compared to the proposed Derivative Level Thresholding method.

2.3. *Derivative Level Thresholding*

A new binarization method which does not produce a threshold level is proposed in this study. Instead, the method derives the final binary image from a combination of binary images that was binarized at

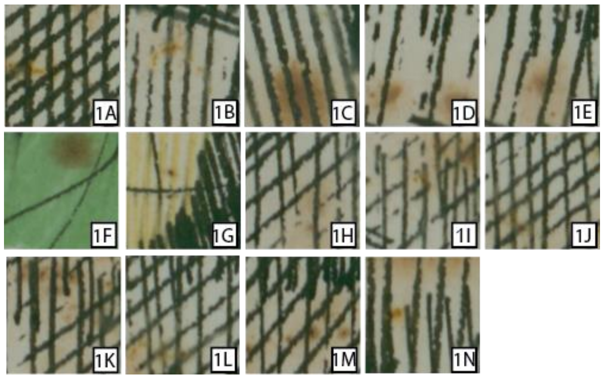


Fig. 3 Samples Selected from Painting 1

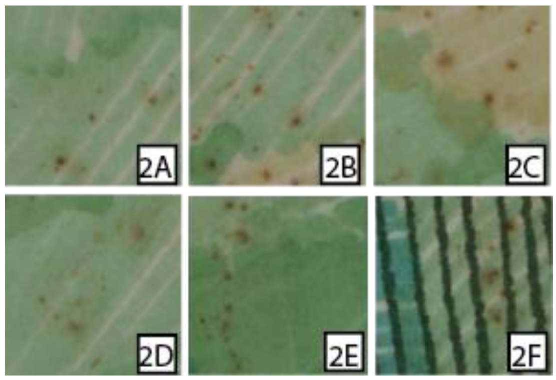


Fig. 4 Samples Selected from Painting 2

different threshold levels. For simplicity, the maximum number of threshold images to be considered is set at 20, which would produce a set of 20 binary images I_j. The binary images are collected by setting the threshold value t_j = *j*/20, where *j* = 1, 2, …, 20.

The resultant image consists of the background and foreground, and the mold defects would be identified in one or more images as black pixels against a white background. The percentage values of black pixels for each 20 images are calculated, and then plotted against the value of *j* = 1, 2, …, 20. The gradient of the curve, *y*, is calculated by taking its first order derivative, *dy/dx*, yielding a bimodal histogram. Both the curves *y* and its derivative, *dy/dx*, is shown in Figure 5 and Figure 6 respectively.

From Figure 5, we can see that the black pixel percentage values increase with the increase in the value of *j* increases. In the grayscale samples, the molds have grayscale values that are lighter than the artwork strokes but are darker than the background values. This results in the bimodal curve having two peaks that contains the artwork strokes and the background grayscale values, as indicated in Figure 6. The mold defects can be extracted by processing the images in between the two peaks.

Subsequently, the locations of the two peaks, $j_{peak\ 1}$ and $j_{peak\ 2}$, where $j_{peak\ 2} > j_{peak\ 1}$ are then located, from which the number of points, r, is determined. The number of points, r, is calculated by:

$$r = j_{peak\ 2} - j_{peak\ 1} \quad \text{for } j_{peak\ 2} - j_{peak\ 1} > 4 \qquad (5)$$

$$r = (j_{peak\ 2} - j_{peak\ 1}) + 4 \quad \text{for } j_{peak\ 2} - j_{peak\ 1} \leq 4 \qquad (6)$$

For the second case in (6), the value of $j_{peak\ 1}$ and $j_{peak\ 2}$ is modified to value of points $j_{peak\ 1}*$ and $j_{peak\ 2}*$:

$$j_{peak\ 1}* = j_{peak\ 1} - 2 \qquad (7)$$

$$j_{peak\ 2}* = j_{peak\ 2} + 2 \qquad (8)$$

where $j_{peak\ 2}* > j_{peak\ 1}*$ to supply two additional points for the case so that $r > 4$, while for the case in (5), the value $j_{peak\ 1}* = j_{peak\ 1}$ and $j_{peak\ 2}* = j_{peak\ 2}$.

The next step is to produce subtracted binary image, I_s, from every combination point pairs possible, from $j_{peak\ 1}*$ to $j_{peak\ 2}*$. The total number of combinations, R,can be calculated by:

$$R = \frac{r!}{2!(r-2)!} \qquad (9)$$

The process can be explained in a pseudocode form as follows:

<u>ALGORITHM 1: PSEUDOCODE</u>

BinaryImage $\boldsymbol{I_n}$;
SubtractedBinaryImage $(\boldsymbol{m})$;
NumberOfPoints **R**; //From Equation (5) and (6)
ModifiedFirstPoint $j_{peak\ 1}*$;
ModifiedSecondPoint $j_{peak\ 2}*$;

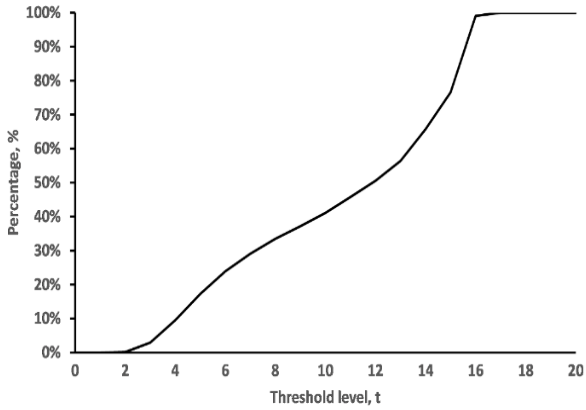


Fig. 5 Percentage of black pixel value, *y*

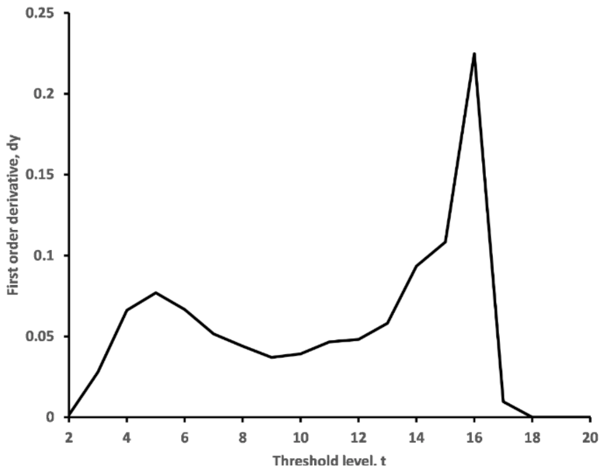


Fig. 6 The derivative of curve y, dy/dx

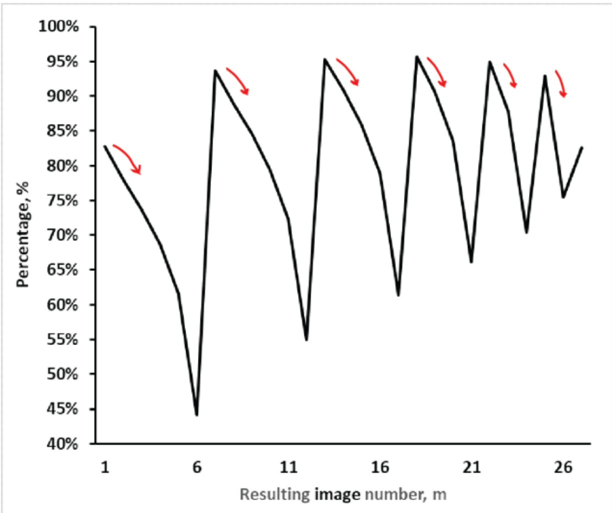


Fig. 7 Curve y_2 showing the peaks and valleys as a results of the image subtraction process.

for $m = j_{peak\ 1} *: j_{peak\ 2} *$
for $n = 1: \boldsymbol{R}$
$S(m) = I_{n+1} - I_n$
end
end
…

The percentage of black pixels of the resulting images $S(m)$, where $m = 1, 2, \ldots, R$, is then plotted and analyzed. The resulting curve, designated as y_2, has multiple peaks and valleys culminated from different sets of pairing between points. Curve y_2 is then plotted, as shown in Figure 7.

Curve y_2 is a 6 results series (indicated with arrows) can be identified where each series begins with a maximum value of black pixel to its minimum, and is separated with a straight line connecting the current series to the next. Ideally, some of the resulting images, $S(m)$, are adequate to be used in identifying the molds. The image that fit this criterion is called the optimized subtracted image, I_s.
To obtain the optimized subtracted image, I_s, with visible mold defects image, the corresponding m values are determined. The m values for obtaining the optimized subtracted image, S, is designated as m_s. The average values for all the peaks, q_{ave}, and valleys, v_{ave}, in the y_2 curve are calculated as follow:

$$q_{ave} = \frac{q_1+q_2+\cdots+q_N}{N_{peaks}} \qquad (10)$$

$$v_{ave} = \frac{v_1+v_2+\cdots+v_N}{N_{valleys}} \qquad (11)$$

The two values derived from both equations cannot be used to determine the value of m directly, as the average value may not be positioned on the y_2 curve. The values of q_{ave} and v_{ave} would be used to calculate the closest value for m_s which is:

$$m_s = round\left(\frac{m_{UL}+m_{LL}}{2}\right) \qquad (12)$$

where the lower limit, m_{LL}, is determined by finding the lowest m value that corresponds to the valleys, v, such that $y_2(m_{LL}) > v_{ave}$ and the upper limit, m_{UL}, is determined by finding the highest m value that corresponds to the peaks, q, such that $y_2(m_{UL}) < q_{ave}$. The obtained value is then rounded to the nearest integer to get m_S.

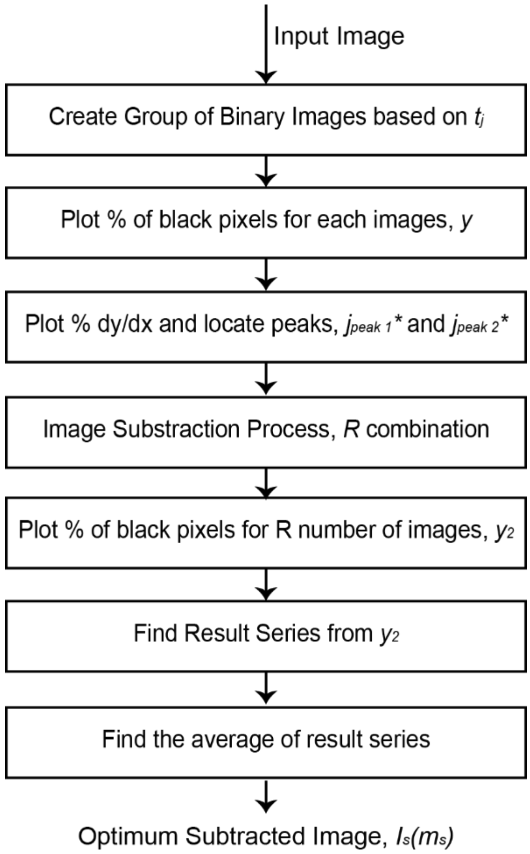


Fig. 8 Flow Diagram of the Derivative Level Method

Finally, the binary subtracted image used for defect detection can be obtained from the labeled image $I_S(m_S)$. The flow diagram shown in Figure 8 summarizes the overall process of the Derivative Level Thresholding.

2.4. *Defect Detection and Analysis*

Filtering and defect analysis can also be conducted on the binary images to locate and characterize defects. In a black and white image, the defects are represented by the white pixels.

2.4.1. Connected Component Filtering for Noise

Figure 9 shows the overall image processing for the defect detection which comprises of the process after the

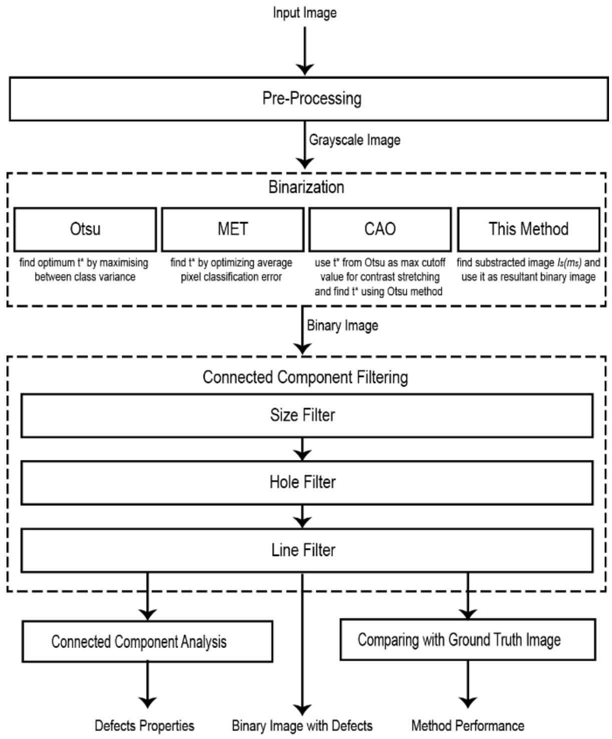


Fig. 9 Flow Diagram of the overall image processing for the defect detection.

Derivative Level Thresholding has been carried out, as well as after the other selected thresholding methods. Once the binarization is complete, the connected component filtering then be conducted. Prior to the filtering process, a connected component analysis using flood-fill algorithm is performed to locate groups of pixels. In this process, an unlabeled pixel is first located and a flood-fill algorithm is used to label adjacent pixels to be in the same group. In this study, 4-connected neighborhood component is used to determine pixel grouping. Each pixel group will be represented with a size, A_k , measured in unit pixel. The connected components are first filtered according to size, and the first filter will only store pixel groups that are larger than a_{min} and smaller than a_{max}:

$$a_{min} < A_k < a_{max} \quad (13)$$

Single pixels will not be stored, as well as group of pixels that are connected in diagonals. The value of a_{min} is set to 20 while a_{max} is set to 5000 pixels. Next, the selected group of pixels will be refined using a hole filter.

The hole filter works by filtering group of pixels which have a difference of filled image area to image area, A_k, that is larger than a preset scale threshold value. A filled image area is the total number of pixels in the same group of pixels with holes filled. By using this hole filter, group of pixels with holes will be filtered out. In this study, the hole filter threshold value is set at 0.1.

The filter is calculated by:

$$\frac{(A_{filled} - A_k)}{A_k} \quad (14)$$

Subsequently, a line filter is also applied on the images. In this process, the pixel group is accepted when it has an area to perimeter ratio of more than 0.6. This filter will eliminate lines - which will usually have a low area to

Table 1 Defect Detection Performance for Sample 5

No.	Binarization Method	Number of Defects Detected	Number of Defects Correctly Detected	Total Defect Surface Area (pixels)	Percentage of Defects (%)
1	Otsu's Method	9	1	13662	33.82
2	Minimum Error Thresholding (MET) Method	7	3	17099	42.32
3	Contrast Adjusted Thresholding Method (CA)	12	1	9614	23.80
4	Derivative Level Thresholding (DLT)	10	3	16614	41.12
5	Ground truth*	3	3	891	2.21

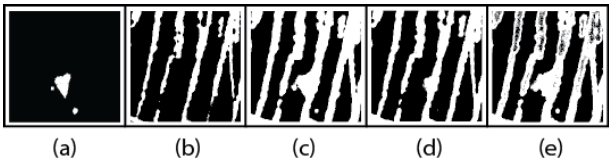


Fig. 10 Binary Images of Sample 1E from left: (a) Ground Truth Binary, and results of (b) Contrast Adjusted Thresholding, (c) MET, (d) Otsu and the (e) proposed Derivative Level Method.

perimeter ratio, ranging from 0.1 to 0.5, especially for straight lines.

2.4.2. Connect Component Analysis for Mold Characterization

The selected group of pixels are then remapped into the final binary image and are subjected to another connected component analysis using a similar 4-connected neighborhood connected component analysis approach. Data on defect properties such as the area size (in pixels), centroid (location) and morphology are collected for comparison. The defect detection is carried out for the newly proposed Derivative Level Thresholding method and the three existing thresholding methods. The results obtained are then compared to determine their relative performance.

3. Results and Discussions

In this study, the proposed Derivative Level Thresholding is compared with three existing binarization by thresholding methods. These methods are implemented in MATLAB R2016a and computed on an Intel(R) Core™ i7-4500U 1.80GHz processor with 8GB RAM on a Windows 10 Pro platform. The images captured using mesoscopic technique are processed using the selected binarization methods in accordance to the flow diagram in Fig. 8.

The comparative performance of all four methods were evaluated experimentally for detecting mold defects on the selected artworks. 20 samples were selected from different locations on the scanned image of the artwork with known mold locations from both the original painting image and the restored painting image. The respective ground truth images are also obtained to compare the performance of each thresholding and detection method. The results will be presented in two sections – the first section will focus on Sample 1E which originated from Painting 1. The second section is an overview of the results obtained from all 20 samples that consists of samples from Painting 1 and Painting 2.

3.1. *Sample 1E Results*

As the images are captured at a resolution of 650 dpi, the sizes of each pixel are approximately 1.6nm^2. The variation in the size of defects of on Sample 1E ranges from 4.8 nm^2 to 705.6 nm^2. Figure 10 shows the comparison of the binarization results from the four methods. The results images visibly suggested that the results for the proposed method has included the defects in the resulting binary image, along with false positives.

The results are compared to a ground truth image indicated in Figure 10(a) to confirm the accuracy and sensitivity of the defect detection performance of the various binarization methods. The ground truth is represented as an image that is manually produced from the feedback of expertise on the actual amount of mold defects. This is used as the benchmark to compare the performance of each method using their resultant defect images. The sensitivity and accuracy of the binarization methods can be calculated by comparing the detection results with the data from the ground truth image. The sensitivity or true positive rate is determined as:

$$Sensitivity = \frac{TP}{TP+FN} \qquad (15)$$

and the accuracy is determined as:

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \qquad (16)$$

where *TP* is True Positive, *FP* is False Positive, *FN* is False Negative and *TN* is True Negative.

As the binary images were processed with the Connected Component Filtering it can be said that while the Derivative Level Thresholding is robust with higher sensitivity (96.75%) and accuracy (60.94%) to MET method in the case of defect detection in Sample 1E.

Table 1 summarizes the defect detection results from Sample 1E with Figure 11 presenting the comparison in graph.

Table 1 shows the detect detection performance of the newly proposed method with comparisons to Otsu's, MET and Contrast Adjusted Thresholding methods for the analysis of Sample 1E. The ground truth values, showing the actual size and number of defects, are also included for comparison. The results show that the Derivative Level Thresholding has managed to correctly identify all 3 defects with lesser false positive as compared to the MET method. This can be seen in Figure 10(e) when compared with the ground truth image in Figure 10(a).

3.2. *Thresholding Method Overall Performance*

To measure the overall performance of the methods, 20 samples were selected from Painting 1 and Painting 2. 14 samples from Painting 1, and 6 from Painting 2 were selected in this overall evaluation. As shown in Figure 3 and Figure 4, the samples selected depict different types of artwork details comprising of different stroke thicknesses and colors. This is purposely chosen as a fair indicator on the general performance of all four methods. The robustness of each method used in mold defect detection can be determined from the calculated values of the accuracy and the sensitivity.

The results showed that the newly proposed Derivative Level Thresholding method has performed better than the three existing binarization methods. The DLT method has the highest average sensitivity at 84.73% as compared to the other methods having average sensitivity values ranging from 17 – 61%. DLT also has a high average accuracy 68.40%, second only to the Contrast Adjusted Thresholding Method with the latter having the lowest average sensitivity value. The average accuracy for DLT is also higher than the accuracy value discussed for Sample 1E. The study has also found that the value of accuracy for the DLT method is generally higher in samples that has less thick strokes.

From the results shown in Figure 12, it can be deduced that the Derivative Level Thresholding Method is efficient in detecting mold defects. The high sensitivity value means that the binary image produced by the DLT

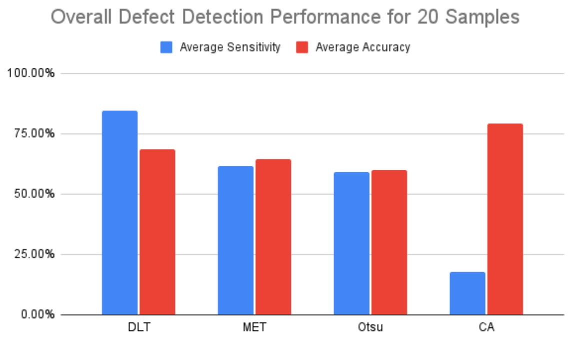


Fig. 12 Overall Defect Detection Performance for 20 Samples

method was able to correctly detect the defects, while the lower accuracy value indicate it has successfully done so at a cost. The method has considered more pixels in the samples as defects, resulting in higher false positive values, thus lowering the accuracy.

This problem can be addressed by having a better filtering after the binary image has been produced, sufficient to reduce false positives while at the same time ensuring that actual defects do not get filtered out. In the future, a better way of classifying the binary images resulting from the methods is required, as it can be concluded that the rudimentary filtering suggested in this study is not fully capable of avoiding false positives in the mold defect detections. The accuracy can be increased by considering all the defects detected separately, rather than using a blanket filtering approach.

In addition, instead of treating the results from the methods suggested in this study as the final mold defects, they can be treated as features extracted that can be used in machine learning methods to correctly classify the mold defects and non-molds. This will improve the final mold detection results and reduce the error generated from the restored image samples.

4. Conclusions

An automatic mold defects detection method, the Derivative Level Thresholding (DLT) Method has been developed to locate mold-type defects on high resolution scanned images of artwork paintings. The performance of this newly proposed method was compared to three existing common binarization methods for the evaluation of the 20 samples. It was found that the DLT method is better in terms of average accuracy (up to 68%) and has the highest average sensitivity of 84.73%. In general, the DLT method is shown to be robust and effective in distinguish molds from various types of painting samples. The methods suggested in the study can be further developed with machine learning methods to optimize their performance.

Acknowledgement

This study is partially sponsored by Impact Oriented Interdisciplinary Research Grant, Universiti Malaya with Project No.: IIRG034B-2019.

References

1. Kokkinaki, O., et al., Comparative study of laser induced breakdown spectroscopy and mass spectrometry for the analysis of cultural heritage materials. Journal of Molecular Structure, 2013. **1044**: p. 160-166.
2. Feng, Y.Q., N.N. Wang, and H.X. Ju, Highly Efficient Electrochemiluminescence of Cyanovinylene-Contained Polymer Dots in Aqueous Medium and Its Application in Imaging Analysis. Analytical Chemistry, 2018. **90**(2): p. 1202-1208.
3. Galli, A. and L. Bonizzoni, True versus forged in the cultural heritage materials: the role of PXRF analysis. X-Ray Spectrometry, 2014. **43**(1): p. 22-28.
4. Buchta, D., et al., Artwork Inspection by Shearography with Adapted Loading. Experimental Mechanics, 2015. **55**(9): p. 1691-1704.
5. Luo, Z.H., et al., Extraction of microcracks in rock images based on heuristic graph searching and application. Computers & Geosciences, 2015. **85**: p. 22-35.
6. Deborah, H., N. Richard, and J.Y. Hardeberg, Hyperspectral Crack Detection in Paintings. 2015 Colour and Visual Computing Symposium (Cvcs), 2015.
7. Guarneri, M., et al., 3D remote colorimetry and watershed segmentation techniques for fresco and artwork decay monitoring and preservation. Journal of Archaeological Science, 2014. **46**: p. 182-190.
8. Desai, S.D., et al., User Intervention Based Detection & Removal of Cracks from Digitized Paintings. 2014 Fifth International Conference on Signal and Image Processing (Icsip 2014), 2014: p. 15-19.
9. Desai, S.D., et al., Detection & Removal of Cracks from Digitized Paintings and Images by User Intervention. 2013 Second International Conference on Advanced Computing, Networking and Security (Adcons 2013), 2013: p. 51-55.
10. Win, M., et al., A Contrast Adjustment Thresholding Method for Surface Defect Detection Based on Mesoscopy. IEEE Transactions on Industrial Informatics, 2015. **11**(3): p. 642-649.
11. Otsu, N., A threshold selection method from gray-level histograms. Automatica, 1975. **11**(285-296): p. 23-27.
12. Kittler, J. and J. Illingworth, Minimum Error Thresholding. Pattern Recognition, 1986. **19**(1): p. 41-47.

Authors Introduction

Mr. Hilman Nordin

Hilman Nordin received the B.Eng degree in computer-aided design and manufacturing engineering and the M.Eng.Sc. degree in mechanical engineering (mechanical systems) from Universiti Malaya, Kuala Lumpur, Malaysia in 2009 and 2013, respectively. He is currently a Ph.D. candidate at Universiti Malaya. His research interest include image processing and analysis, high-resolution scanning and its industrial applications and mechanical systems.

Dr. Bushroa Abd Razak

Assoc. Prof. Datin Ir. Dr. Bushroa Abd Razak received her B.Eng. in Electrical and Electronic Engineering from University of Manchester Institute Science and Technology, UK in 1997, which was merged to Manchester University. Consequently, she received her Master in Engineering Science and PhD from University of Malaya, Malaysia in 2004 and 2012, respectively. She started her career in 2004 at the Department of Engineering Design and Manufacture, UM, and later in 2013 at the Department of Mechanical Engineering in the same university when both departments were merged. Apart from teaching, she is actively involved in research, management and consultation works. In the late of 2020, she was appointed as a Head of Department of Mechanical Engineering, University Malaya. During her tenure, she has contributed impactful in securing the Engineering Accreditation Council (EAC) as well as revising new curriculum review for courses in Mechanical Engineering. She holds a few memberships including a board member of Advanced Manufacturing and Material Processing Centre, UM. She is awarded a Chartered Engineer of IET, UK. In July 2016, she was promoted to Associate Professor. In 2015-2016, she went to Univ of California, LA, USA for her post-doctoral research program sponsored by the Ministry of Higher Education, Malaysia. She was also awarded as a Professional Engineer by the Board of Engineers (BEM) in March 2017. Till now, she is the author and co-author of more than 100 publications in international journals and proceedings with H-index reaches above 25. Her research interest includes fundamental, application, technology and commercialization in Surface Engineering and its related field of study. She is passionate in research that contributes to society.

Dr. Norrima Mokhtar

Norrima Mokhtar received the Bachelor of Engineering (B. Eng) degree in Electrical Engineering from University of Malaya in 2000. After working two years with International Telecommunication Industry with attachment at Echo Broadband GmbH, she managed to secure Panasonic Scholarship which required intensive screening at the national level in 2002. She finished her Master of Engineering (Oita, Japan) under financial support received from Panasonic Scholarship in 2006. Consequently, 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 held several important positions such as Coordinator for Industrial Training, Coordinator for Research Laboratory, Coordinator for EBA-Consortium(Keio University), Coordinator for Nagaoka University of Technology and Coordinator for Manchester Metropolitan University(UK). Dr. Norrima was also the principal investigator for High Impact Research Grant (MOHE) worth of RM 1 million along with Science Fund and CREST grant. She is the author and co-author of more than 50 publications in international journals and proceedings in Sensors, Automation, Image Processing, Human-Computer Interface, Brain-Computer Interface, UAV and Robotics. To date, she had successfully supervised 3 PhD and 3 Master Engineering Science students (by research). She is also active as reviewer for many reputable journals and several international conferences.

Mr. Fadzil Jamaludin

Mohd Fadzil Jamaludin is a Senior Research Officer at the Centre of Advanced Manufacturing and Material Processing (AMMP Centre), Universiti Malaya. He has received the Bachelor of Engineering (B. Eng) degree in Computer Aided Design and Manufacturing (CAD/CAM) and Masters of Engineering Science (M.Eng.Sc) in 2003 and 2010, respectively, from Universiti Malaya, Kuala Lumpur, Malaysia. His research interest includes precision joining technologies, additive manufacturing and mixed reality development for research and education.

Imaginary Finger Control Detection Algorithm Using Deep Learning with Brain Computer Interface (BCI)

Suresh Gobee, Norrima Mokhtar, Hamzah Arof, Noraisyah Md Shah
Department of Electrical Engineering, Faculty of Engineering, University of Malaya, Malaysia
E-mail: sureshgobee@gmail.com, norrimamokhtar@um.edu.my

Wan Khairunizam
Faculty of Electrical Engineering Technology,University Malaysia Perlis, 02600 Perlis, Malaysia
E-mail: khairunizam@unimap.edu.my

Abstract

Before the advancement of deep learning technology, the brain signals are to be analysed manually by the neuroscientists on how the brain signals reacts in proportion with the human body. This process is very time consuming and unreliable. Therefore, this project aims to develop a brain signal detection system based on deep learning algorithm in response to the output of EEG device on the imagery finger movements. These fingers include thumb, index, middle, ring and little of right hand. There are 4 CNN classification models being developed in this project. They differ with each other in terms of the pre-processing requirements and the neural network architecture. The best results for offline classification obtained in this project are 69.07% and 82.83% respectively in terms of average accuracy from 6-class and 2-class tests. Moreover, this project has also developed a proof of concept for applying the trained models in online or real-time classification.

Keywords: BCI, Imaginary finger movement, CNN, EEG.

.

1 INTRODUCTION

Brain-computer Interface (BCI) is a device used to detect brain signals[1]. BCI has wide range of applications, such as stroke rehabilitation, controlling smart devices, communication device for patients without motor abilities, mental disease diagnosis and many more. There are three types of BCI devices: invasive, partially invasive, and non-invasive.

Due to the clinical risks and impracticality of invasive and partially invasive BCI, this research will only focus on electroencephalogram (EEG) based non-invasive BCI device. According to[2], EEG are mostly used in other BCI research because of its reliability, ease of use and non-invasiveness. EEG functions by detecting the brain signals in terms of voltage fluctuations from the human scalp through the electrodes. Besides, EEG is high in temporal resolution and relatively low in cost.

According to[3], there are six classes of brain signals: motor imagery (MI), sleep stage scoring, emotion recognition, mental workload, seizure detection and event-related potential detection. Since that this research is about imagery finger detection, therefore, it falls under the category of MI. MI refers to the mental's thought on motor movements [4]. This phenomenon can be detected by the EEG devices as the brain signals are induced in motor cortex during imagination of motor movements.

To classify these signals, deep learning approach are often being used. Deep learning is a subset of machine learning in which a computer model learns to perform certain tasks based on a set of data [5]. The deep learning architecture is inspired by the structure of human brain; hence, they are often called as artificial neural networks. When properly trained, deep learning models can achieve high accuracy for tasks like classification, regression, and text generation [6].

The research done by [7-9] do not implement any feature extraction technique to pre-process the training data. The raw EEG signals were used directly as the training data for the deep learning algorithm. Due to the nature of Convolutional Neural Network (CNN) architecture, the learnt parameters in this algorithm form layers of filters capable of producing the feature maps from the raw EEG signals. Therefore, by stacking more convolutional layers, the deep learning algorithm can produce higher order of feature maps automatically, increasing the classification accuracy. However, the increase in number of learnable parameters causes more computational power required to execute the algorithm.

The Common Spatial Pattern (CSP) and Fast Fourier Transform Energy Map (FFTEM) algorithms was used by

[10]for feature computation and feature selection, respectively. The CSP algorithm works by applying spatial filters to obtain the distinctive features between two classes. The FFTEM algorithm then used these features to compute the energy maps. These energy maps are used as the training data. This method can reduce the computational power needed for running the deep learning algorithm as lesser neural network layers are required. However, some of the features from the signals might be discarded when applying the CSP algorithm which leads to lower classification accuracy.

In the papers written by [11-12], Short Time Fourier Transform (STFT) was applied to convert the EEG signals into 2D images. The resulting images contain 3 information: time, frequency, and intensity. This technique can greatly increase the classification accuracy of the algorithm since CNN architecture is very effective on image recognition. However, performing STFT for feature extraction increases the complexity of the classification algorithm because higher dimensions of data is being used.

The combinations of artefact removal with bandpass filters, CSP feature extraction and random forest (RF) classifier to classify the imagery finger movements was done by [13]. This research has successfully obtained the best accuracy of 54% with 5 classes (thumb, index, middle, ring and little) of imagery finger movements.

The research done by [14] uses the combinations of artefact removal with bandpass filters, ERD/ERS feature extraction and Support Vector Machine (SVM) classifier. This research obtains average accuracy of 62.5% in classifying the 2 classes of imagery movements for left and right index fingers.

2 METHODS

2.1. Overall system

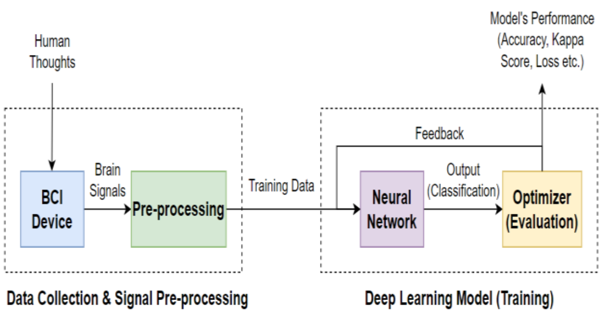


Fig 1: Block diagram for the system in training the deep learning based classification algorithm.

The system shown in Fig 1 can be separated into 2 sections: data collection and signal pre-processing as well as training and evaluation of the classification algorithm. The former prepares the MI dataset recorded from the EEG device. In this stage, artefacts removal, feature extraction and data conversion are performed at the raw EEG signals. These pre-processed signals are being used as the training data for the deep learning model. The deep learning model in turn trains the neural network to classify the MI tasks based on the pre-processed training data. The deep learning model's performance will be evaluated in terms of accuracy and loss.

2.1.1 Data Pre-processing Technique

Summarizing the outcomes obtained from literature review, there are 5 feature extraction techniques that can be applied in this research as shown in Table .

Table 1: Description of pre-processing techniques that can be applied in this research.

Pre-processing Technique	Description
No feature extraction	Raw EEG signals are used directly as the input to the deep learning algorithm.
EEG channels selection	Only certain EEG channels that contains the most MI related information are selected as the input to the deep learning algorithm.
Short-time Fourier transform (STFT)	This technique converts the time-based EEG signals into 2D images as the input to the deep learning algorithm.
Common spatial pattern (CSP)	CSP algorithm is applied on the EEG signals to obtain the distinctive features between each class. The extracted features are used as the input to the deep learning algorithm.
Cropped training	Each MI trials are cropped with a sliding window of smaller period. This technique increases the number of training data and improves the performance during online classification.

In this research, combinations of each of these data pre-processing techniques will be applied. The results are compared to obtain the classification algorithm with the highest performance.

2.1.2 Deep Learning Architecture

Summarizing the outcomes obtained from the literature review, there are 3 convolutional layer architectures that can be applied in this research as shown in Table .

Table 2: Description of different convolutional layer architectures.

No.	Deep Learning Architecture
1	2 convolutional layers with one-dimensional filters. These filters are applied across the time and channel axis of EEG data, respectively. This architecture allows the filter parameters to be learnt separately and produce feature maps on the temporal and spatial information from the input signals.
2	2 convolutional layers with two-dimensional filters. This architecture is applicable only when the EEG signals are converted into 2D images via Short Time Fourier Transform (STFT). This technique can greatly reduce the number of learnable parameters in the deep learning algorithm, hence, reducing the computational power needed to run the algorithm.
3	The Channel-wise CNN architecture. This architecture consists of 2 convolution layers both with 1D filters applied across the time axis of the EEG signals. Higher order of feature maps can be obtained by applying more than one filters across the temporal axis of the signals. Therefore, this can improve the classification accuracy of the trained model.

In this research, each of CNN architectures will be applied and compared to obtain the technique that can yield to the highest performance.

2.1.3 Preliminary Design for Data Acquisition

The first step for acquiring the EEG dataset is to determine the timings where the subjects perform the imagery finger movement tasks. To ensure that the timings used are consistent across each subject, videos will be used as the guidance for collecting the dataset. To achieve this, a timing diagram is developed as the videos framework as shown in Fig .

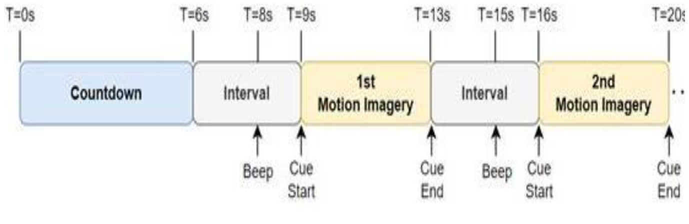


Fig 2: Timing diagram for data acquisition.

10 samples per class per session will be used in this research. The cue time or the period of MI task is 4 seconds. The time interval between each MI is 3 seconds to allow space for the subject in preparing the next MI task. A beeping sound will be played 1 second before the cue start to notify the subject for the next MI task. Therefore, combining the time required for initial countdown, 76 seconds are needed to record each MI class per subject.

2.1.4 Preliminary Design for Artefact Removal

Since that battery-based EEG device will be used in this research, the use of notch filter (bandstop filter) to remove the line noise is not required. This is because the battery power supply can provide perfect direct current (DC) to the device, hence, the line noise is not present in the EEG signals. Besides, 8-30Hz bandpass filter will be applied to the EEG signals to remove the noise in the signals due to muscle movements. This frequency range is chosen as it covers the mu (8-13Hz) and beta (13-30Hz) frequency bands of brain signals, consisting of the most optimal signal-to-noise ratio for MI related classification.

2.1.5 EEG Dataset Collection

In stage 1, the EEG headset is used to record the brain signals of the subjects. Two recording sessions were conducted for each subject, with 10 MI trials per class per session. Therefore, there will be a total of 20 MI trials for each class per subject. The EEG signals are sampled at the rate of 250Hz and 16 channels. The electrode locations (based on 10-20 EEG electrode placement standard) and their respective channels were illustrated in Fig **1** 3 and summarized in Table 1.

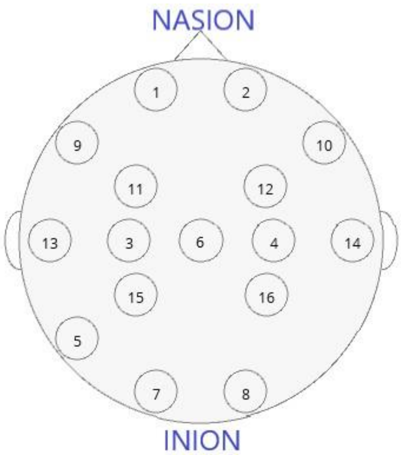


Fig 1: Illustration of electrode locations and their respective channel.

Table 1: The list of electrode locations and their respective channel.

Channel	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
Location	Fp1	Fp2	C3	C4	T5	Cz	O1	O2	F7	F8	F3	F4	T3	T4	P3	P4

Before the recording session, the subjects were asked to ensure the cleanliness of their hair and scalp. This is to prevent that the EEG signals obtained contains noises due to the poor conductivity of the electrodes caused by blockage of grease or dirt on the scalp. The subjects were briefed about the details and procedures of the recording session, while any enquiry from the subjects were answered. After that, they were asked to fill in a consent form as an agreement to participate in this research.

During the recording session, the subjects were asked to sit on a comfortable chair in a quiet room, with their right arm resting on a table and palm facing upwards. This is to help the subjects on focusing during the recording session. Then, the subjects were asked to wear the EEG headset. The electrodes were tuned such that they were touching the subjects' scalp. A video was played for each class and the subjects were asked to perform MI trial according to the instructions given in the video as shown in Fig 4. The subjects were requested not to do any muscle movements such as blinking when performing the MI trials. This is because the muscle movements will decrease the quality of the EEG dataset, as it confuses the deep learning algorithm during the training session.

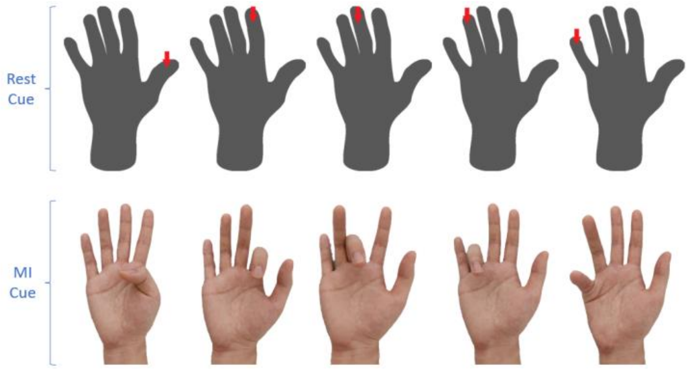


Fig 2: Image montage of the cues used in the video for recording EEG dataset.

After the first recording session was completed, the subjects were given 5 minutes of break before continuing with the second session. The total duration of recording sessions is estimated to be around 15 to 20 minutes per subject excluding the time taken for briefing.

2.2 Classification Model Development

In this stage, the EEG signals data was pre-processed before being used as the training and testing data for the CNN classification models. The EEG signals were passed through an 8-30Hz bandpass filter to remove the artefacts.

After artefact removal, the EEG signals are trimmed according to the MI task for each class and the rest interval. The trimmed dataset is then separated into testing and training data with the ratio of 9:1. After that, cropped training algorithm is applied on the dataset to further increase the number of testing and training data.

After trimming, the EEG dataset are down sampled from 250Hz to 125Hz. During dataset collection, the EEG signal values were stored directly into the SD card from the EEG headset. Using SD card as storage allows more consistent and higher sampling frequency (250Hz) as compared with streaming via Bluetooth connection and storing the data in computer. However, during online classification, the EEG signals are streamed directly into the computer via Bluetooth connection at the sampling frequency of 125Hz. Therefore, down sampling on the EEG dataset is required so that the trained CNN classification models can be used in online classification. The flow is shown in Fig 5.

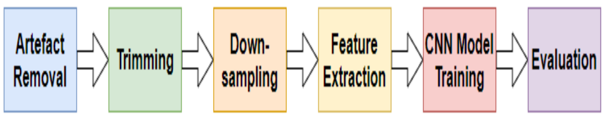


Fig 5: Block diagram of classification model development.

Feature extraction is applied to the EEG dataset before training and testing the CNN classification model. 4 types of CNN classification model are developed in this project to classify the EEG signals. These models differ with each other in terms of feature extraction requirement and the neural network architecture as summarized in Table 2.

Table 2: Summary of 4 types of CNN classification model. The abbreviations used are power spectral density (PSD), short-time Fourier transform (STFT), frequency band splitting (FBS), standard deviation (STD), convolutional (Conv) and fully connected (Fc).

Model	**Feature Extraction**					**Number of CNN Layers**	
	PSD	**STFT**	**FBS**	**Mean**	**STD**	**Conv**	**Fc**
A						2	2
B	x					2	1
C		x				2	3
D			x	x	x	2	2

Each of the CNN classification model was trained with 6 different sets of classes. The first set contains 6 classes including thumb, index, middle, ring, little and rest MI. The other 5 sets (2 classes) contain the combination of thumb, index, middle, ring and little MI each with respect to the rest MI. Details about the CNN classification models and the test results will be discussed in sections 错误!未找到引用源。 and 0 respectively.

2.2.1 Models

Model A

Model A uses the raw EEG signals, which contain the sampled voltage values per unit time for each channel, as the input to the CNN architecture. Therefore, feature extraction is not required in this model. Instead, these features are produced by the filters in the convolutional layers during the training process as shown in Fig 6.

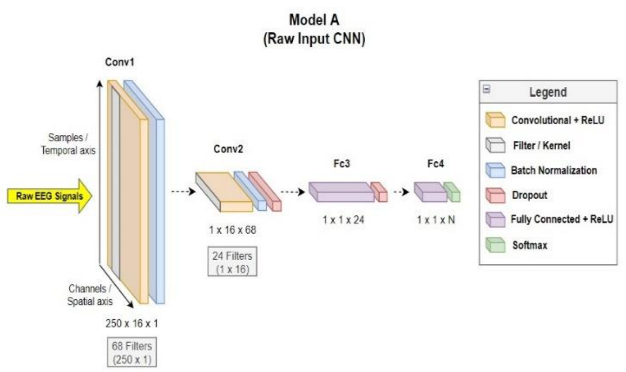


Fig 6: Block diagram of Model A.

The CNN architecture of Model A consists of 9 layers of neural network components as illustrated in 错误!未找到引用源。. The first convolutional layer Conv1 consists of 68 one-dimensional filters applied across the temporal axis of the EEG signals. This allows the neural network to learn and extract the features on the temporal axis for each channel on the EEG signals. The feature maps produced by Conv1 layer are fed into the second convolutional layer Conv2. As shown in Fig. 7, this layer consists of 24 one-dimensional filters applied across the spatial axis on the EEG signals. This allows the neural network to learn and extract the features related to the spatial information from the EEG signals. The fully connected layer Fc3 is added to allow the neural network in recognizing more complex features. The layers labelled as Fc4 act as the classification layer of Model A, with the number of output classes as the number of nodes (N).

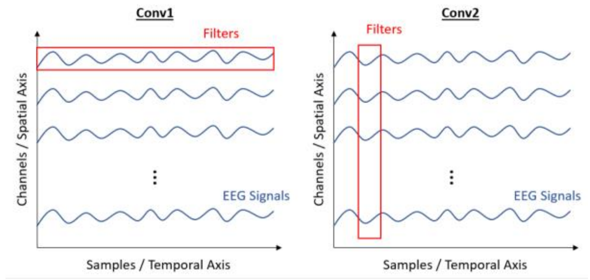


Fig 7: Illustration on the filter size with respect to EEG signals for each convolutional layer in Model A.

Similar method was applied for the other 3 models. Model B, Model C and Model D as describe in Table 4.

2.2.2 Testing of the Proposed Design

There are 6 subjects participated in the EEG data collection for this project. The test for online classification requires the participation of the same subject in separate session after CNN classification models were trained. Therefore, this test could only be conducted on one subject due to the limitation of time for the subject's participation.

3 RESULTS AND DISCUSSIONS

Table 3: Test results of offline classification for Model A.

Subject	**6-Class Model**	**2-Class Model**
	Accuracy (%)	**Accuracy (%)**
1	62.22	84.00
2	51.67	65.00

3	51.67	67.00
4	54.45	71.00
5	71.11	89.00
6	41.67	69.00
Avg	**55.46**	**74.17**

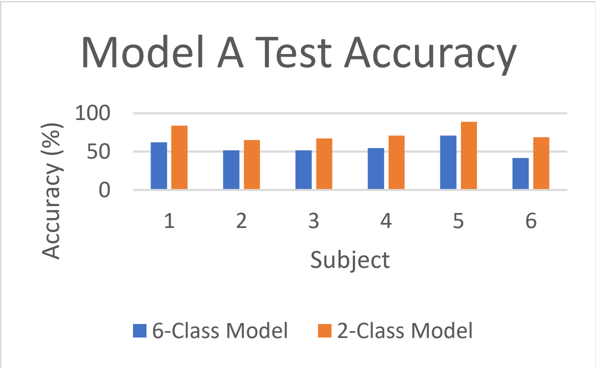


Fig 8: Bar chart showing test results of offline classification for Model A.

As shown in Table 5 and Fig 8. Model A has the average accuracies of 55.46% and 74.17% on the 6-class and 2-class tests respectively. Although the accuracy in 2-class test is 18.71% higher than the 6-class test.

Table 6: Test results of offline classification for Model B.

Subject	**6-Class Model**	**2-Class Model**
	Accuracy (%)	**Accuracy (%)**
1	73.33	98.00
2	52.22	77.00
3	53.33	83.00
4	54.45	79.00
5	53.89	83.00
6	46.60	71.00
Avg	**55.64**	**81.83**

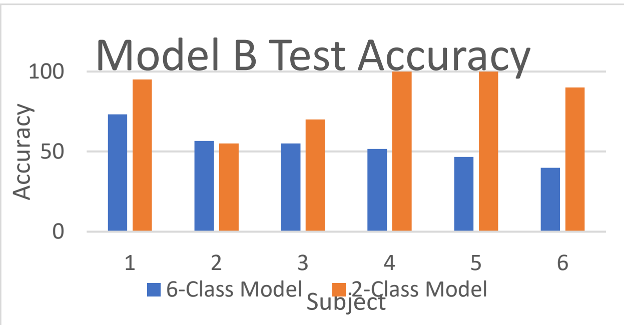


Figure 9: Bar chart showing test results of offline classification for Model B.

Model B has the average accuracies of 55.64% and 81.83% on the 6-class and 2-class tests respectively as shown in table 6 and Fig 9. This result indicates that Model B is better in performance when classifying lower number of classes.

Table 7: Test results of offline classification for Model C.

Subject	**6-Class Model**	**2-Class Model**
	Accuracy (%)	**Accuracy (%)**
1	61.11	86.00
2	30.00	61.00
3	34.45	63.00
4	50.56	70.00
5	66.67	87.00
6	36.11	53.00
Avg	**46.48**	**70.00**

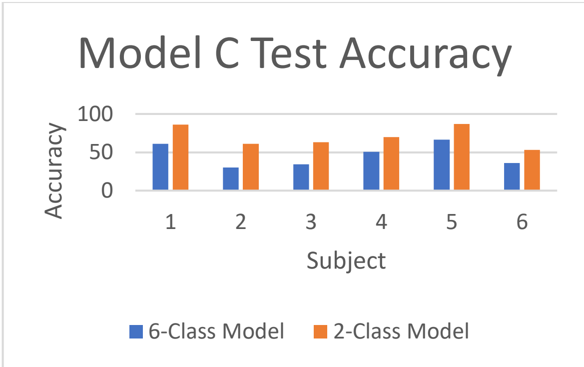


Fig 10: Bar chart showing test results of offline classification for Model C.

Model C has the average accuracies of 46.48% and 70% on the 6-class and 2-class tests respectively. Although the accuracy in 2-class test is 23.52% higher than the 6-class test as shown in Table 7 and Fig. 10.

Table 8: Test results of offline classification for Model D.

Subject	**6-Class Model**	**2-Class Model**
	Accuracy (%)	**Accuracy (%)**
1	78.33	97.00
2	56.11	75.00
3	69.45	79.00
4	71.67	84.00
5	73.89	88.00
6	65.00	74.00
Avg	**69.07**	**82.83**

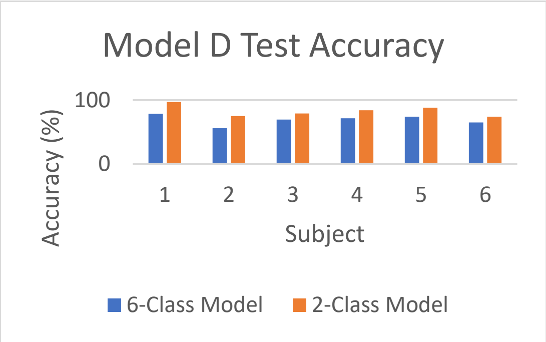


Fig 11: Bar chart showing test results of offline classification for Model D.

Model D has the average accuracies of 69.07% and 82.83% on the 6-class and 2-class tests respectively. Model B shows the best accuracy for 2 and 6 Classes. The overall result shows that Model D has the highest accuracy both 6-class and 2-class tests. This indicates that the combinations of feature extraction techniques with frequency band splitting, mean and standard deviation are most optimized for MI related EEG classification by using CNN architecture. As shown in Table 9.

Table 9: Summary of test results for all models in offline classification.

Model	6-Class Model	2-Class Model
	Accuracy (%)	Accuracy (%)
A	55.46	74.17
B	55.64	81.83
C	46.48	70.00
D	69.07	82.83

Table 10: Standard deviations of classification accuracy for each subject.

Model	Standard Deviation		
	6-Class Model	2-Class Model	Average
A	10.11	9.89	10.00
B	11.29	18.44	14.86
C	15.25	13.89	14.57
D	7.75	8.75	8.25

The consistency of the CNN classification model to be implemented across different individuals can be indicated by calculating the standard deviation on the classification accuracy across each subject as shown in
Table . The lower the value of standard deviation, the more consistent is the CNN model. This is because lower standard deviation indicates that the difference in accuracies across each subject is low. Therefore, Model D has the highest consistency among the other CNN classification models.

4 CONCLUSIONS

This project contributes to the creation of 4 techniques to classify the imagery finger movements on EEG signals by using the state-of-the-art deep learning technology. The best results achieved was 82.83% and 69.07% in terms of accuracy on classifying the MI of 2 classes and 6 classes respectively. In addition, this project also contributes to the development of EEG classification algorithm that is both lightweight and accurate. The training, testing and online classification could be run entirely on CPU without the need of discrete graphics. Hence, the low computational power requirement contributes to reducing the cost for BCI solution and make it affordable to more people. Moreover, the low computational power requirement also indicates that it is possible to integrate the BCI device and classification algorithm as a battery-based standalone system, which can be useful when being used as the controller for other devices. This project has successfully achieved all the research objectives. 4 types of CNN based deep learning algorithm were developed for classifying the imagery finger movement consisting of 6 classes and 2 classes. The best results obtained are 69.07% and 82.83% in model D, respectively in terms of average accuracy for 6 classes and 2 classes.

REFERENCES

1. Gu, X. et al., 2020. A Survey of Recent Studies on Signal Sensing Technologies and Computational Intelligence Approaches and Their Applications. EEG-based Brain-Computer Interfaces (BCIs).
2. Zhang, X., Yao, L., Wang, X., Monaghan, J., Mcalpine, D., & Zhang, Y. (2020). A survey on deep learning-based non-invasive brain signals: recent advances and new frontiers. Journal of neural engineering, 10.1088/1741-2552/abc902. Advance online publication.
3. Cho, H., Ahn, M., Moonyoung, K. & Sung, C. J., 2018. A Step-by-Step Tutorial for a Motor Imagery–Based BCI.
4. Brownlee, J., 2020. What is Deep Learning?
5. Mathworks, 2021. What Is Deep Learning? How it works, Techniques & Applications.
6. Tang, Z., Li, C. & Sun, S., 2016. Single-trial EEG classification of motor imagery using deep convolutional neural networks.
7. Schirrmeister, R. T. et al., 2017. Deep Learning With Convolutional Neural Networks for EEG Decoding and Visualization.
8. Dose, H., Møller, J. S., Iversen, H. K. & Puthusserypady, S., 2018. An End-to-end Deep Learning Approach to MI-EEG Signal Classification for BCIs. Expert Systems with Applications.
9. Abbas, W. & Khan, N. A., 2018. DeepMI: Deep Learning for Multiclass Motor Imagery Classification.] (Tabar & Halici, 2016)

10. Anam, K., Nuh, M. & Al-Jumaily, A., 2019. Comparison of EEG Pattern Recognition of Motor Imagery for Finger. Bandung, Indonesia, 2019 6th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)

Authors Introduction

Mr. Suresh Gobee

He received his Master's degree from the Department of Mechanical Engineering, University Science Malaysia, in 2002. He received the BTECH in 1998 form University Science Malaysia Currently, he is a lecturer in Asia Pacific University of Innovation and Technology (APU), KL, Malaysia. His research interest includes image processing, artificial intelligence and BCI.

Dr. Norrima Mokhtar

Norrima Mokhtar received the Bachelor of Engineering (B. Eng) degree in Electrical Engineering from University of Malaya in 2000. After working two years with International Telecommunication Industry with attachment at Echo Broadband GmbH, she managed to secure Panasonic Scholarship which required intensive screening at the national level in 2002. She finished her Master of Engineering (Oita, Japan) under financial support received from Panasonic Scholarship in 2006. Consequently, 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.

Professor Dr. Hamzah Arof

He received the B.Eng. degree from Michigan University 1992 and Ph.D. degree from University of Wales, Cardiff. He is currently serving as senior professor at the Department of Electrical Engineering, University of Malaya. His research interests are image processing, machine learning and photonics.

Dr. Noraisyah Binti Mohamed Shah

She received the B.Eng. degree from University of Malaya in 1999, Master Degree from Oita Universitym, Japan, and the Ph.D. degree from George Mason University. She is currently a senior lecturer at the Department of Electrical Engineering, University of Malaya. Her research interests are image processing, machine learning and satellite communication.

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.

Investigating the Effect of Individuality Factors in Measuring Aggression induced by Human Brain

Wan Khairunizam, Kai Xu tung, M. Lugieswaran, Wan Azani Mustafa, Hashimah Ali
Faculty of Electrical Engineering Technology, University Malaysia Perlis, 02600 Perlis, Malaysia

Zuradzman M. Razlan, Shahriman AB
Faculty of Mechanical Engineering Technology, University Malaysia Perlis,02600 Perlis, Malaysia

Norrima Mokhtar
Department of Electrical Engineering, Faculty of Engineering, University of Malaya, Malaysia

E-mail: khairunizam@unimap.edu.my

Abstract

Aggression is a behaviour of human that may cause physical or emotional harm to others. Several factors that cause aggressive behaviour such as physical health, mental health and socioeconomic. Many previous researchers reported that aggression could be measured through either questionnaire or the brain signals. This paper proposes the experimental studies to collect the brain signal of the human subject for investigating the effect of individuality in aggression. Ten subjects are selected to perform the aggression activities. The experimental protocol for inducing aggression is proposed. In general, there are four tasks which is collecting brain data in relaxing state before and after the experiments, and data collection while playing game in muted and maximum volume levels. In the experiments, the subject are required to play a popular non-violence smart phone game named "Subway Surfers" and at the same time the EEG signals are recorded from the subject's brain. In the signal pre-processing stage, a Butterworth filter is used to remove the noises contain in the signals. A windowing technique is employed for extracting significant features. A Pearson correlation technique is used to reduce and remain the less and most significant features. In the methodologies, the aggressiveness level *A*, is defined to investigate the effect of individuality in inducing the aggression signals. The proposed experimental protocol and signal processing techniques are seen able to generate level of aggression.

Keywords: Aggression, EEG, Aggression measure, Aggression level

.

1. INTRODUCTION

The word aggression is interpreted in the literature to reflect all the actions of aggressive goal hostilities. Instead, the term aggressiveness extracted from aggression is not limited to violence and hostility but has been used to quantify positive behavioral responses from the emotions of a subject [1,2,3,4]. To test the aggressiveness of the subject, traditional approaches use questionnaires method, which is Buss Perry Aggressive Questionnaire, (BPAQ) [1], however, recently the researchers start measuring the aggression by using EEG [5,6,7,8].

This paper proposes the experimental studies to measure aggression through brain signals. An experimental protocol is designed to induce aggression. The raw brain signals are processed by using signal processing approaches. The aggressiveness level is defined and is used to measure it in the aggression experiments.

The flow of this paper is followed. Introduction discusses the background including the related information to the research as published by the earlier researchers. Methodologies discuss the proposed experimental studies for inducing and obtaining aggression. Result discusses the finding from the investigation, and the paper is concluded in the conclusion.

2. METHODS

2.1. Traditional survey (BPQA) and brain signal experiments (EEG)

The experiment included in this research aimed at investigating and measuring aggressiveness through their effect on the human brain. The evaluation starts with the subjects to answer Buss Perry Aggressive Questionnaire (BPQA) questionnaire without the involvement of EEG data collection [5,7,9]. The evaluation was done before the EEG recording experiments to analyze their survey-based aggressive score.

2.2. EEG Recording Materials

The data collection for the analysis moment was taken by using the mindset 24 amplifier as presented in **Figure 1**, which able to cover most of the EEG points according to the "10-20 electrode placement system" as shown in **Figure 2**. The 19 channel electrodes (FP1, FP2, F7, F3, FZ, F4, F8, T3, T5, C3, CZ, C4, T4, T6, P3, PZ, P4, O1 and O2) were placed on the scalp using 10-20 electrode positioning system and the reference electrode were placed on the left and right mastoids. The sensors that were used in this experiment were the silver chloride ($AgCl_2$) disc electrodes. The Lycra Stretch Cap fix 19 electrodes closely to the subject's head. Electrodes were pre-positioned following the international 10/20 montage, helping novice EEG researchers to minimize electrode placement errors.

Fig. 1. Mindset 24 EEG amplifier

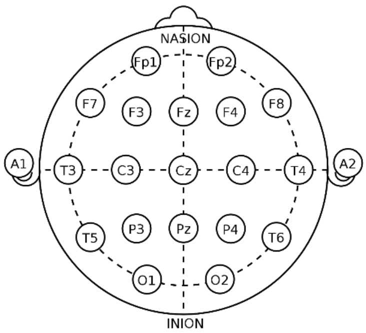


Fig. 2. EEG electrode's location according to 10-20 system

2.3. Subject selections

In the experiments, 10 subjects were chosen to perform the tasks. The subjects were asked to fill up the Buss-Perry Questionnaire (BPQA) and to involve in the Electroencephalography (EEG) experiments.

2.4. Aggression Induction Via Mobile Game

The subjects were taught to play the smartphone game "Subway Surfers" until he/she is familiar with the game for 5 min [9]. Firstly, the game was played in the silence of mute mode. The subjects were asked to achieve a certain score (1500m/500gold coin collected) for the game and survive for at least one minute. The gold coin collecting while surviving the game was meant for inducing aggressiveness onto the subject. The urge to play the game after several fail attempts to achieve the goal is a form of aggressiveness that intent to be monitored through this experiment. Provided the goal is yet to be achieved, the game has to be repeated until the goal was successfully met. For subjects that fail to achieve the minimal goal after multiple training session and they felt hard to cope with that and wish to opt-out, their recording was excluded from the research analysis as the training and trying out process might apply too much stress for them affecting adding more uncertainty into the statistic pool.

2.5. Aggression induction experiments

For the whole EEG recording, the subjects had to complete four mental tasks, including resting tasks (Task #1 and Task #4) and mental active tasks (Task #2 and Task #3), the brief description of each task is explained in **Table 1**. Three trials session were recorded per experiment task for every subject. All trial sessions were conducted on the same day.

During experiment period, facial expression and posture of the subject while playing game and writing were recorded. This is important to record the moment where the subject was producing significant body movement that might induce noise in the signal recorded.

Table 1. List of tasks completed by each subject

Task #	Condition of Experiment
1	Resting state of subject before playing the game.
2	Playing "Subway Surfers" with sound muted.
3	Playing "Subway Surfers" with maximum volume sound.
4	Resting state of subject after finishing the game.

2.6. Filtering the raw EEG Signals

In this work, the Butterworth bandpass filter was used at cut-off frequencies of 0.5 Hz and 49 Hz to filter the signal from unwanted signals and noise. The Butterworth filter with orders 2, 4 and 6 was applied to filter the EEG signals of Task #1 to Task #4. To validate the effectiveness of different filter used, data loss was calculated by taking the ratio of the filtered data to the raw data. The 19 channels from both filtered and raw signal are being compared out of the 120 recording trials, which formed by 10 subjects performed 3 recordings for all 4 tasks (10 x 3 x 4=120).

$$\text{Data\_loss} = F_{ki} / G_{ki} \quad (1)$$

Where i = {1,2,3, …,*i*,…,19} and k = {1,2,3, …, k, …, 120}.

F is a filtered data and G is a raw data, *i* is the number of channel and *k* is the number of recording. The effect of Butterworth filter on the EEG signals was analysed by comparing the average time used for filtering and the ratio of data loss after filtering using the Butterworth filter.

2.7. Measuring the aggressiveness level, *A*

The point relationship of the task, R_k, was then compared to the baseline signal, R_B, to find out the differential signal, D_i, and was summed up to measure the aggressiveness level, *A*. The *D* is a row vector consisted of 171 values and *A* is a value that indicates the aggressiveness level.

$$D_i = \text{abs}\,(R_{iB} - R_{ik}) \quad (2)$$
$$A = \text{Sum}D \quad (3)$$

Where i = {1,2,3, …,i,…,19} and k = {1,2,3, …, k, …, 120}.

In measuring aggressiveness level, *A*, R_{ik} induced by the brain are compared with R_{iB} at the rest state. Two R_{iB} values are tested, which is the universal and the individual R_{iB}. The universal R_{iB} are measured by averaging the values of individual reference data.

3. RESULTS AND DISCUSSIONS

3.1. Universal and individual rest signal as the reference

Table 2. Aggressiveness level of subjects using universal rest data as reference

Subject	Task #1: Pre-game rest	Task #2: Mute gaming	Task #3: Max sound gaming	Task #4: Post-game rest	BPAI
1	24.70	33.27	35.05	36.39	32.60
2	16.89	48.28	43.69	25.06	42.68
3	35.72	<u>20.95</u>	<u>19.47</u>	35.67	48.48
4	21.96	<u>18.20</u>	23.05	20.41	32.74
5	47.05	<u>22.95</u>	<u>23.86</u>	46.26	22.01
6	14.22	21.54	22.22	36.14	41.99
7	14.63	<u>19.75</u>	22.22	37.14	36.51
8	47.76	<u>22.50</u>	<u>20.91</u>	45.93	53.81
9	34.16	<u>22.18</u>	<u>24.03</u>	28.21	31.27
10	50.01	<u>46.22</u>	<u>42.97</u>	75.48	29.57

Table 2 shows the aggressiveness level for all 10 subjects. The hypothesis is the aggressiveness level at the highest level at Task #2 and #3. However, look at the subject #3, the aggressiveness level at the Task #2 and #3 lower than the Task #1 and #4. The aggressiveness level at the Task #1 and #4 should be lower because the subject is at the relaxing state.

The further investigation is conducted to use the reference value of individual subject instead of using the universal reference. The investigation results are shown in Table 3. The table shows that for all 10 subject, the aggressiveness levels at the resting states are lower compared with the gaming state. The aggressiveness level at Task #1 and Task #4 are lower than the level at the Task #2 and #3.

The reason it happens is the effect of the individuality when inducing aggression. Further investigation could be conducted to see how far is the aggression level different among subjects.

Table 3. Aggressiveness level of subjects using individual rest as reference

Subject	Task 1: Pre-game rest	Task 2: Mute gaming	Task 3: Max sound gaming	Task 4: Post-game rest	BPAI
1	15.51	40.19	40.33	23.50	32.60
2	18.99	51.21	46.52	29.55	42.68
3	27.08	35.34	34.21	19.95	48.48
4	13.98	39.24	52.53	27.03	32.74
5	10.89	40.56	39.54	14.32	22.01
6	26.76	59.88	55.49	47.70	41.99
7	15.74	43.95	47.74	37.48	36.51
8	10.97	79.88	77.75	12.81	53.81
9	17.49	28.99	32.80	48.84	31.27
10	11.72	56.81	43.16	52.10	29.57

4. CONCLUSIONS

The paper proposes the design of an experimental protocol and the signal processing approach to measure aggressiveness level of the human subject. An experimental protocol is designed to induce aggression and the brain signals are measured by using EEG. The filtering technique which is Butterworth filter are employed in the investigation to remove the noises. Several parameters of Butterworth filter are tested in the experiments and the performance is measured by using Data_loss measure, however, its do not affect the performance the filter. The proposed aggressiveness levelling technique is capable to index the aggressiveness level and could be used for conducting further investigation in the future.

ACKNOWLEDGEMENT

The author would like to acknowledge the support from the Fundamental Research Grant Scheme (FRGS) under a grant number of FRGS/1/2019/ICT04/UNIMAP/02/1 from the Ministry of Education Malaysia.

REFERENCES

1. Buss 1961 *The Psychology of Aggression* (New York: Wiley).
2. Richardson DS Hammock GS 2003 Defining non-direct aggression raises general questions about the definition of aggression *Int. Review of Social Psychology* 16 pp 5–10.
3. Ramirez JM Andreu JM 2006 Aggression, and some related psychological constructs (Anger, Hostility, and Impulsivity) Some comments from a research project *Neuroscience and Biobehavioral Reviews* 30(3) pp 276–91.
4. Gothelf D Apter A Van Praag HM 1997 Measurement of aggression in psychiatric patients *Psychiatry Research* 71(2) pp 83-95.
5. Paulraj MP Xu TK 2013 Application-specific multichannel. EEG data acquisition system for portable biomedical application.
6. Abootalebi V Hassan M Khalilzadeh MMA 2009 A new approach for EEG feature extraction in P300-based lie detection *Computer Methods and Programs in Biomedicine* 94(1) pp 48–57.
7. Xu T K Paulraj M P 2015 Aggressiveness level assessment using EEG inter channel correlation coefficients *Indian Journal of Science and Technology* 8(21) pp 1-10
8. *Hema, C. R., & Osman, A. A. (2010, May). Single trial analysis on EEG signatures to identify individuals. In 2010 6th International Colloquium on Signal Processing & its Applications* (pp. 1-3)
9. Kai Xu Tung, Wan Khairunizam, Wan Azani et. al (2020): *Preliminary Analysis of Human Behaviors Based on Buss-Perry Questionnaire Score for Designing Measures of Aggression Lecture Notes in Mechanical Engineering* book series (LNME), https://doi.org/10.1

Authors Introduction

Dr. 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 Faculty of Electrical Engineering Technology, Universiti Malaysia Perlis. His research interest is in Human-Computer Interaction (HCI), Intelligent Transportation System, Artificial Intelligence and Robotics.

Mr. Kai Xu tung

He recieved his Degree in mechatronic engineering from Universiti Malaysia Perlis in 2012. He is currently a Ph.D student at Universiti of Malaysia Perlis. His research interests include signal processing and application of EEG, particularly on brain activity for human rest and aggressiveness.

Mr. Lugieswaran A/L Munian

He received his B.Eng with Honours degree in Mechatronic Engineering from Universiti Malaysia Perlis in 2021. He is currently a Process Engineer at SMT Technology. His research interests include Automation, Image Processing, Machine Learning, Artificial Intelligence, and Robotics.

Dr. Wan Azani Mustafa

Ts. Dr. Wan Azani Mustafa obtained his PhD in Mechatronic engineering from Universiti 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.

Dr. Hashimah Ali

Hasimah Ali is a senior lecturer at Faculty of Electrical Engineering Technology, UniMAP. She received her PhD in Mechatronic Engineering from UniMAP (2015), MSc in Mechatronic Engineering from IIUM (2008) and BEng in Mechatronic Engineering from IIUM (2004). Her research interests are signal and image processing (facial expression), ground penetrating radar, robotics gripper and IoT. She is responsible for Manager Division in Signal Processing and Artificial Intelligent and Head of Ground Penetrating Radar Research Group.

Dr. Zuradzman M. Razlan

He received his Bachelor of Mechanical Engineering from Yamagata University, Japan from Apr. 1993-Mar. 1997 and Ph.D. in Engineering from Mie University, Japan. He is currently an Assoc. Prof. at Faculty of Mechanical Engineering Technology, Universiti Malaysia Perlis. His research interest is in Energy, Thermo-Fluid, Two Phase Flow, Air Flow System, Heat-Pump and refrigeration system.

Dr. Shahriman AB

Shahriman A.B. received the B.Eng, M.Eng, and PhD in Mechanical engineering from Mie University, Mie, Japan in 1997 2004, 2010 respectively. He is currently an Assoc. Prof. at Faculty of Mechanical Engineering Technology, Universiti Malaysia Perlis.. His current research interest is in Medical Engineering, Industrial Agriculture Automation and sustainable energy.

Dr. Norrima Mokhtar

She received the B.Eng. degree from University of Malaya, the M.Eng. and the Ph.D. degree from Oita University, 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.

Voice User Interface(VuI) Smart Office Door Application in the Context of Covid-19 Pandemic

Muhammad Zharif Aiman Alias, Wan Norsyafizan W. Muhamad
School of Electrical Engineering, College of Engineering, University Teknologi MARA, Malaysia

Darmawaty Mohd Ali
School of Electrical Engineering, College of Engineering, University Teknologi MARA, Malaysia

Azlina Idris
School of Electrical Engineering, College of Engineering, University Teknologi MARA, Malaysia

E-mail:syafizan@uitm.edu.my

Abstract

Nowadays, all countries around the globe are trying their best to prevent the spread of Covid-19 from reaching their people. However, the virus spread is through the transmission of close contact with the infected person and touching a surface that already contains the virus. This project presents Voice User Interface (VUI) Smart Office Door System that interact with the Internet of Things (IoT) to execute the command given by the user. This project use Raspberry Pi as microcontroller which Thonny Python software used for running the coding script of the system. The electrical components such as dc motors and LEDs are connected to General-Purpose Input Output (GPIO) pin of Raspberry Pi and motor driver. One of the dc motors used for controlling the lock and another one used for controlling the door. Blue LED used as locked door notification while green LED used as unlocked notification. Microphone and speaker connect to the Raspberry Pi through usb port and Bluetooth. This system used Google Assistant as its VUI to control the smart door contactless. The user can give the command to the system through microphone to control the output of the system. The purpose of this project is to invent the smart door with IoT technology that help prevent the spread of Covid-19 at the workplaces.

Keywords:Voice User Interface (VUI), Internet of things (IoT), Google Assistant, voice command and dc motor

1 INTRODUCTION

Nowadays, the novel coronavirus, also known as COVID-19, became unpredictable and caused a massive loss to people worldwide. Coronavirus can be spread through the surface appliances that have been touch by the infected person. The virus can live on that surface at most 48 hours. The workers get affected by the virus when using the same appliances at the workplace such as door, switch, water dispenser and photostat machine. The invention of contactless appliance has been emphasized as it will help reduce the possibility of the people get affected by the virus caused by touching appliance's surfaces. The appliance can be controlled by voice command with the help of Voice User Interface (VUI), Internet of Thing (IoT) and smart controller [1].

Voice user interface (VUI) is a speech recognition technology that allows people to interact with a computer, smartphone, or other devices through voice commands [2][3]. Google Assistant is one of the VUI that are available in the market today [4]. This project propose the voice command to control the smart office door when interacting with the Internet of Things (IoT). The smart door will help people to eliminate the needs to touch the surface. To develop a voice-based command controller, a microphone will be used in data transmission communication and a voice signal processor to implement the voice recognition method.

Google Assistant is an artificial intelligence-powered virtual assistant developed by Google[5][6]. Users interact with the Google Assistant through natural voice or keyboard input. The Google Assistant can search the Internet, schedule events and alarms, adjust hardware settings on the user's device, and show information from the user's Google account. In the future, Google says that the Google Assistant will identify objects, gather visual information through the device's camera, support purchasing products and identify songs. For this project, Google Assistant will be used as VUI to control the smart office door with the help of IoT technology.

The Internet of Things brings a new node to a marketplace and knocks on the door with new options for innovations[7]. IoT is a system that integrates with sensors, computers, and digital devices through the internet and connects with each other to exchange and transmit information via special Identifications (IDs). The emphasis on automation technology has risen significantly with the new invention of automation appliances [8]. Some automation technology use voice command as input because it is more efficient compared to the sensors. In IoT, the voice command signal must be transmitted immediately between adjacent nodes throughout the IoT-enabled network. Thus, the nodes must be in the overlapping communication range of one another for this to occur. The IoT-based system helps the user to control the smart home system without ever having any physical contact [9][10]. The IoT emphasizes the use of ambient technology, robotics and IoT to reduce screen-based interaction, meaning that the user interface is blended with the physical gadget. Besides IoT, microcontroller such as Raspberry Pi plays the main role for the whole system.

The Raspberry Pi is a Linux single-board computer which use SD card as storage. It has an Ethernet, two USB interfaces, a USB interface, HDMI (support sound output) and RCA terminal output support. One of the applications of Raspberry pi is to operate a system that operates itself and can access the wireless network and Bluetooth chips[11].The Raspberry Pi has also been used as a processing chip and underlying architecture to build up a personal assistant[12][3][13]. The proposed voice-operated personal assistant focused on Raspberry Pi gives individuals more comfort and ease.

There are several works in literature that proposed the voice command system to control appliances. Mtshali et al presented a smart home appliance control system for physically disabled people [14]. The system consists of a smart plug, smart camera, smart power strips and digital assistants such as Google Assistant, Siri, or Alexa. This system captured the voice command from the user and provide the output signal for controlling the targeted appliances, such as turning it on/off. Having this system could be a tremendous benefit for disabled people. This system also helps to enhance the quality of nondisabled people.

Nefy Puteri Novani and Mohamad Hafiz proposed electrical household appliances control using voice command based on a microcontroller [1]. In this system, EasyVR Commander been used as a sensor that can receive voice input to control electrical appliances such as fan, light, and door. Then, voice recognition method been implemented to captured voice signal to recognize the type of voice pattern. This method is important to provide safety feature to the system. This system only allows the homeowner to give the command to control electrical appliances in that house.

Yash Mittal et al proposed a voice-controlled multi-functional smart home automation system [2]. This system allows users to use voice command to control the home appliances and gadgets for different functions and purposes. This system can adapt to the user's voice and recognize the voice command, depending on the speakers' accents. Arduino microcontroller and dedicated hardware module used in this system for commands processing and control the appliances. This system aims to be affordable, flexible, and robust.

For the propose smart office door, the office door, microcontroller, motor driver, dc motor, LED and microphone are integrated as a smart door control system. The proposed smart office door system provides users with a feature that can control the office door contactless using voice command [15]. This system consists of several modules: LED, dc motor, motor driver, microphone, IoT, operating system device, voice user interface (VUI), and Raspberry pi. The Raspberry Pi 400 is the the microcontroller for the whole system. The Google Assistant is implemented into the microcontroller and provide the ability of voice command controller with the help of IoT. For the door system, there will be two parts that will be controlled by the user, which are the lock of the door and the door itself. User can control the system's lock through the voice command system by giving the specific command and opening the door only when the door is unlocked by given a specific command. There will be an LED notification to let the user know either the door is locked or unlocked. This system provides touchless feature for opening/closing of the door which prevent spread of Covid-19.

2 METHODS

2.1. Block Diagram

As mentioned, the proposed Voice User Interface (VUI) Smart Office Door provides user with the feature to control the office door using voice command. Figure 1 shows the block diagram of the VUI Smart Office Door system. This block diagram consists of several modules: voice command, Raspberry Pi, Cloud, Google Server, LED, motor driver, and dc motor. In this project, Raspberry Pi 400 is used as a microcontroller for the system.

Figure 1 shows the Block diagram of the proposed VUI smart door system. A microphone captures the voice command from the user and convert it into an input signal for the system. Raspberry Pi which is a microcontroller for the system will receive the input signal and sent it to the Cloud through the internet. Cloud refers to the servers that are accessed over the internet, and software and databases that run on those servers. The Cloud will directly transfer the signal to the Google Server for processing. Google Server function is to execute the coding script when the event fulfils the condition. In this system, when Google Server received the signal, it executes the output signal based on the function setup. The output signal will be sent to the Raspberry Pi via an internet connection. Following that, the microcontroller distributes the output signal to the user and the targeted components such as LED and DC motor to execute the command given by users. For controlling the dc motor, the signal should be sent to the L298N motor driver as it controls the movement of the motor.

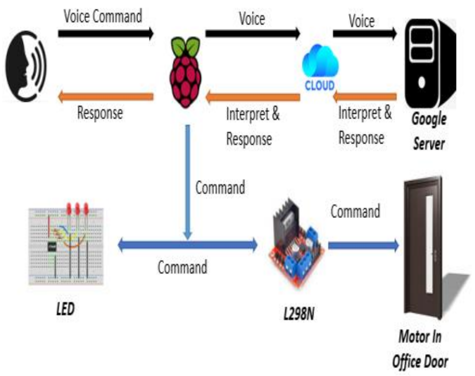


Figure 1: Block diagram show the operation of Voice User Interface Smart Office Door *Flowchart*

Figure 2 illustrate the flowchart of the proposed system. To access the Smart Office Door system, user needs to say "Hello Friend". This command is used as a password for security purpose to the Smart Office Door system. When password command been detected by the system, GPIO pin 21 at Raspberry Pi will produce the high output signal to turn on the LED notification for the locked door and give the user access to the Smart Office Door system. This LED notification will let the user know the door is locked. After getting access to system, the user needs to give the command "Unlock the Door" to the system. When this command is detected, the lock motor will rotate to the right to unlock the door. The GPIO pin 3 will produce a high output signal, and GPIO pin 2 will produce the PWM wave output signal with half of the max value for 0.7 seconds. These output signals will be sent to the L298N motor driver to rotate the lock motor in the right direction with half of the speed for 0.7 seconds to unlock the door. Next, GPIO pin 2 will produce the PMW wave signal with zero value, and this signal will be sent to the motor driver to stop the lock motor rotation. Then, GPIO pin 21 will produce the low output signal to turn off the LED notification for the locked door, and GPIO pin 20 will produce the high output signal to turn on the LED notification for the unlocked door. This LED notification will let the user know the door is unlocked.

Users need to give command "Open the Door" to open the smart door. When this command is detected by the system, the smart door will open, and after several seconds, the smart door will automatically close. The GPIO pin 22 will produce a high output signal, and GPIO pin 7 will produce a PWM wave output signal with half of the max value for 0.8 seconds. Both signals will be sent to the motor driver to rotate the door motor in the left direction with half the speed for 0.8 seconds to open the door. Next, the GPIO pin 7 will produce the PWM wave output signal with zero value for 2 seconds. This signal will be sent to the motor driver to stop the door motor rotation for 2 seconds. Then, the GPIO pin 22 will produce the low output signal, GPIO pin 25 will produce the high output signal, and GPIO pin 7 will produce the PWM wave output signal with half of the max value for 0.8 seconds. These signals will be sent to the motor driver to rotate the door motor in the right direction with half of the speed for 0.8 seconds to close the door. Lastly, GPIO pin 7 will produce the PMW wave signal with zero value, and this signal will be sent to the motor driver to stop the door motor rotation.

Users need to give the command "Lock the Door" to lock the smart door. When the system detects this command, the lock motor will rotate to the left side to lock the smart door. The GPIO pin 3 will produce a low output signal, GPIO pin 4 will produce high output signal and GPIO pin 2 will produce the PWM wave output signal with half of the max value for 0.7 seconds. These signals will be sent to the motor driver to rotate the lock motor in the left direction for 0.7 seconds to lock the smart door. Next, GPIO pin 2 will produce the PMW wave signal with zero value, and this signal will be sent to the motor driver to stop the lock motor rotation. Then, the GPIO pin will produce a high output signal to turn on the LED notification of the locked door, and GPIO pin 20 will produce the low output signal to turn off the LED notification for the unlocked door.

This LED notification will notify user that the door is locked.

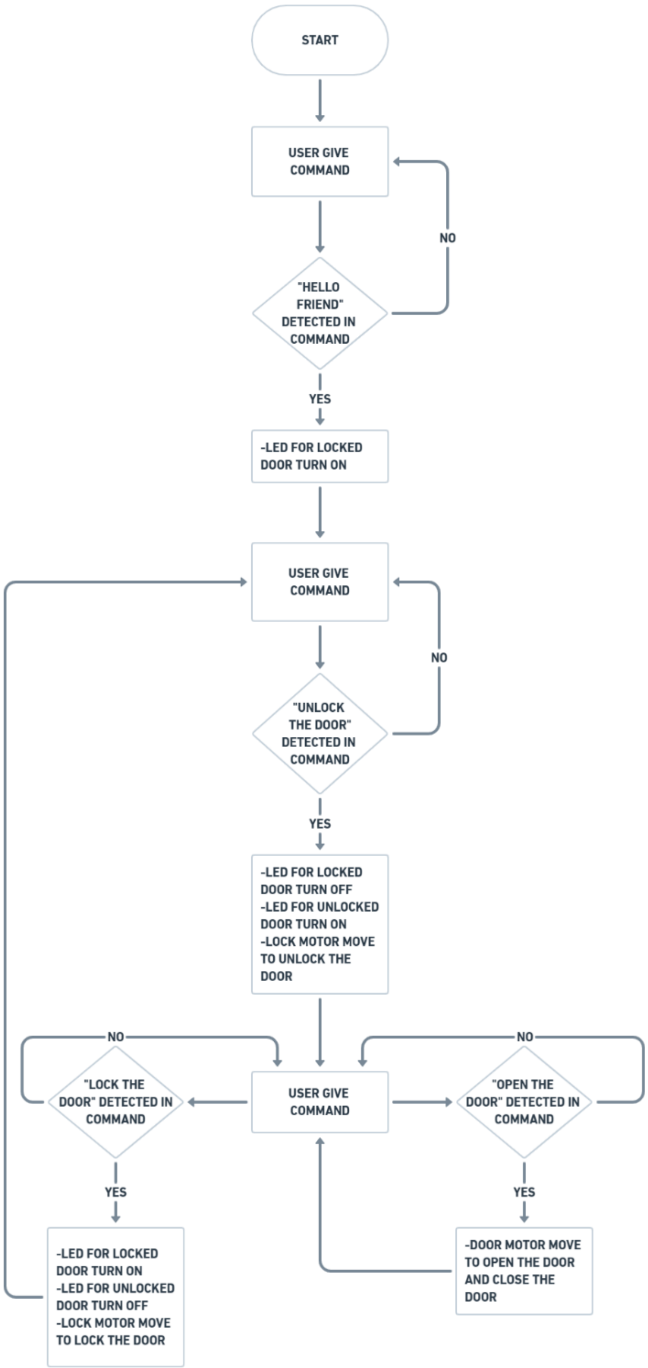


Figure 2: Flowchart of VUI Smart Office Door.

2.2. Circuit Diagram

Figure 3 illustrates the circuit diagram of the Voice User Interface (VUI) Smart Office Door system. As mentioned, this system has one input which is the microphone. Microphone used to capture the voice command from the user. This microphone will be connected to the Raspberry Pi 400 using a USB port.

As for the internet connection, the system will use the build-in Wi-Fi receiver in Raspberry Pi to interact with the Google server to execute the user's command. For the output part, this system will have four output which are 2 LEDs for notification purpose and two dc motors for controlling the lock and the door. LEDs notification will connect to GPIO pin 20 and 21, and dc motors will connect to the motor driver.

Then, the motor driver will connect to GPIO pin 2, 3 and 4 to control the lock motor and connect to the GPIO pin 7, 22 and 25 to control the door motor. The resistor used in this circuit reduce the voltage that are received by the LEDs notification.

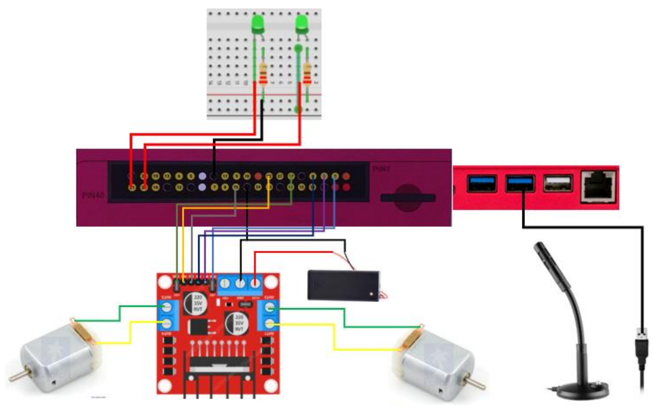


Figure 3: Circuit diagram of Smart Office Door Combine with Voice User Interface (VUI) and Internet of Things (IoT)

3 RESULTS AND DISCUSSIONS

3.1. Hardware Development

For hardware development, the initial stage is shown in Figure 4.1, where the dc motors are connected to the motor driver. Then, motor driver, microphone, and LEDs are connected to the microcontroller. The Voice User Interface Smart Door system used power supply 5A from the power socket to run the system. As for controlling the dc motors, the motor driver will connect to an extra power supplier, which is batteries with 6V, as shown in Figure 4.2.

Dc motors is connected to the door and lock using gears as a connecter to transfer the rotation movement from motors to the door and lock. Connection of dc motor with the door is shown in Figure 4.3 and connection for dc motor with lock shown in Figure 4.4. LEDs notification is placed above the door, as shown in Figure 4.5. The connection between LEDs and motor driver with Raspberry pi is executed using breadboard as shown in Figure 4.6. The breadboard and motor driver are mounted below the model. Figure 4.7 shows the microphone is connected to Raspberry Pi via the USB port.

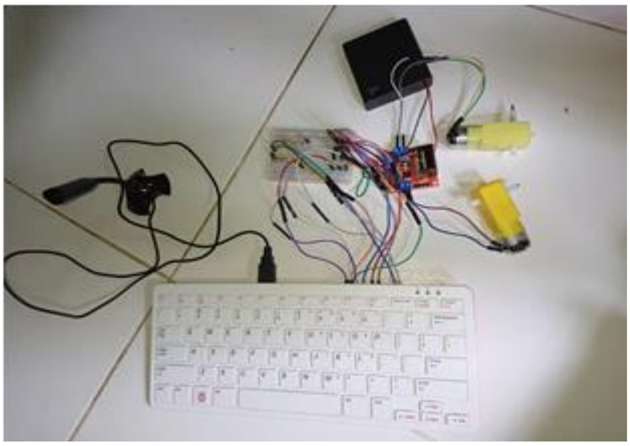


Figure 4.1: Hardware before mounted to Smart Office Door model

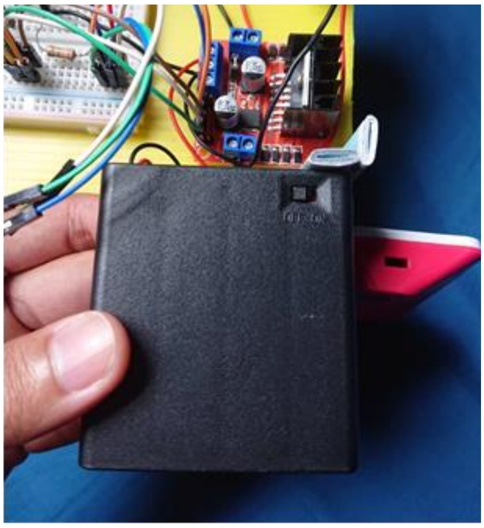


Figure 4.2: Batteries supply to motor driver.

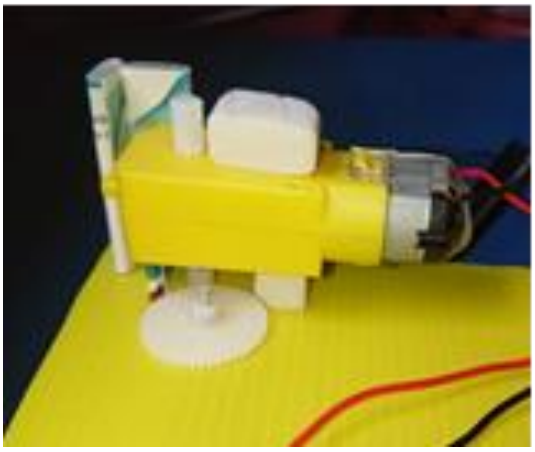


Figure 4.3: Dc motor connect to the door.

Figure 4.4: Dc motor connect to the lock.

Figure 4.5: Office room model with LED's positioned above the door.

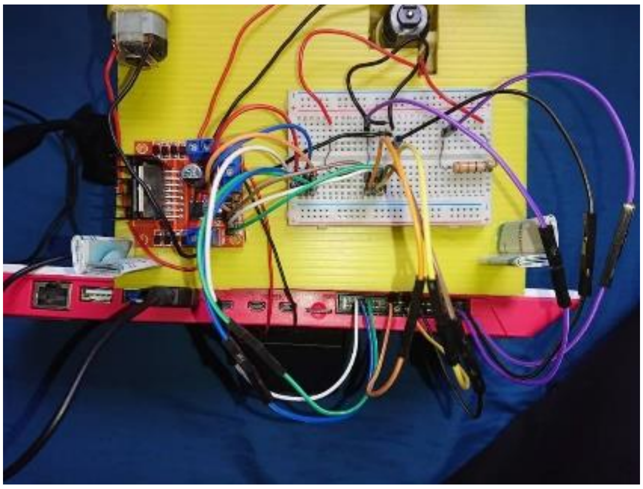


Figure 4.6: The circuit connection in the model.

Figure 4.7: Microphone connect to the Raspberry Pi

3.2. VUI Smart Office Door locked state

The model of the VUI Smart Office Door in the locked condition is shown in Figures 5.1 and 5.2. The blue LED notify users that the door has been locked. Figure 5.2 demonstrates the lock's position, which the door is in the locked state. This state is executed with voice user's command 'lock the door'. When this command is detected, the green LED is turned off and blue LED is turned on to notify user that the door has been locked.

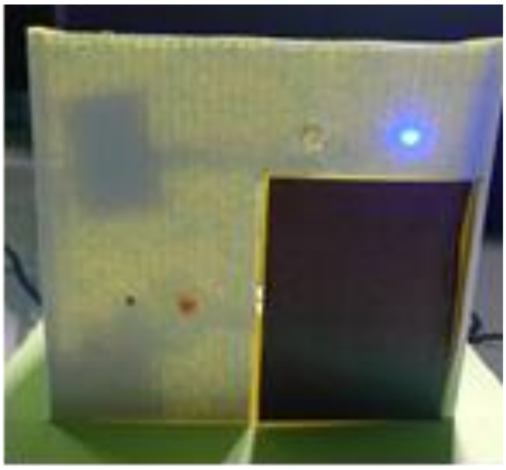


Figure 5.1: Notification for locked door condition.

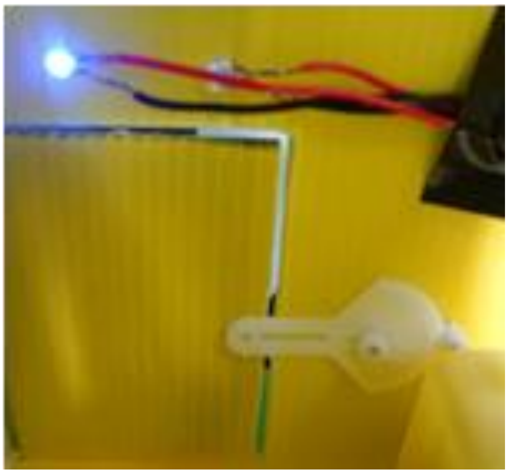


Figure 5.2: Position of the lock in locked door state.

3.3. VUI Smart Office Door unlocked state

Figure 6.1 and Figure 6.2 illustrates the model of VUI Smart Office Door in unlocked state. For this condition, the blue LED is turned off and the green LED is turned on to alert the user that the door has been unlocked. Figure 6.2 illustrates the lock's position, which permits the door to be opened. This state is executed with voice user's command 'unlock the door'. In unlocked state, user can give command in order to open the door. When the command of "opens the door" has been detected, the door will be opened as shown in Figure 6.3 and it will be closed automatically after 2 seconds, as set in the algorithm.

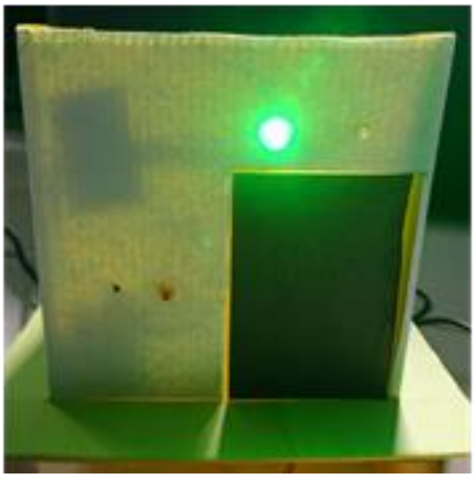


Figure 6.1: Notification for unlocked door condition.

Figure 6.2: Position of the lock in unlocked state.

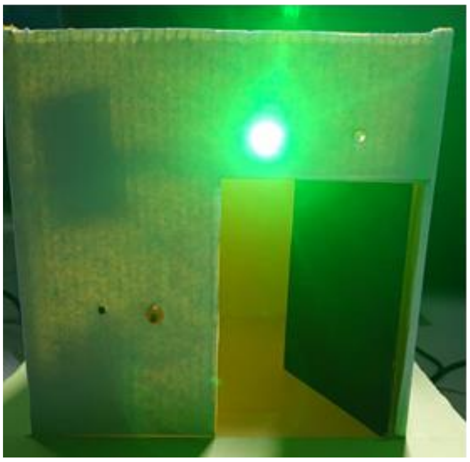


Figure 6.3: The door position when been opened.

4 CONCLUSION AND RECOMMENDATION

There are many ways to control the office door, such as using an infrared sensor or using a weight sensor. In this project, Voice User Interface Smart Office Door had been proposed to control the office door using voice command. The use of IoT technology in this system provides the feature of controlling the office door just by using voice command with the help of Google Assistant. This system helps to eliminate the need for the workers to touch the appliances. This system also helps prevent the spread of coronavirus among the workers and provide a safer environment for working in the office during this pandemic.

For the recommendation, this system still has a room that can be further upgraded by linking the system with the apps that allow user to monitor the appliances from long distances and control them for emergency purposes. This system also needs to add a special controller for mute people to use. As known by everyone, mute people cannot talk thus they will have difficulties when using this system.

REFERENCES

[1] N. P. Novani, M. H. Hersyah, and R. Hamdanu, “Electrical household appliances control using voice command based on microcontroller,” *2020 Int. Conf. Inf. Technol. Syst. Innov. ICITSI 2020 - Proc.*, pp. 288–293, 2020,

[2] S. Sharma, D. Singhal, and R. Gupta, “A Voice-Controlled Multi-Functional Smart Home Automation System,” pp. 1–6, 2015.

[3] D. Yang, L. Ma, and F. Liao, “An Intelligent Voice Interaction System Based on Raspberry Pi,” *Proc. - 2019 11th Int. Conf. Intell. Human-Machine Syst. Cybern. IHMSC 2019*, vol. 1, pp. 237–240, 2019,

[4] V. Kepuska and G. Bohouta, “Next-generation of virtual personal assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home),” *2018 IEEE 8th Annu. Comput. Commun. Work. Conf. CCWC 2018*, vol. 2018-Janua, no. c, pp. 99–103, 2018,

[5] K. Bhanuabhiram, L. T. Kumar, and N. Srinivasan, “Google assistant controlled home automation,” *J. Comput. Theor. Nanosci.*, vol. 16, no. 8, pp. 3259–3264, 2019,

[6] L. K. Vanathi, B. Mahalakshmi, S. Madhusudan, M. Srinivasaperumal, S. Srikanth, and R. Sathish Kumar, “Smart Control of Home Amenities Using Google Assistant and Clap Switch Circuit,” *2019 5th Int. Conf. Adv. Comput. Commun. Syst. ICACCS 2019*, pp. 350–352, 2019,

[7] S. Chaudhary, R. Johari, R. Bhatia, K. Gupta, and A. Bhatnagar, “CRAIoT: Concept, Review and Application(s) of IoT,” *Proc. - 2019 4th Int. Conf. Internet Things Smart Innov. Usages, IoT-SIU 2019*, pp. 2–5, 2019,

[8] S. K. Vishwakarma, P. Upadhyaya, B. Kumari, and A. K. Mishra, "Smart Energy Efficient Home Automation System Using IoT," *Proc. - 2019 4th Int. Conf. Internet Things Smart Innov. Usages, IoT-SIU 2019*, pp. 17–20, 2019,

[9] A. K. Gupta and R. Johari, "IOT based Electrical Device Surveillance and Control System," *Proc. - 2019 4th Int. Conf. Internet Things Smart Innov. Usages, IoT-SIU 2019*, 2019,

[10] A. Jain, P. Tanwar, and S. Mehra, "Home Automation System using Internet of Things (IOT)," *Proc. Int. Conf. Mach. Learn. Big Data, Cloud Parallel Comput. Trends, Prespectives Prospect. Com. 2019*, pp. 300–305, 2019,

[11] P. Vashistha, J. P. Singh, P. Jain, and J. Kumar, "Raspberry Pi based voice-operated personal assistant (Neobot)," *Proc. 3rd Int. Conf. Electron. Commun. Aerosp. Technol. ICECA 2019*, pp. 974–978, 2019,

[12] P. Singh, P. Nayak, A. Datta, D. Sani, G. Raghav, and R. Tejpal, "Voice Control Device using Raspberry Pi," *Proc. - 2019 Amity Int. Conf. Artif. Intell. AICAI 2019*, pp. 723–728, 2019,

[13] C. Y. Peng and R. C. Chen, "Voice recognition by Google Home and Raspberry Pi for smart socket control," *Proc. - 2018 10th Int. Conf. Adv. Comput. Intell. ICACI 2018*, pp. 324–329, 2018,

[14] P. Mtshali and F. Khubisa, "A smart home appliance control system for physically disabled people," *2019 Conf. Inf. Commun. Technol. Soc. ICTAS 2019*, pp. 1–5, 2019,

[15] M. D. A. P. Mummaka Sai Srinath, Manepalli Nanda Kishore, "Interactive Home Automation System With Google Assistant," *Int. J. Pure Appl. Math.*, vol. Vol 119, no. 12, p. Page 12, 2018.

Authors Introduction

Mr. Muhammad Zharif Aiman bin Alias

He received his B. Eng. degree in Electrical Engineering from MARA University of Technology in 2021. He is currently a Technical Trainee under BriTay MyStep programme at BriTay Asia (M) Sdn Bhd. His final year project about Voice User Interface (VUI) SMART Office Door.

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 is in the area of wireless communication (Physical and MAC cross layer optimization),

Dr, Darmawaty Mohd Ali

She is an Associate Professor in the Department of Communication Engineering, Universiti Teknologi MARA (UiTM). She obtained her Ph.D in 2012 from Universiti Malaya, Malaysia. She has received the Master of Engineering in Electrical in 2002 from Universiti Teknologi Malaysia. Previously, she obtained her first degree from Universiti Kebangsaan Malaysia with Honours, in Electrical, Electronic and System, graduating in 1999. She is a member of Wireless Communication Technology (WiCOT) Research Interest Group (RIG) and her research interests include Wireless Access Technology and Quality of Service in Wireless Broadband.

Dr. Azlina Idris

She is a Senior Lecturer in the Universiti Teknologi MARA (UiTM). She obtained her Ph.D in 2012 from Universiti Malaya (UM), Malaysia. She has received the Master of Engineering in Electrical from Universiti Teknologi Malaysia (UTM). Previously, she obtained her first degree from Leeds Metropolitan University, United Kingdom with Honours, in Applied Computer. She is a member of Wireless Communication Technology (WiCOT) Research Interest Group (RIG) and her research interests include OFDM/OFDMA transmission, single and multiuser precoding, modulation, MIMO transmission techniques and receivers, channel coding, interference management and mitigation, and channel modelling (channel estimation).

Endometrial Cell Images Segmentation: A Comparative Study

Wan Azani Mustafa*
Advanced Computing (AdvComp), Centre of Excellence (CoE), Universiti Malaysia Perlis, Pauh Putra Campus, 02600 Arau, Perlis, Malaysia

Nurul Umaira Salim
Faculty of Electrical Engineering Technology, UniCITI Alam Campus, Universiti Malaysia Perlis, 02100 Padang Besar, Perlis, Malaysia

Wan Khairunizam
Faculty of Electrical Engineering Technology, UniCITI Alam Campus, Universiti Malaysia Perlis, 02100 Padang Besar, Perlis, Malaysia

Shahrina Ismail
Faculty of Science and Technology, Universiti Sains Islam Malaysia (USIM), Bandar Baru Nilai, 71800, Negeri Sembilan, Malaysia

Hiam Alquran
Biomedical Systems and Medical Informatics Engineering, Yarmouk University, 21163 Irbid, Jordan

Corresponding Author: wanazani@unimap.edu.my

Abstract

Uterine cancer, also known as endometrial cancer, is a form of cancer that affects the female reproductive system. Nowadays, there are 2 step methods that the physician or health care provider tend to use to diagnose cancer, which is using ultrasound technique and endometrial biopsy. The biopsy procedure is used to extract the cell and sent to the pathologist for histopathological image analysis. The histopathological image analysis is the crucial step in all the procedures because it determines the situation for the patient, whether positive or negative. They are two types of cell images known as high grade squamous intraepithelial lesion (HSIL) and low grade squamous intraepithelial lesion (LSIL). The problem occurs when both LSIL and HSIL are different, needing different medical treatment techniques but showing slighter differences in nucleus size cell histopathological image analysis. Therefore, the pathologist usually requires more time to identify whether it is LSIL or HSIL. Based on the limitation, the paper aims to compare a few popular detection methods, which are the Wolf method, Bernsen method, Otsu method and Feng method. Based on the Image Quality Assessment (IQA), the Wolf method shows good performance compared to the others. In a precise term, this finding could benefit the health care community to reduce the diagnosis time to categorize the cell and lead to early treatment of endometrial cancer.

Keywords: Uterus, Endometrial, Cancer, Detection, Segmentation, Nucleus

1 INTRODUCTION

Endometrial cancer or uterus cancer is the fifth-deadliest cancer, accounting for more deaths than any other cancer of the female reproductive system[1]–[3]. Women above the age of 50 are more likely to get uterine cancer. It is estimated that around 21,750 women will receive a new diagnosis of ovarian cancer this year alone, and approximately 13,940 of those women will die from it. Like other forms of cancer, this also benefits from earlier detection because treatment could start as soon as possible[4]. Endometrial cancer develops when cells in the endometrium (the inner lining of the uterus) begin to proliferate uncontrollably. Cells in almost any part of the body can develop into cancer and spread to other parts of the body. The uterus is a hollow organ that is approximately the size and shape of a medium-sized pear. When a woman is pregnant, the uterus is where the fetus grows and develops. It is divided into two sections (as shown in Figure 1)[5], [6].

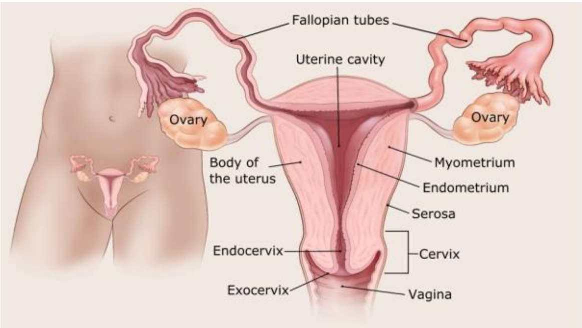


Fig. 1. Structure of Uterus

Due to the complexity of human cells, some pathologists may have a hard time differentiating between Low-Grade Squamous Intraepithelial Lesion (LSIL) and High-Grade Squamous Intraepithelial Lesion (HSIL) cells[7]. The outcome for this problem is that the pathologist may take a longer time to diagnose the cell. This might cause the pathologist to miss diagnosing the condition of the cell. However, by quickly categorizing the type of cell between LSIL and HSIL, quick treatment can soon be started for that particular patient. Thus, like all cancer, the patient can be saved with quick treatment.

Currently, they are 2 step methods that the physician or healthcare provider uses to diagnose cancer, which is by using ultrasound technique and endometrial biopsy. If the patient is suspected of having cancer, ultrasound is first used to measure the length of the endometrial lining. Then, if the lining is greater than four millimetres, the endometrial biopsy will be performed. At this stage, a small tubular instrument is gently placed inside through the endocervical canal into the endometrial cavity. This instrument is used to extract the cell and sent to the pathologist for histopathological image analysis.

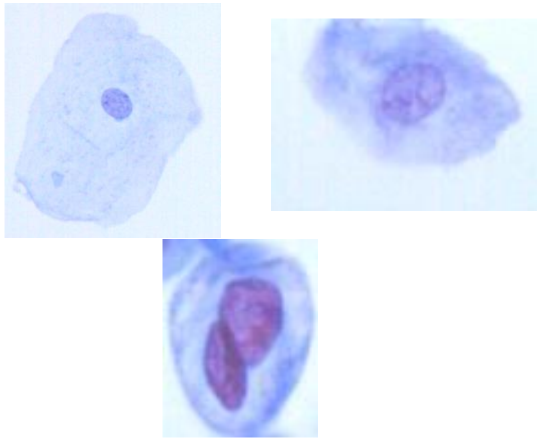


Fig. 2. Sample image of endometrial cell

The histopathological image analysis is the most important step in all the procedures because it determines the situation for the patient, whether positive or negative[8], [9]. They are two types of cell results, known as LSIL and HSIL. The difference for LSIL is a low grade that can be false positive even though the test comes as a low grade but may not tell the abnormality on the cell. For HSIL or the high grade, it always points to something abnormal that can point to cancer. The HSIL are so important even if other tests are out negative. Furthermore, the physician will still have to look further into that case. Another approach employed by the doctor was content-based picture retrieval (CBIR). Using machine vision techniques, this methodology was utilized to locate pictures in large databases. Physicians may only give effective decision-making as a means of detecting illness so that patients can rest and receive treatment at the right moment. When the CBIR model receives a query, it extracts the same set of characteristics and compares them to the indexed features to find comparable pictures in the database. Medical pictures saved on both distributed and centralized servers are also utilized for teaching, information, and diagnosis.

The gold standard for identifying endometrial cancer is histological image analysis. This method involves using traditional machine learning techniques used in previous computer-aided diagnostic (CAD) methodologies[10], [11]. But this method frequently failed to produce satisfactory results. Hence, a new technique was introduced to detect the endometrial cancer cell involved in developing a CAD approach based on a convolutional neural network (CNN) and attention mechanisms, called HIENet. A computer-aided diagnostic (CAD) system works intending to be able to identify endometrial cancer early.

By standardizing texture features, choosing multifeatured in a consistent manner, and presenting physicians with comparative distributions of extracted texture features, the CAD system aids reproducibility. In addition, the RGB images were gamma-corrected before being converted to HSV and YCrCb to ensure consistency.

2 METHODOLOGY

The algorithm is developed using MATLAB 2017 from Toshiba laptop (L50A) with processor Intel® Core™ i5-4200M CPU @ 2.50GHz. All the images were provided by an established database using the microscopic capturing technique. All the processed images are in color level, where the size of each image is random. Four methods which are Otsu, Wolf, Bernsen and Feng, are selected to be applied on cell images.

2.1 Otsu Method

The most known method is the classical segmentation invented by Nobuyuki Otsu in 1979[12]. It is also a global segmentation method that separates the pixels into foreground and background by choosing the optimal threshold. The algorithm works by exhaustively searching for the threshold that minimizes the weighted within-class variance. In other words, it maximizes the between-class variance. The weighted within-class variance of two classes are given in Equation 1 as below:

$$\sigma^2 p(t) = p_1(t)\sigma_1^2 + p_2(t)\sigma_2^2 \,. \tag{1}$$

This method works best when it is applied to images with a clear bi-modal pattern. Due to that, images with uneven illumination and shadow will not yield the best performance for this method.

2.2 Wolf Method

Wolf Method is a modification of the Sauvola Method, which is claimed to normalize the contrast and mean grey value of the image and compute the threshold as in Equation 2[13]:

$$T = (1-k)m + kM + k\frac{s}{R}(m-M), \tag{2}$$

where the value of k is fixed to 0.5, m and s represent mean and standard deviation, while M is the minimum grey value of the image and R is a maximum value of the standard deviations over all local windows. Thus, misleading in the calculation of binarization thresholds may happen when even a tiny, noisy patch influences M and R values, although this method performs better than other previous methods before it.

2.3 Bernsen Method

This method uses the contrast of images[14]. It includes the estimation of an average of the highest and lowest intensity values in the window. The local contrast of the window is calculated as in Equation 3:

$$C(i.j) = I_{max} - I_{min} \,. \tag{3}$$

The local contrast is compared to the threshold value based on the pixels, which are foreground and background. When the local contrast is less than the threshold value, the pixel is classified as background and vice-versa.

2.4 Feng method

This method calculates the notion of two local windows contained within one another instead of calculating the dynamic range of grey value's standard deviation from the whole image[15]. As a result, the binarization threshold is computed as follows:

$$T = (1-a_1)m + a_2\frac{s}{R_s}(m-M) + a_3 M, \tag{4}$$

where R_s is a dynamic range of grey value's standard deviation, m is a mean value, s is standard deviation, α is coefficient, while M is the minimum value of the grey levels.

3 IMAGE QUALITY ASSESSMENT (IQA)

IQA works to evaluate the magnitude of differences (or amount of similarity) between the original (reference) image and processed image[16]. In this study, a few selected IQAs were reviewed, including specificity, accuracy and precision. The IQA of specificity, accuracy and precision are generally described in terms of true positive (TP), true negative (TN), false negative (FN) and false positive (FP)[17], [18]. The diagnostic test is considered a true positive if a disease is proven present in the patient. Similarly, the test result is a true negative if the disease is proven absent in the patient. On the other hand, the false positive indicates a patient with no disease, but the diagnose test indicates the presence of disease. Similarly, the test result is false negative when the disease is absent for a patient with the disease. Therefore, the test results are opposite to the actual condition if both are false positive and false negative.

3.1 Specificity

As suggested in equation [5], specificity is the proportion of the true negatives correctly identified by a diagnostic test. It suggests how good the test is at identifying the normal (negative) condition.

$$specificity = \frac{TN}{TN+FP}. \quad (5)$$

The probability of diagnostic tests identifying patients with a particular disease without giving false positive results represents the numerical value of specificity. For instance, there is a 99% chance for the patient without a certain disease to be identified as negative when the specificity test is 99% and vice versa.

3.2 Accuracy

Based on equation (6), accuracy is the proportion of true results, either true positive or true negative, in a population. Thus, it measures the degree of veracity of a diagnostic test on a condition.

$$accuracy = \frac{TN}{TN+TP+FN+FP} \quad (6)$$

The value of accuracy numerically represents the proportion of true positive results (both true positive and true negative) in the selected population. Regardless of positive or negative, the test result is said to be accurate if the accuracy rate is 99%. In other conditions, the accuracy does not suggest to equally high as well, even though both sensitivity and specificity are high.

3.3 Precision

From equation (7), we can immediately see that precision talks about how precise/accurate the model is. In other words, out of those predicted positive, how many of them are actually positive. Thus, precision is a good measure to determine when the cost of false positive is high.

$$precision = \frac{TP}{TP+FP} \quad (7)$$

4 RESULT AND DISCUSSION

Detection is a well-known image segmentation method due to its wide application in digital image processing. For example, it is an effective technique for distinguishing objects from their surroundings. On the other hand, segmentation is a technique for binarizing images based on pixel intensities and is used as a preprocessing step in many applications. For instance, it could be used in medical image processing to detect an abnormal nucleus size for cancer diagnosis. In this study, Otsu, Bernsen, Feng, and Wolf segmentation techniques were studied and implemented on endometrial cancer images. Figure 3 shows the resulting performance comparison of four different segmentation methods.

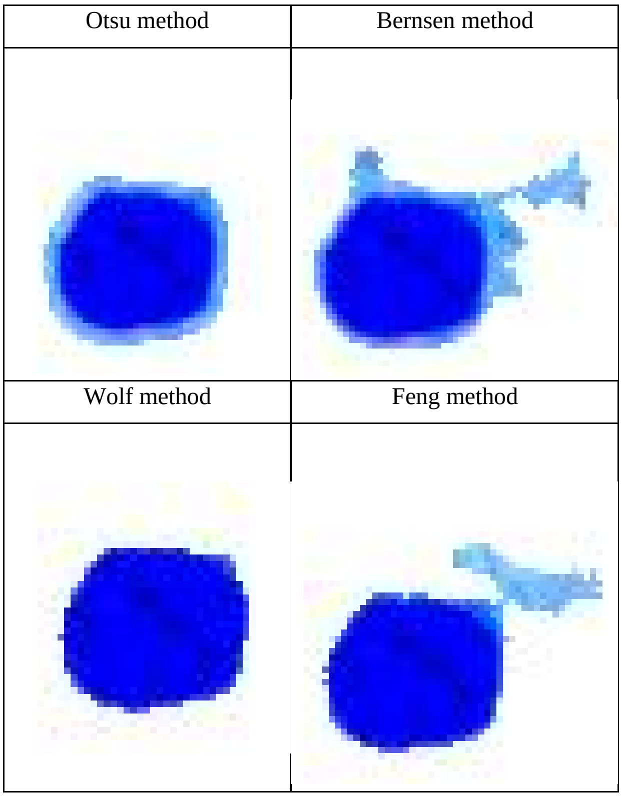


Fig. 3. Comparison images after applying the few selected methods

Based on Figure 3, result images for Otsu and Wolf segmentation have almost identical and obvious forms of nucleus region compared to the Bernsen and Feng method. Moreover, the Feng method also slightly detect unwanted or noise regions. To prove its effectiveness, the resulting images were evaluated using Image Quality Assessments (IQA), assessing the segmented images. In this experiment, IQA is a process to calculate the values of sensitivity (%), accuracy (%), and precision. Thus, the IQA can be used to monitor the image quality of the segmented image. The segmentation results after applying IQA are presented in Table 1.

Table 1. Comparison evaluation each segmentation techniques.

Methods	Specificity (%)	Accuracy (%)	Precision
Otsu	98.90	98.99	0.90
Bernsen	98.52	98.70	0.87
Wolf	99.73	99.76	0.97
Feng	99.40	99.45	0.94

Based on the Specificity, the Wolf method yields the highest value (99.73%), followed by the Feng method (99.40%). In contrast, the lowest value came from the Otsu method (98.90%). In terms of accuracy and precision, again, the Wolf method achieved 99.76% and 0.97, respectively.

5 CONCLUSION

Current development in terms of image processing mostly utilized the CAD approach. Also, the use of CAD systems such as deep learning could increase the accuracy of the result. In this study, the main purpose is to detect the segmented cell nucleus. The segmentation method will be applied to the image so that the nuclear cell image can be more clearly visible for easy recognition and diagnosis. The findings of this study suggest that the Wolf method is an effective approach that can be applied to detect the nucleus region. Future research should therefore concentrate on the proposed automatic classification in order to separate between LSIL and HSIL stages.

References

[1] American cancer society, “What is endometrial cancer?,” *Natl. Heal. Counc.*, pp. 1–11, 2015.

[2] M. Meissnitzer and R. Forstner, “MRI of endometrium cancer - How we do it,” *Cancer Imaging*, vol. 16, no. 1, 2016.

[3] F. Amant, M. R. Mirza, M. Koskas, and C. L. Creutzberg, “Cancer of the corpus uteri,” *Int. J. Gynecol. Obstet.*, vol. 143, pp. 37–50, 2018.

[4] G. Trojano, C. Olivieri, R. Tinelli, G. R. Damiani, A. Pellegrino, and E. Cicinelli, “Conservative treatment in early stage endometrial cancer: A review,” *Acta Biomedica*, vol. 90, no. 4. pp. 405–410, 2019.

[5] H. van der Merwe, “Endometrial cancer – What is new?,” *Obstetrics and Gynaecology Forum*, vol. 29, no. 1. pp. 10–15, 2019.

[6] A. Billiau, “What is Endometrial Cancer? Let us Explain it To You,” *Med. Oncol.*, 2016.

[7] P. J. Eifel, “High intermediate risk endometrial cancer. What is it?,” *International Journal of Gynecological Cancer*, vol. 29, no. 7. pp. 1084–1085, 2019.

[8] P. Gupta, N. Gupta, and P. Dey, “Artificial neural network for cytodiagnosis of endometrial carcinoma,” *Anal. Quant. Cytopathol. Histopathol.*, vol. 38, no. 4, pp. 249–254, 2016.

[9] N. Kim, J. S. Chang, Y. B. Kim, and J. S. Kim, “Atlas-based auto-segmentation for postoperative radiotherapy planning in endometrial and cervical cancers,” *Radiat. Oncol.*, vol. 15, no. 1, 2020.

[10] I. P. Constantinou, C. A. Koumourou, M. S. Neofytou, V. Tanos, C. S. Pattichis, and E. C. Kyriakou, “An integrated CAD system facilitating the endometrial cancer diagnosis,” in *Final Program and Abstract Book - 9th International Conference on Information Technology and Applications in Biomedicine, ITAB 2009*, 2009.

[11] H. Sun, X. Zeng, T. Xu, G. Peng, and Y. Ma, “Computer-Aided Diagnosis in Histopathological Images of the Endometrium Using a Convolutional Neural Network and Attention Mechanisms,” *IEEE J. Biomed. Heal. Informatics*, vol. 24, no. 6, pp. 1664–1676, 2020.

[12] N. Otsu, “A Threshold Selection Method from Gray-Level Histograms,” in *IEEE Transactions On Systems, Man, And Cybernetics*, 1979, vol. 20, no. 1, pp. 62–66.

[13] C. Wolf, J. Jolion, R. F. V Rr-, I. De Lyon, and A. A. Einstein, “Extraction and Recognition of Artificial Text in Multimedia Documents,” *Pattern Anal. Appl.*, vol. 6, no. 4, pp. 309–326, 2003.

[14] J. Bernsen, “Dynamic thresholding of grey-level images.,” in *Proceedings of the Eighth International Conference on Pattern Recognition*, 1986, pp. 1251–1255.

[15] M. Feng and Y. Tan, “Contrast adaptive binarization of low quality,” *Electron. Express*, vol. 1, no. 16, pp. 501–506, 2004.

[16] W. A. Mustafa, H. Aziz, W. Khairunizam, Z. Ibrahim, S. Ab, and Z. M.Razlan, “Review of Different Binarization Approaches on Degraded Document Images,” in *IEEE International Conference on Computational Approach in Smart Systems Design and Applications (ICASSDA)*, 2018, pp. 1–8.

[17] Wan Azani Mustafa, Low Zhe Wei, and Khairul Shakir Ab Rahman, “Automated Cell Nuclei Segmentation on Cervical Smear Images Using Structure Analysis,” *J. Biomimetics, Biomater. Biomed. Eng.*, vol. 51, pp. 105–115, 2021.

[18] W. A. Mustafa, I. Jamaluddin, W. Khairunizam, R. Santiagoo, N. S. Othman, and M. N. K. H. Rohani, “Comparison of Detection Method on Malaria Cell Images,” in *IEEE International Conference on Computational Approach in Smart Systems Design and Applications (ICASSDA)*, 2018, pp. 1–6.

Authors Introduction

Dr. 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.

Ms. Nurul Umaira Salim

Nurul Umaira obtained her KPM Matriculation Certificate in Science from Negeri Sembilan Matriculation College in 2017 and 2018 respectively. She is currently an undergraduate student of Bachelor of Biomedical Electronic Engineering with Honour at School of Electronic Engineering, University Malaysia Perlis. His research interests include Image and Signal Processing

Dr. 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.

Ms. Shahrina Ismail

She obtained her B. Sc. Ed. in Mathematics from Universiti Putra Malaysia (UPM) in 2019 and M. Sc. in Pure Mathematics from INSPEM, UPM in 2011. She then traveled to Australia to pursue her Ph. D in Mathematics at University of Queensland, Brisbane. She currently serves as a senior lecturer in the Mathematics Program at Universiti Sains Islam Malaysia, Negeri Sembilan. Her research interests includes elliptic curves, homotopy analysis, diophantine equations and financial mathematics.

Dr. 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.

Temperature Control Using Fuzzy Controller for Variable Speed Vapor Compression Refrigerator System

Siti Qurrata Ain,
Faculty of Electrical Engineering Technology, Unicity Alam Sg. Chuchuh Campus, University Malaysia Perlis padang Besar, 02100, Perlis

M. Saifizi*
Faculty of Electrical Engineering Technology, Pauh Putra Campus, Universiti Malaysia Perlis Arau, 02600, Perlis
S.M.Othman

Faculty of Electrical Engineering Technology, Pauh Putra Campus, Universiti Malaysia Perlis Arau, 02600, Perlis
Azri A. Aziz

Faculty of Electrical Engineering Technology, Pauh Putra Campus, Universiti Malaysia Perlis Arau, 02600, Perlis
Wan Azani Mustafa

Advanced Computing (AdvComp), Centre of Excellence (CoE), Universiti Malaysia Perlis, Pauh Putra Campus, 02600 Arau, Perlis, Malaysia

Wan Khairunizam
Faculty of Electrical Engineering Technology, Pauh Putra Campus, Universiti Malaysia Perlis Arau, 02600, Perlis

Corresponding Author: saifizi@unimap.edu.my

Abstract

Keeping the cold chain vaccine is crucial to a stable immunisation programme; however, faulty processes may occur more frequently than are often thought in developing nations. This paper discusses the quick and accurate control process for designing fuzzy controllers for variable speed vapor compression refrigerator system. The suggested controller is based on the fuzzy logic intended to improve performance while keeping the cooler's constant internal temperature and increasing the refrigerator efficiency. Despite the external changes such as the outside temperature change or the volume change in the refrigerator vaccine, the fuzzy logic controller is utilised to maintain the interior temperature. However, a variable speed compressor (VSC) must be used to control the thermophysical characteristics, which dramatically alter the temperature with a small pressure change. In this case, fuzzy rules of the sort developed by Mamdani are used to build up the system. The programming platforms utilised to implement the model include MATLAB, SIMULINK, and Fuzzy Logic Toolbox (FLT). The efficiency of fuzzy logic controller design membership will be compared to ensure that the refrigerator temperature is more accurate and until it achieves the best performance, maintains a temperature of 5°C, and adapts to its surroundings. From the research done, the membership 2 with load shows the near accurate temperature of 5°C with steady-state error ±1.97°C.

Keywords: Fuzzy controller, membership, variable speed compressor, temperature control, control effort

1 INTRODUCTION

Refrigerators work by converting the liquid refrigerant that circulates through them into a gas phase. This process, known as evaporation, cools the environment surrounding it and resulting in the intended outcome[1]. There are many different types of a refrigerator in the refrigerator industry, such as vapour compression refrigerators,air and gas expansion refrigerators, vapor absorption refrigerators, and others. Refrigerators systems are widespread in the natural gas processing sector, petroleum refining operations, petrochemical industries and chemical processes in the synthetic rubber manufacture, textile, plastic, etc.[2].

Fuzzy logic is a primary control system based on the degrees of input and output depending on the input and changing rates. In other words, the idea of assigning a particular output relies on the probability of the input condition. Fuzzy logic focuses on output definition according to assumptions[3]. It operates based on sets. Each set has specific linguistic characteristics that define the state of the prospective output. Depending on the outcome, each possible input state and state degree of change is part of the set. It operates on if-the-mind. These fuzzy sets are represented graphically using membership functions, and the output may be decided according to the membership level of each function. IF-ELSE logic determines the membership settings[4]. Three parts are essential in the fuzzy controller: the fuzzification process, rule extraction, and defuzzification process. The decomposition of a system's input or output into one or more fuzzy sets is known as fuzzification. The fuzzification technique allows system inputs and outputs to be represented in language terms, allowing simple rules to define a complex system[5].

This research needs to study a controller that can control temperature precision and analyse the membership of fuzzy controllers. To reach the desired temperature of 5°C, the refrigerator's input current must be appropriately regulated. The conventional circuit is updated in this segment with a resistance temperature controller known as an RTD. It also contains a newly updated circuit and a controller based on the Arduino MEGA microcontroller.

2 METHODOLOGY

In this project, a mini-refrigerator has been selected to carry on. The mini refrigerator had been modified with an RTD sensor, Arduino Mega and controller. The modification was made to improve the refrigeration system's temperature output precision. The refrigerator is adjusted until it achieves the best performance, maintains a temperature of 5°C, and adapts to its surroundings.

2.1. Experimental Setup of Refrigerator System

During this research, a minibar-sized Vapour Compression Refrigeration System-type refrigerator with dimensions of 51 cm x 48.9 cm x 48.8 cm, the storage capacity of 60 L, and a weight of 13.5 kg were used. The Hitachi CL 1588- DA variable speed compressors were used. The refrigeration system will function because of the pressure generated by the compressor. When the pressure of the refrigerant R600a changes, the characteristics of the refrigerant change as well. However, for the purpose of control, the system's input is represented by current, and the temperature response of the system represents its output. The refrigerator original control circuit was replaced with a custom build system. As input, the system makes use of a Resistive Temperature Detector (RTD)- sensor, and as a controller, and it makes use of an Arduino control unit that serves as a pass-through to Matlab. The system makes use of an output unit that consists of a single-phase diode rectifier, a three-phase inverter, a small DC-link film capacitor, and a permanent magnet synchronous motor as well as other components to perform.When design a fuzzy controller, the parameter of each output and input can set from the calculated frequency. Maximum frequency of this system is 350Hz and the minimum frequency is 0. When the frequency is high the speed of compressor is increase.

2.2. The Design of Frequency Logic Controller

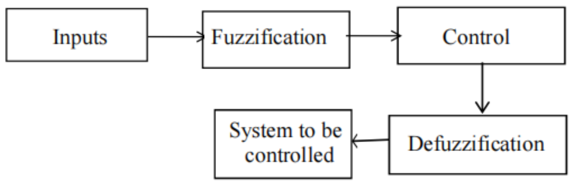


Fig. 1. Block Diagram of the Fuzzy Control System

A fuzzy logic controller was designed by using Matlab software. It is necessary to do three phases to build up a fuzzy logic controller: fuzzification, Inference, and finally defuzzification[6].

2.2.1 Fuzzification

The decomposition of a system's input or output into one or more fuzzy sets is known as fuzzification. It uses the membership contained in the fuzzy knowledge base to convert the crisp input to a linguistic variable. When it comes to systems, fuzzification is the process of identifying their input, output, and membership function[7]. Before designing the membership, the minimum range for temperature is 0°C, and the maximum is 35°C. The temperature needs to maintain at 5°C. For the speed, the minimum range is 0 Hz, and the maximum is 350 Hz. The function of the frequency is to control the speed of the refrigerator.

2.2.2 Inference

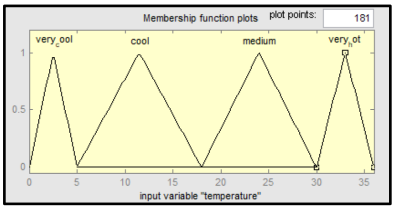
By extracting IF-THEN rules from a decision tree, a rule extraction classifier is created. When extracting a rule from a decision tree, one rule is constructed for each path from the root to the leaf node. Each splitting criterion is logically ANDed to generate a rule antecedent[8][9]. The class prediction is stored in theleaf node, which forms the rule consequent. So, for the fuzzy refrigerator, it had four phases: very cool, cool, medium, and very hot. The temperature needs to maintain at medium temperature. So, if the temperature is hot, then the speed will be faster. And if the temperature is cold, then the speed will be slower to maintain the medium temperature set.

2.2.3 Deffuzification

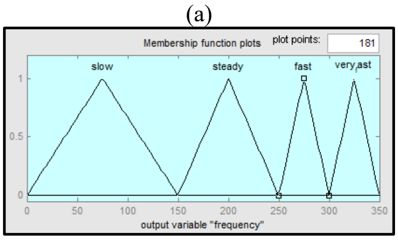
Defuzzification is the method through which a fugitive set with a crisp number is represented. Internal data displays in a fuzzy system are generally fuzzy sets. Fuzzy sets are commonly used as internal data representations in fuzzy systems[10]. On the other hand, the output is usually required to be a precise number that may be utilised to conduct a function. The centre of area approach (COA), also known as the centroid approach, is the most often used defuzzification method. This function returns the crisp value corresponding to the fuzzy set's centre of the area[11][12]. After the range had been confirmed, need to try the suitable membership and the size of membership to get the best output to reach the desired temperature of 5°C. The crisp output is the combination output that will analyse using Matlab.

2.2.4 Input-Output Membership of Fuzzy Controller

The membership function for the input temperature of the fuzzy logic controller is shown in Figure 2 (a). It was imperative to provide four different membership functions for the input temperature: very cool, cool, medium, and very hot. The membership function for the output frequency of the fuzzy logic controller is shown in Figure 2 (b). It was imperative to provide four different membership functions for the input temperature: slow, steady, fast and very fast. Trapezoidal and triangular geometric also had been used in this research with another data for membership. The parameter had been set for each of membership function based on the stability of the refrigerator.



(a)



(b)

Fig. 2 The Relation of Input-Ouput Membership of Fuzzy Controller

2.2.5 Mamdani Law

The Mamdani law is shown in Figure 3 and is used in the decision-making process of the fuzzy logic controller. If the Fuzzy Inference System is of the Mamdani type, then both the input and output variables will include a collection of fuzzy sets[13]. The Mamdani law is a collection of principles that describe the connection between the temperature of the input and the output speed.[14] First rule objectsrepresent fuzzy if-then rules that relate input membership function conditions to corresponding output membership function conditions in a declarative method. The antecedent of a fuzzy rule is the part of the rule that describes the membership function for each of the variables in the rule's input [15].

1. If (temperature is very_cool) then (frequency is slow) (1)
2. If (temperature is cool) then (frequency is steady) (1)
3. If (temperature is medium) then (frequency is fast) (1)
4. If (temperature is very_hot) then (frequency is very_fast) (1)

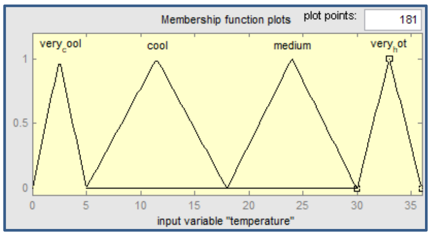
Fig. 3. Mamdani Laws of the Fuzzy Logic

3 RESULT & DISCUSSION

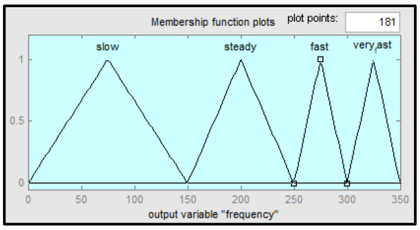
A few memberships had been tested to find the most accurate temperature for the fuzzy controller refrigerator.

3.1. The Performance of Control Effort and Temperature Response for Membership 1

Figure 4 shows that the fuzzy controller of membership 1. All the four membership functions for temperature and speed use a triangular geometric shape. For temperature input, very cool has the parameter of [0 2.5 5], cool has a parameter [5 11.5 18], the medium has a parameter of [18 24 and 30] and lastly, very hot has the parameter of [30 33 36]. For output speed, slow has a parameter [0 75 150], steady has a parameter of [150 200 250], fast has a parameter [250 275 300] and lastly, very fast has a parameter of [300 325 350].



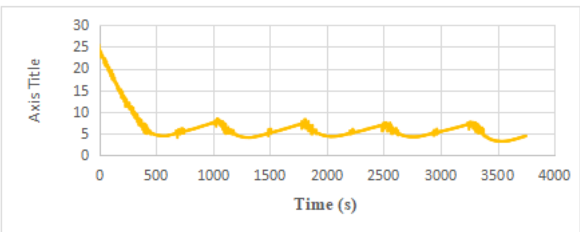
(a) Input variable (Temperature)



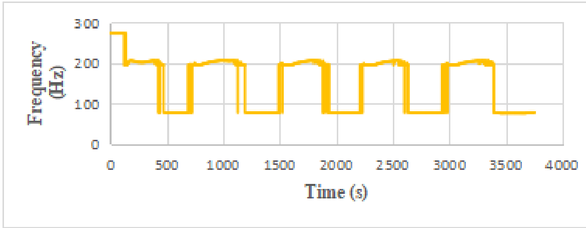
(b) Output variable (Speed)

Fig. 4. Membership 1 of Fuzzy Controller

Figure 5 shows the relationship between the control effort of the controller and its measured temperature for a Fuzzy-Logic Controller without load. The peak overshoot temperature for data 1 (a) is 8.24°C to achieve the desired temperature with a still respectable steady-state error of ±1.30°C. From Figure 5 (b), it can be seen that the control effort of the Fuzzy-Logic controller initially run at 272Hz before throttling down to 194Hz and then alternating between 78Hz and 194Hz. The alternating of the frequency follows the Mamdani law established before, where when the input temperature is medium, the output speed is fast while the temperature is very cool, the output speed will slow.



(a)

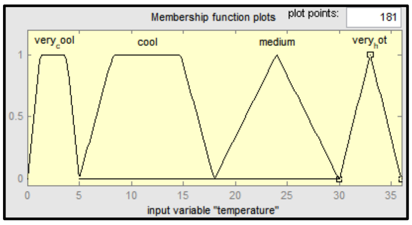


(b)

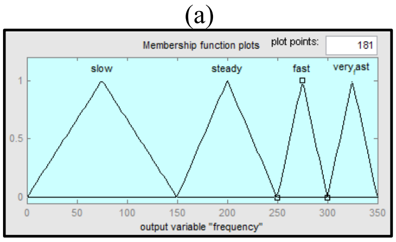
Fig 5. The Fuzzy-Logic controller performance (a) control effort, (b) temperature response for membership 1

3.2. The Performance of Control Effort and Temperature Response for Membership 2

Figure 6 shows the fuzzy controller of membership 2. For the input variable membership, trapezoid and triangular membership were used, and for the output variable speed, all four members used triangular membership. For temperature input, very cool has the parameter of [0 1.25 3.75 5], cool has a parameter [5 8.25 14.8 18], the medium has a parameter of [18 24 and 30] and lastly, very hot has the parameter of [30 33 36]. For output speed, the parameters use the same parameters as previous for membership 1.



(a)

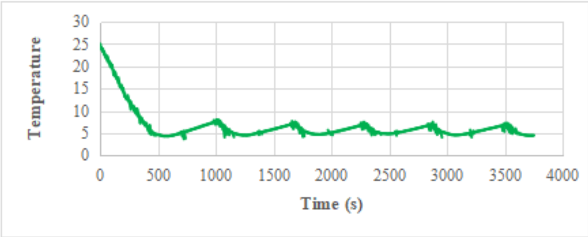


(b)

Fig 6. Membership 2 of Fuzzy Controller

Figure 7 shows the measured temperature for a Fuzzy-Logic Controller. The peak overshoot temperature for data 1 (a) is 7.54°C to achieve the desired temperature with a still respectable steady-state error of ±1.5°C. From Figure 7 (b), it can be seen that the control effort of the Fuzzy-Logic controller initially run at 272Hz before throttling down to 204Hz and then alternating between

77Hz and 194Hz. The compressor will stop entirely at 77Hz then will work again when the temperature gets higher.



(a)

(b)

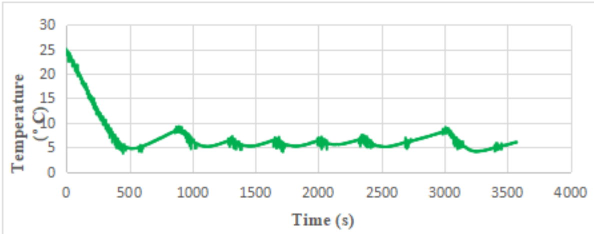
Fig. 7. The Fuzzy-Logic controller performance (a) control effort, (b) temperature response for membership 2

3.3. The Performance of Control Effort and Temperature Response for Membership 2 with load

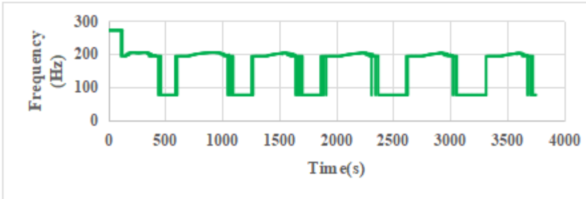
Two fuzzy controllers had been tested to find the most precise temperature control for this refrigerator, with the desired temperature with the acceptable steady-state error of $\underline{+}$1°C. From the fuzzy membership without load, membership 1 is most precise from membership 2 with the average $\underline{+}$1.3°C. The water of 500ml also had been tested with the refrigerator to find out the refrigerator's performance. The result shows that with the load, data 2 is more precise than the fuzzy membership controller of data 1.

Table 1. Steady State Error of Two Membership

Figure 8 shows the relationship between the control effort of the controller and its measured temperature for a Fuzzy-Logic Controller with the load. The peak overshoot temperature for data 2 (a) is 8°C with a still respectable steady-state error of ±1.97°C. After the peak overshoot, the temperature has a steady temperature until 3000sec. Figure 5 (b) shows that the control effort of Fuzzy-Logic controller initially runs at 200Hz before throttling down to200Hz and then alternating between 194Hz and 205Hz. The graph frequency of data 2(b) with load almost constantly working.



(a)



(b)

Fig 8. The Fuzzy-Logic controller performance (a) control effort, (b) temperature response for membership with load.

4. CONCLUSION

In conclusion, the two fuzzy controllers could be designed, studied and compared the performance of each type of controller. The first fuzzy controller designed was able to achieve a steady-state error of ±2.02°C with a load while the second fuzzy controller was able to achieve a steady-state error of ±1.97°C. Even though the first fuzzy controller has lower steady-state error at ±1.30°C compared to the second fuzzy controller with ±1.52°C. The system also tested based on load is to know the performance of the refrigerator with the load. By comparing the temperature response between each fuzzy controller, it can be concluded that the fuzzy controller for membership 2 with load has the fastest settling time and a low steady-state error.

	Membership 1	Membership 2
Without Load	1.300610739	1.518598611
With Load	2.0234474	1.966741478

5. REFERENCES

1. Refrigeration Principles and how a Refrigeration System Works: Berg Chilling Systems. Berg Chilling Systems Inc. (2020, August 25).

2. Buel, Z. (2018, December 27). The four types of refrigeration systems you need to know. Refrigeration School, Inc. (RSI). Retrieved
3. Jaekel, Jens & Mikut, Ralf & Bretthauer, Georg. (2004). Fuzzy Control Systems.
4. Ordoobadi, S. M. (2008). Fuzzy logic and evaluation of advanced technologies. Industrial Management & Data Systems, 108(7), 928–946.(5)
5. T. A., & , N. (2020, April 22). Fuzzy Logic - How Does Fuzzy Logic Work: Architecture and Applications.
6. Omrane, H., Masmoudi, M. S., & Masmoudi, M. (2016). Fuzzy Logic Based Control for Autonomous Mobile Robot Navigation. Computational Intelligence and Neuroscience, 2016, 1–10
7. Czabanski R., Jezewski M., Leski J. (2017) Introduction to Fuzzy Systems. In: Prokopowicz P., Czerniak J., Mikołajewski D., Apiecionek Ł., Ślęzak D. (eds) Theory and Applications of Ordered Fuzzy Numbers. Studies in Fuzziness and Soft Computing, vol 356. Springer, Cham.
8. Kalogirou, S. A. (2009, July 27). Designing and modeling solar energy systems. Solar Energy Engineering.Retrieved October 31, 2021
9. Kayacan, E., & Khanesar, M. A. (2016). Fundamentals of Type-1 Fuzzy Logic Theory. Fuzzy Neural Networks for Real Time Control Applications, 13–24.
10. Sato, M., & Tsukimoto, H. (2001). Rule extraction from neural networks via decision tree induction. IJCNN'01. International Joint Conference on Neural Networks. Proceedings (Cat. No.01CH37222), 3, 1870-1875 vol.3
11. Athif, M., & Daluwatte, C. (2017). Combination of rule based classification and decision trees to identify low quality ECG. 2017 IEEE International Conference on Industrial and Information Systems (ICIIS).
12. Talon, A., & Curt, C. (2017). Selection of appropriate defuzzification methods: Application to the assessment of dam performance. Expert Systems with Applications, 70, 160–174.
13. Guo, Z. X., & Wong, W. K. (2013). Fundamentals of artificial intelligence techniques for apparelmanagement applications. Optimising Decision Making in the Apparel Supply Chain Using Artificial Intelligence (AI), 13–40.
14. Alavi, N. (2013). Quality determination of Mozafati dates using Mamdani fuzzy inference system. Journal of the Saudi Society of Agricultural Sciences, 12(2), 137–142.
15. Buriboev, Kang, Ko, Oh, Abduvaitov, & Jeon. (2019). Application of Fuzzy Logic for Problems of Evaluating States of a Computing System. Applied Sciences, 9(15), 3021.

Authors Introduction

Siti Qurrata Ain

She received her Diploma in Mechatronic from Politeknik Sultanah Bahiyah in 2017. She is currently submitted her Degree of Engineering Electrical with University of Malaysia Perlis, Malaysia. From 2018 to 2020 until now.

Dr Mohd Saifizi Bin Saidon

He received the B. Eng. degree in Electrical & Electronic Engineering from Fukui University, M.Sc and PhD in Mechatronic Engineering from Universiti Malaysia Perlis, 2008, 2013 and 2016, respectively. He is currently a senior lecturer at the Faculty of Electrical Engineering Technology, University Malaysia Perlis. Her research interests are power electronic system, system identification, control design and artificial intelligence.

Dr Siti Marhainis Othman

She received the B.Eng, M.Eng and PhD degree from University of Technology Malaysia in 2005, 2007 and 2019, respectively. She is currently a senior lecturer at the Faculty of Electrical Engineering Technology, University Malaysia Perlis. Her research interests are system identification, nonlinear controller design, organic optimization and machine learning.

Mohd Azri Abd Aziz

He received the B.Eng.(Mechatronics) from International Islamic University Malaysia in 2009 and M.Sc(Mechatronic) from Universiti Malaysia Perlis in 2017. He is currently a teaching engineer at the Faculty of Electrical Engineering Technology, University Malaysia Perlis. His research interests are system identification, variable frequency drive, robotic and automation .

Ts. Dr. 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.

Dr. Khairunizam Wan

He received his B. Eng. degree in Electrical & Electronic Eng. from Yamaguchi University and Ph.D. in Mechatronic Eng. from Kagawa University, in 1999 and 2009 respectively. He is currently 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.

Automatic Dry Waste Classification for Recycling Purpose

Muhammad Nuzul Naim Baharuddin, Hassan Mehmood Khan, Norrima Mokhtar*, Wan Amirul Wan Mahiyiddin
Department of Electrical Engineering, Faculty of Engineering, University of Malaya, Malaysia

Heshalini Rajagopal
Department of Electrical and Electronic Engineering, Manipal International University, Malaysia

Tarmizi Adam
Faculty of Engineering, School of Computing, Universiti Teknologi Malaysia 81310 Johor Bahru, Johor, Malaysia

Jafferi Jamaluddin
UM Power Energy Dedicated Advanced Centre, Universiti Malaya, 50603, Kuala Lumpur Malaysia

E-mail: nuzulnaim140297@gmail.comnorrimamokhtar@um.edu.my, heshalini@gmail.com

Abstract

There has been a serious increment in solid waste in the past decades due to rapid urbanization and industrialization. Therefore, it becomes a big issue and challenges which need to have a great concern, as accumulation of solid waste would result in environmental pollution. Recycling is a method which has been prominent in order to deal with the problems, as it is assumed to be economically and environmentally beneficial. It is important to have a wide number of intelligent waste management system and several methods to overcome this challenge. This paper explores the application of image processing techniques in recyclable variety type of dry waste. An automated vision-based recognition system is modelled on image analysis which involves image acquisition, feature extraction, and classification. In this study, an intelligent waste material classification system is proposed to extract 11 features from each dry waste image. There are 4 classifiers, Quadratic Support Vector Machine, Cubic Support Vector Machine, Fine K-Nearest Neighbor and Weighted K-Nearest Neighbor, were used to classify the waste into different type such as bottle, box, crumble, flat, cup, food container and tin. A Cubic Support Vector Machine (C-SVM) classifier led to promising results with accuracy of training and testing, 83.3% and 81.43%, respectively. The performance of C-SVM classifier is considerably good which provides consistent performance and faster computation time. Further classification process is improved by utilization of Speeded-Up Robust Features (SURF) method with some limitations such as longer response and computation time.

Keywords: Support Vector Machine, Recycling, Feature Extraction, Classification.

1. Introduction

Solid wastes refer here to all non-liquid wastes. In general, this does not include excreta, although sometimes nappies and the faeces of young children may be mixed with solid waste. Solid waste can create significant health problems and a very unpleasant living environment if not disposed of safely and appropriately. If not correctly disposed of, waster may provide breeding sites for insect vectors, pests, snakes and vermin (rats) that can increase the likelihood of disease transmission. It may also pollute water sources and the environment. This is why solid wastes are known as the most critical problems of our time. There is no place in the corner of the earth that currently immune from the municipal solid waste. The statistics show that, since the beginning of last decade, an increasing sharply amount of solid wastes has been generated[1]. A large quantity of solid waste production is correlated with the GDP. High GDP tends to produce large quantity of solid waste. Updated report data are shown by the world bank report that there are about 4 billion tons of waste generated every year globally which urban area is one of the main contributors to the huge numbers and the waste is estimated to be up until 70% by 2025. In the next 25 years, number of wastes accumulated will be rapidly increased in underdeveloped nations due to accelerated pace of urbanization and industrialization [2].

The population growth rate has continued of 2.4% per year since 1994 based on Department of Statistic Malaysia in 2012. The higher the growing number of people with higher consumption rates, the higher the amount of waste generation. There is a correlation between the income rate and urbanization. As the disposable income and living standard increase, the consumption of products and services, as well as the quantity of waste produces rise correspondingly. With an incrementing number of industries in the urban area, thus, solid waste management becomes a critical concern and challenging for municipal authorities worldwide since the drastically amount of waste generated. Solid waste management confronts more complex problems in the developing nations due to limited door-to-door collection, inefficient treatment and inadequate disposal facilities [3]. Majority of the solid waste is comprised of waste that mainly found in public which consists of paper, plastics, and glass waste material.

The main method to manage waste is landfilling. In U.S.A, 80% of the trash is managed by landfilling which present serious health and ecological problems. This is why landfilling is inefficient since its high-cost operation and most important, polluting the environment. For instance, people who stay around at landfill site can affect their own health. Another solution to manage waste by reducing the volume of the waste is burning waste in Incinerators. However, this method more driven to negative effects, caused severe health problems such as cancer due to human exposure to polycyclic aromatic hydrocarbons (PAHs), which is hazardous and might spread into the air during the air pollution. This method costs a lot of money to be invested in order to build, operate and maintain the machine. Today, the most effective in solving waste disposal problems is recycling and reusing. Based on recent research data, the amount summed up is about 150 million tons. Solid waste is often rich source of various recyclable materials. According to Gundupalli et al. (2017), these recyclable materials can be recovered, become useful and reduce the negative impact on the environment. Waste sorting practices on the other hand, is a beginning step in solid waste management for the recycling of materials. This waste sorting technique is implemented in order for the separation of waste into their different categorization component which can be recycled using different techniques. Hence, recycling is an important tool to assure that our environment is protected as well as human's health [2].

Many studies have been conducted and there are a lot of research papers published in the area of waste sorting and classification using different methods. The main focus in this study is automated sorting of dry wastes such as plastic, bottle, paper and tin cans by means of classification and vision system. Thus, we study literature by the types of materials classified and techniques used for sorting multi-material classification. This contributes to the motivation of this work which are to improve the lacking of prominent, outstanding and discriminative features for excellent classification accuracy.

2. Methods

2.1 Vision-based recognition system

The proposed dry waste classification system by using vision inspection method consists of several module: image acquisition, image processing, feature extraction, classification and finally determined the decision as shown in Fig. 1.

2.2 Image Acquisition

Image acquisition always be the first step when involving image processing. Fig. 2 shows the inspection zone of the waste sorting system which have been covered with housing made up of box where the images were taken for classification later. RGB images are manually taken using a web camera (Logitech QuickCam V-UAP41 USB) within the inspection zone. The camera was attached on top-middle surface of the box after the box is upside down. The setup distance between the camera and the test sample is secured at 25 cm. All the properties setting such as the brightness, contrast and saturation are adjusted based on their respective scales. For the illumination technique, homogenous lighting is adopted for this experiment to obtain a set of geometric properties. This can be done by having the size, shape, orientation and position of the waste samples.

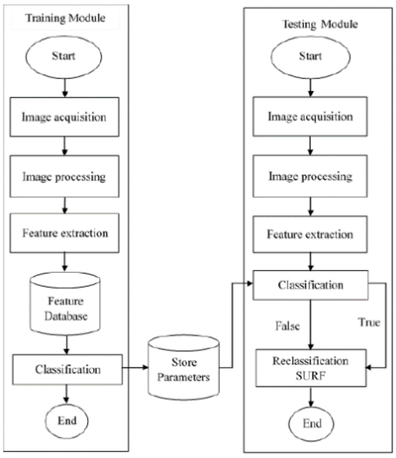


Fig. 1. Conceptual framework for training and testing stage proposed by the system.

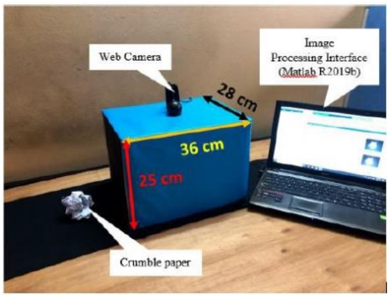


Fig. 2. A photographic image of a setup for vision-based recognition system.

In this experiment, 210 samples are used for training purpose while 70 samples are collected from different place of trash bins such as homes, offices, shops and markets. The test samples consist of 7 different type of class which are 10 test samples of crumble garbage, 10 test samples of flat garbage, 10 test samples of tin cans, 10 test samples of bottles, 10 test samples of food container, 10 test samples of cup and 10 test samples of box. For example, Fig. 3 shows a set of test sample that were randomly picked from each of the class. All the test samples have different looks; shape, size, diameter and orientation. The resolution of each captured photograph is 640×480 pixels.

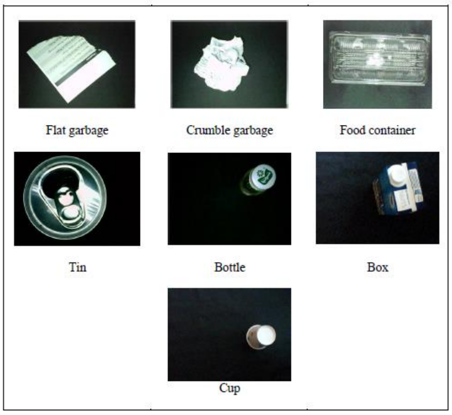
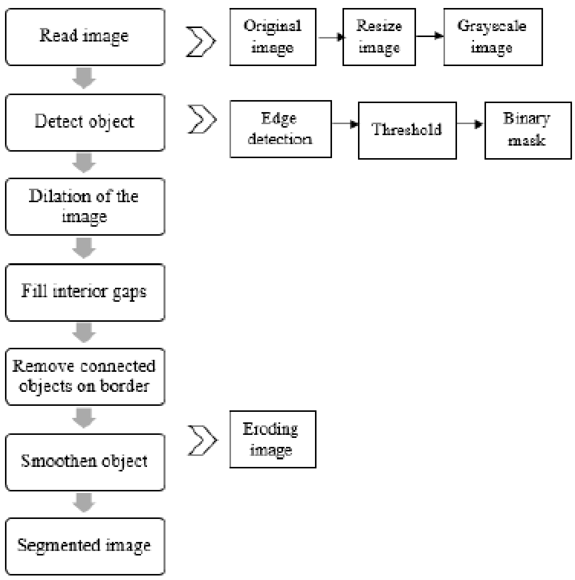


Fig.3. Representative image of dry waste samples.

2.3 Image Processing

In this section, image pre-processing technique such as segmentation or thresholding is discussed. The basic idea of image segmentation is to extract the area of effect of the test sample. Hence, object detection is a starter pack module in this part. There are many techniques developed to find an object from an image, such as object detection algorithm to extract objects in 2-D intensity images using boasted cascade [4], histogram of gradient [5], shift invariant feature transform [6], its background. Segmentation flow of the processing sample image are illustrated in Fig. 4. From Fig. 4, an original image with size 640×480 pixels, is read into the system by using MATLAB R2019b. This RGB sample image is resized to 320×240 pixels by factor of 0.5. This true colour RGB image is converted to greyscale format, where an integer value between 0 to 255, in order to reduce the complexity since RGB image is 3 dimensions and grayscale image is of only 2 dimensions. In order to detect the entire area of the sample image, the background information of the images was eliminated using edge detection and morphological operations. The difference or change in contrast between the object's image and background can be detected by operators. This is done by performing a calculation of the gradient of an image. In this experimental study, a fudge factor value is set to 0.9. The edge detector used in this experimental study is Sobel filter, which is suitable for this condition as Sobel operator more sensitive to diagonal edge than horizontal and vertical edges. Edge detection was performed so that the resulting outline of the images are very precisely. Sobel operator also is used to calculate the threshold values, and edge detection to get the binary mask. The outline images were dilated using linear structure elements to produce more accurate gradient mask. The holes in the dilated images were filled up using morphological operations to fill gaps in images at unfilled area. The cell of interest has been successfully segmented, but it is not the only object that has been found. Any objects that are connected to the border of the image can be removed. Basically, the edge detection and morphological operations are done to perform the segmentation of the samples from the background. Therefore, the area of crumble garbage is extracted by using the mentioned morphological operators. Segmentation flow process for each type of class samples is shown in Fig. 5.



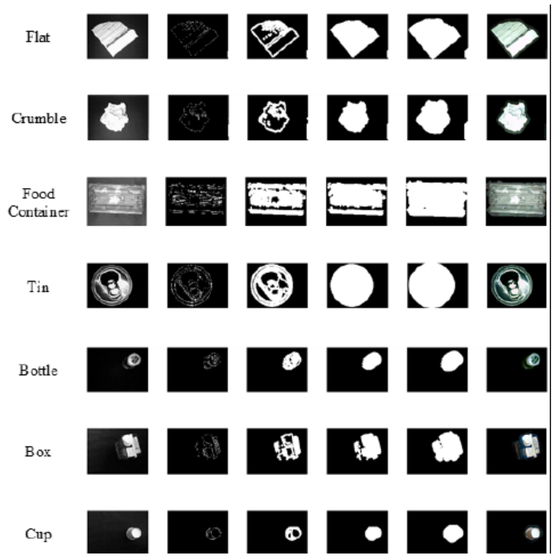


Fig. 4. Flowchart of the sample image in image processing

Fig. 5. The segmentation steps of processing dry waste sample images

2.4 Feature Extraction

Features are used to build classifier model in order to determine the material type and location or coordinates. The feature extraction from the test samples was done using image processing. The operational flow on the test samples for image segmentation is shown in Fig. 6. There are 2 different datasets proposed in this experimental study; the first dataset was extracted from white pixel plot of the sample and the second dataset was extracted from segmented grey image of the sample. Two datasets for training and testing are created by combining statistical and non-statistical features. The dataset is extracted from segmented grey image and white column matrix. From segmented grey image, features such as grey level co-occurrence matrix (GLCM), ratio of grey level, entropy and standard deviation are extracted. This feature dataset is also extracted from the plot of binary mask images of the sample. The white pixel plot of each sample images has different shape which made it possible to extract features for classification. The statistical data applied from the white pixel plot is quantile. The GLCM depicts second order statistical analysis of an image by analyzing how often the pairs of pixels which consist of specific values and spatial relationship take place in an image. The probability, $p(m,n)$ is computed using Eq. (1) [7]:

$$p(m,n) = \{C(m,n)|(d,\theta)\} \quad (1)$$

where d is the inter-pixels displacement distance, θ denotes orientation and $C(m,n)$ denotes the frequency of gray level occurrence in MSCN of the image. Four statistical textures such as contrast, correlation, energy, and homogeneity were extracted from the GLCM matrix.

Contrast calculates the local variations in the gray-level co-occurrence matrix and is defined as Eq. (2) [7]:

$$\text{Contrast} = \sum_{m,n} |m-n|^2 p(m,n) \quad (2)$$

Correlation computes the joint probability occurrence of the specified pixel pairs and is defined as Eq. (3) [7]:

$$\text{Correlation} = \sum_{m,n} \frac{(m-\mu m)(n-\mu n)p(m,n)}{\sigma_m \sigma_n} \quad (3)$$

Energy calculates the sum of squared components in the GLCM. It is also known as uniformity or the angular second moment. The energy parameter is computed as Eq. (4) [7]:

$$\text{Energy} = \sum_{m,n} p(m,n)^2 \quad (4)$$

Homogeneity calculates the closeness of the distribution of elements in the GLCM to the GLCM diagonal and is computed as Eq. (5) [7]:

$$\text{Homogeneity} = \sum_{m,n} \frac{p(m,n)}{1+|m-n|} \quad (5)$$

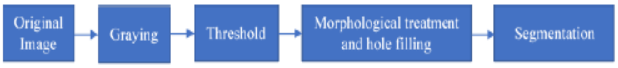


Fig. 6. Summary of image processing used in this experimental study

Grey level is fundamental in study of image processing. The grey level or grey value indicates the brightness of a pixel. The maximum grey value depends on the depth of an image. For example, 8-bit-deep image contain levels up to 255, which they can take any value in the range. However, binary image can only take either value 0 or 255. Table 1 shows the summary of grey level. The program has been setup to calculate ratio of grey level (L) and ratio of grey level (H). Here, ratio of grey level (H) is denote as 40 < x ≤ 110, and ratio of grey level (L) is denote as 181 ≤ x ≤ 255.

Grey level	Colour
0	Black
0 < x < 255	Grey
255	White

Table 1. Grey level and its respective color

In image processing, entropy is a statistical measure of randomness that can be used to characterize the texture of the input image [8]. The higher the value of entropy will result as the more detailed information of an image. Entropy is a measure image information content, which is interpreted as the average uncertainty of information source. A vector with relatively "low" entropy is a vector with relatively low information content, such as it might be [0 1 0 1 1 1 0]. A vector with relatively "high" entropy is a vector with relatively high information content such as it might be [0 242 124 222 149 13]. Hence, it is very important to have higher entropy in order to have precise segmented image after image post-processing method so that it can classify accordingly to its own type of groups. Entropy algorithm is shown in Eq. (6):

$$\mu_m = \sum_{m,n=0}^{i-1} p_{m,n}\left(-\ln\left(p_{m,n}\right)\right) \quad (6)$$

Standard deviation is a most common method used to calculate variability or diversity in statistic. In image processing, it shows how much discrepancy, or "dispersion" exists from the mean value. Total standard deviation used as a more accurate expression of the statistical distribution of each class [9].

Maximum quantile finds the quantities between the data values using linear interpolation which using linear. Initially, quantile assigns the sorted values in x to the (0.5/n), …., ([n-0.5]/n) quantiles [10].

2.5 Classification

2.5.1 Support Vector Machine (SVM)

The classification process SVM in the experiment is carried out by using Classifier Learner Application in MATLAB [11]. There are about six classification models under category SVM, which are linear SVM, quadratic SVM, cubic SVM, fine gaussian SVM, medium gaussian SVM and coarse gaussian SVM. Experiment using the 5-fold cross-validation is applied to evaluate the prediction accuracy of the model, which optimum number of k-fold applied is 5. If the number of k-fold increase, it will result in lower accuracy. In contrast, the accuracy will increase but it does not protect from the overfitting data. Cost matrix applied in this experiment are using default setting for misclassification costs. Other parameters used for the SVM is shown in Table 2.

Table 2. Parameters setting for SVM using classifier learner application

Type of SVM	Parameters			
	Kernel scale	Box constraint	Multiclass method	Standardized data
Linear	Automatic	1	One-vs-One	True
Quadratic	Automatic	1	One-vs-One	True
Cubic	Automatic	1	One-vs-One	True
Fine Gaussian	0.79	1	One-vs-One	True
Medium Gaussian	3.2	1	One-vs-One	True
Coarse Gaussian	13	1	One-vs-One	True

2.5.2 K-Nearest Neighbor

The training examples are vectors in a multidimensional feature space, each with a class label. The training phase of the algorithm consists only of storing the feature vectors and class labels of the training samples. k is a user-defined constant at the classification stage, and a non-labelled vector (a query or test point) is classified by allocating the tag that is the one most common in the k training datasets closest to that query point. The Euclidian distance is a popular distance metric for sustained variables, for which a different metric may be used, such as a differentiation metric (or Hamming distance) for discrete variables such as text classification. In the context of micro-array data on gene expression, k-NN for example was used as metric for coefficients such as Pearson and Spearman. Often, if a distance metric is learned with special algorithms, such as Large Margin Nearest Neighbor and Neighborhood components, the grading accuracy of k-NNN can significantly be enhanced. When the class allocation is skewed, there is a drawback in the fundamental "majority voting" classification. That is, examples of a more regular class tend to dominate the new example prediction because they are common among k neighbors, because they are numerous. One way to overcome this problem is by assessing the distance between the test point and each of its closest neighbors. Each of the nearest k points is multiplied by a 35 weight proportional to the opposite distance of that point to the test point by a class (or value, in regression problems). The parameters used for the KNN is shown in Table 3. Cost matrix applied in this experiment are using default setting for misclassification costs.

Table 3. Parameters setting for KNN using classifier learner application

Type of KNN	Parameters			
	Number of neighbor	Distance metric	Distance weight	Standardized data
Fine	1	Euclidean	Equal	True
Medium	10	Euclidean	Equal	True
Coarse	100	Euclidean	Equal	True
Cosine	10	Cosine	Equal	True
Cubic	10	Cubic	Equal	True
Weighted	13	Euclidean	Squared Inverse	True

3. Results and Discussions

3.1 Preparing training, validation and testing sets

In this experiment, seven type of dry waste are used, which are bottle, box, crumble garbage, flat garbage, cup, food container and tin cans. For each of the class, 30 images of the samples were collected. The experiments were carried out in two stages: training phase and testing phase. Table 4 shows the dataset of dry waste images.

Table 4: Dataset of dry waste images.

Type of dry wastes	Number of training samples	Number of testing samples
Bottle	30	10
Box	30	10
Crumble garbage	30	10
Flat garbage	30	10
Cup	30	10
Food container	30	10
Tin can	30	10
Total	**210**	**70**

3.2 Training Results

Table 5 shows the training results obtained by 4 classifiers used in this experiment. Table 6 shows the average classification accuracy for each classifier. From the experiment that has been conducted, all of the classifier were trained as the predictive model for validation purpose.

Table 5. Training classification accuracy results

Classifier	Number of training									
	1	2	3	4	5	6	7	8	9	10
Quadratic SVM	80	79	76.7	77.1	79	75.7	79	73.3	78.6	75.7
Cubic SVM	83.3	81.9	81	80	82.4	79.5	80	78.1	81	80.5
Fine KNN	81.4	83.8	82.4	81.9	84.8	80.5	83.3	81	82.4	81.9
Weighted KNN	81.4	83.8	82.4	81.9	84.8	80.5	83.3	81	82.4	81.9

Table 6. Average training accuracy

Classifier	Average training accuracy (%)
Quadratic SVM	77.41
Cubic SVM	80.77
Fine KNN	82.34
Weighted KNN	79.62

3.2 Testing Results

Table 7 shows the testing classification accuracy for each classifier. Noted that "Models" column in Table 7 refers to the best classification models which are been saved up as predictive models used for testing experiment.

Table 7. Training classification accuracy result

Type of classifier	Models (training session)	Number of testing sample predicted correctly	Classification accuracy (%)
Quadratic SVM	1	52	74.29
Cubic SVM	1	57	81.43
Fine KNN	5	45	64.29
Weighted KNN	9	44	62.86

From Table 7, it can be seen that Cubic SVM has the highest classification accuracy for testing phase with 81.43%, even though Fine KNN have the highest training accuracy with 82.34%, but the results for classification accuracy during testing phase for Fine KNN is much lower.

4 Conclusion

In this research study, types of dry waste recognition for recycling purpose by using image processing based on some features extraction and image pre-processing approaches is submitted. 7 types of dry waste (bottle, box, crumble garbage, flat garbage, tin can, food container and cup) have being success to recognize. Image pre-processing approaches such as edge detection, image dilating, image filling and image smoothing were able to be applied in order, to differentiate all those types of dry wastes into 7 different classes.

White pixel plot also played important role in separated all those data in the beginning so that a clear picture of each class can be differentiate earlier. Then, other approaches help in recognizing each type of dry waste based on load sample image.

References

1. Guerrero L A, Maas G, Hogland W. Solid Waste Management Challenges For Cities In Developing Countries[J]. Waste Management, 2013, 33(1): 220–232. DOI:10.1016/J.Wasman.2012.09.008.
2. Adedeji O, Wang Z. Intelligent Waste Classification System Using Deep Learning Convolutional Neural Network[J/Ol]. Procedia Manufacturing, 2019, 35: 607–612. Doi:10.1016/J.Promfg.2019.05.086.
3. Gundupalli S P, Hait S, Thakur A. A Review On Automated Sorting Of Source-Separated Municipal Solid Waste For Recycling[J]. Waste Management, 2017,60:56–74. DOI:10.1016/J.Wasman.2016.09.015.
4. Viola P, Jones M. Rapid Object Detection Using A Boosted Cascade Of Simple Features[C]//Proceedings Of The 2001 Ieee Computer Society Conference On Computer Vision And Pattern Recognition. CVPR 2001. Ieee Comput. Soc, 2001: I-511-I–518. DOI:10.1109/Cvpr.2001.990517.
5. Dalal N, Triggs B. Histograms Of Oriented Gradients For Human Detection[C]//2005 Ieee Computer Society Conference On Computer Vision And Pattern Recognition (CVPR'05). Ieee, 2005: 886–893. DOI:10.1109/Cvpr.2005.177.
6. Lowe D G. Distinctive Image Features From Scale-Invariant Keypoints[J]. International Journal Of Computer Vision, 2004, 60(2): 91–110. DOI:10.1023/B:Visi.0000029664.99615.94.
7. Abd Latif, Mh, Md. Yusof, H, Sidek, S.N, Rusli N. Implementation Of GLCM Features In Thermal Imaging Affective State Detection[C]//IEEE International Symposium On Robotics And Intelligent Sensors. . DOI:10.1016/J.Procs.2015.12.298.
8. Mente, Rajivkumar, B. V. Dhandra And G M. Image Recognition Using Texture And Color[J]. International Journal Of Computer Applications, 2014: 33–35.
9. Mario Mastriani A E G. Enhanced Directional Smoothing Algorithm For Edge-Preserving Smoothing Of Synthetic-Aperture Radar Images[J]. Journal Of Measurement Science Review, 2004, 4(3): 1–11.
10. Langford E. Quartiles In Elementary Statistics[J]. Journal Of Statistics Education, 2006, 14(3). DOI:10.1080/10691898.2006.11910589.
11. Chang, Chih-Chung And Lin C-J. Libsvm: A Library For Support Vector Machines[J]. Acm Transactions On Intelligent Systems And Technology, 2011, 2(3): 27:1--27:27.

Authors Introduction

Mr. Muhammad Nuzul Naim Baharuddin

He received his B.E (Electrical) in 2020 from University of Malaya, Malaysia. Currently, he is attached to Department of Statistics Malaysia. His research interest includes image processing, artificial intelligence and machine learning.

Mr. Hassan Mehmood Khan

He received his B.E in Electronics from Dawood University of Engineering and Technology, Pakistan in 2009. He is currently submitted his Master of Engineering Science with University of Malaya, Malaysia. From 2015 to 2020, he has been as Head of Department(Mechatronics) at MIT Academy Sdn. Bhd. His research interests include automation, image processing, machine learning, and artificial intelligence.

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. Heshalini Rajagopal

She received her PhD from University of Malaya, Malaysia in 2021. She received the B.E (Electrical) in 2013 and MEngSc in 2016 from University of Malaya, Malaysia. Currently, she is a lecturer in Manipal International University, Nilai, Malaysia. Her research interest include image processing, artificial intelligence and machine learning.

Dr. Tarmizi Adam

He received a B. Eng degree in electronics, computer and information engineering in 2011, an MSC degree in 2013 and PhD in electrical engineering in 2019 from International Islamic University Malaysia, Universiti Teknologi Malaysia and Universiti Malaya respectively. His research interests include image processing, signal processing and mathematical optimization.

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. Jafferi Jamaludin

He received a B. Eng degree from Universiti Tenaga Nasional Putrajaya, an M. Eng. Sc. and PhD from Universiti Malaya, Kuala Lumpur, Malaysia. He received Hitachi Global Foundation for postdoctoral research at Keio University, Japan. His research interests include image processing, signal processing, control system and energy management.

A Low-Cost Smart Parcel Box System with Enhanced Security

Ahmad Luqmanulhakim bin Mohd Rusli, Wan Norsyafizan W.Muhamad, Suzi Seroja Sarnin, Meor Mohd Azreen Meor Hamzah

School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, Malaysia

E-mail:syafizan@uitm.edu.my

Abstract

A global pandemic of covid-19 has hastened the growth of online shopping or E-commerce. Nowadays, E-commerce transactions provide various products from luxury goods and services to everyday necessities. While the popularity of online shopping has grown fast, there are several common issues that shoppers experience. Common problems that customers experienced are failed delivery, missing parcels and even the criminal case. Malaysia has introduced Pos Laju Ezi Box Kiosk, but still contemporary Kiosk type approach for the delivery system has some problems such as high initial installation cost, the expense of management system and especially security weaknesses in the wireless communication. To overcome the aforementioned problems, this paper proposes a low-cost smart parcel box system that will be installed at individual homes with enhanced security. This system used Arduino Mega 2560 to control all the processes of the developed system. The system will be initiated when couriers send the parcel's tracking number as a message to the user via applications to get the password. For security purpose, password will be provided once the courier's message is the same as the message specified by the user. Couriers then can enter the password provided and insert the parcel into the smart parcel box. The proposed low-cost smart parcel box system ensures that parcels are delivered safely and securely to the customer's door.

Keywords: smart parcel, Arduino mega 2560, infrared sensor (IR), Internet of Things.

1 INTRODUCTION

The coronavirus outbreak has accelerated the shift towards a more digital trends and initiated changes in E-commerce or online shopping behavior. Based on survey managed by Rakuten Insight [1], within Malaysia, almost 64% of female's customer expand their online shopping due to reason minimizing chances to meet people as well as adapting social distancing. In addition to that, purchasing anything on websites such as Shopee, Amazon, Lelong, and others is more convenient for the customer. Among all those benefits, there are also several major problems that occurring nowadays [2]. Sometime the parcel is being lost or even damaged due to not properly put in place.

Nowadays, most of parcel/letter box are now in the regular and manual forms. Therefore, it will be less secure if someone leave the parcel for a long period in the box. There are some difficulties that might be faced, mostly in condo, office and apartment buildings where residents have limited time to check mail or parcel due to the central location of mailboxes. In addition to that, Although the parcel is delivered successfully by the couriers, the condition of the parcel is still unsafe even the parcel is inside the mailbox. This is because the mailbox can still be opened by thief even it locked with key or padlock. So, the parcel could be potential to get stolen if the parcel were left inside the mailbox.

Thus, the parcel that did not delivered will be returned back to the post office and will be categorized as 'failed on delivery'. There is some case which is frequently happened by clients where some items went missing when it returned back to the post office. As a result, customer will incur losses due to the missing items. Negligence caused by the courier company will cause a lot of bad impact towards the courier and postal world. Moreover, the recipient will give bad comment and rate to the seller because of the missing item.

Problems such as failed deliveries, loss of packages, and late delivery processes are not only faced by people in Malaysia but also across the country. As stated in the article [3], in Japan, the strict business regulations were barriers for operators' innovative ideas and network expansion effort in the process of development in the parcel delivery market. On the other hand, some operators caused problems such as cargo accidents, delivery delays, and leaving items to neighborhood without permission when receiver was not at home [3]. This means that some operators could not deliver the parcel by hand to hand to the customer with the secure packaging. When this happened, parcel will be delayed and it would exceed the time limit of the delivery process. This will get worse if the parcel missing.

In another article written by Castle, J. [4] the paper state that, “a quarter of parcel-receiving in Australian experienced a failure to be delivered. The author mentioned that the delivery delays problem has been around since 2008, when the Postal Industry Ombudsman (PIO) says it has seen a significant increase in the number of complaints relating to delivery and notification cards. That said, it has also seen an increase in complaints overall. To overcome problems encountered by couriers and online shoppers, in Malaysia, few solutions were proposed such as Pos Laju Ezibox Parcel Lockers, Pos Laju Ezidrop Self-Service Machines and Pos Laju Kiosks. It enables customers to collect or post their parcels any time at their convenience[3]. Even when customers were not at home during delivery process, they just can get their parcel through the nearest Pos Laju Ezibox Parcel Lockers or Kiosk to claim their parcel. However, during this pandemic of Covid-19, people are trying to minimize time going out of home and avoiding crowded places. Thus, it would be an advantage if the parcel box is placed at customer’s home.

However, based on author knowledge, there is none of the proposed solutions installed at home instead of at specific locations. Devi Pujari, A. et al[5] developed a system which is smart letter box system using obstacle sensor to notify users via Android applications. The android application is used to get the warning out of the pack. A web association SIM card is also included in the equipment unit. The unit is given reinforcement of the battery to avoid the power disappointment issue. This will store the information using the MySQL database. The information that stores the letter check will be dropped in the letter box which is the place where the letter was received. This framework is IoT's best methodology.

Nowadays people rarely available at home during the day especially employee because they went out to work. When the parcel out for delivery (normally on office hour), they were not even at home. So, their parcel is simply left in the lobby for pick up. This case often happening in the apartment residential because their mailboxes located at the ground floor of each block. All residents’ mailbox was located there and there was no security in that area. Thus, it will be creating concerns about security, high risk of losing parcel and re-delivery attempts for those who are not at home during the item out for delivery

Based on the aforementioned problem, Smart Parcel Box System is invented to reduce the problems encountered by the courier and customers. This project is also designed in providing reliable, safe, convenience and energy efficient intelligent parcel system. This Smart Parcel Box System will be operating by using Arduino Mega 2560, Keypad, Infrared sensor (IR sensor), Liquid Crystal Display (LCD), Global System for Mobile (GSM), Weight sensor and Application (App). This Smart Parcel Box System provides user with three main features which are keypad for security and smart notification system (GSM module and Application).

The parcel box is placed on specific place (in front of the house gate) at home. The system is responsible to allow the person in charge for delivery process to deliver the parcel into the parcel box. Smart Parcel Box System is installed with keypad that is responsible to unlock the door of the parcel box after the password entered by couriers is correct and allow couriers to gain access to insert the parcel into the parcel box. Password can only be given once the couriers sent the tracking number of the parcel to the user. Infrared sensors are also used in this project to detect incoming parcel and to detect any object or heat at the parcel box’s door in order to close the door after the parcel has been placed. Following that, GSM module will send the notification to the owner of Smart Parcel Box System through short messaging system (SMS) to notify

2 METHODS

2.1. Components

Generally, this system will be initialized when voltage source is supplied from the direct current (DC) to the system. At that time, all systems start up and ready to operate for the next step. The system will identify the password entered by couriers or the person in charge for delivery whether is the same as it programmed or not. The password can be obtained by sending the tracking number of the parcel as message to the recipient via apps. Then, it will automatically receive the password if the tracking number is the same as the recipient specified in apps. If one of the tracking number characters missing or does not follow the message characteristics specified by the recipient, the password will not be given and the person in charge for delivery need to send the tracking number again. Next, the door of the parcel box will be open indicating that the access was granted due to correct password and allowing the person in charge for delivery to insert the parcel into the parcel box. After that, the door will automatically be closed as the parcel is inserted. Finally, the GSM module will send a message to the recipient to give information that the parcel had been sent to their home.

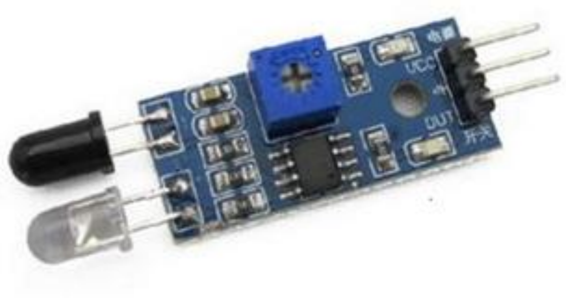


Fig. 1. Infrared Sensor

An infrared sensor is an electronic device that is used to sense certain physiognomies of its surroundings. Infrared sensor is capable of measuring the heat being emitted by an object and detecting Infrared. Infrared waves usually have wavelengths of 0.75 and 1000µm somewhere [6]. In this project, it will act as the sensor to detect the hand of the person in charge to deliver the parcel. It will send information to the Arduino Mega 2560 and the system will execute the program.

Fig. 2. Arduino Mega 2560

Arduino Mega 2560 is an open source microcontroller which can be effectively customized, eradicated and reconstructed at any moment of time [7]. The Arduino Mega 2560 is a microcontroller board based on the ATmega2560 (datasheet). It has 54 digital input/output pins (of which 14 can be used as PWM outputs), 16 analog inputs, and more. It contains everything needed to support the microcontroller; simply connect it to a computer with a USB cable or power it with a AC-to-DC adapter or battery to get started. For this project, it will function to execute the program inside it such as opening and closing of the door, and the infrared sensor.

Fig. 3. Keypad

A 4×4 matrix keypad is used to give commands and the password to the MCU. It consists of 16 keys (S2-S17 arranged in the form of a square matrix of four rows and four columns). Each key in the matrix is labeled according to the operation assigned to it. Rows 1 through 4 are connected to pins RB3, RB2, RB1 and RB0 of Port B of the MCU, respectively. Columns 1 through 4 are connected to pins RB4 through RB7 of Port B respectively [8]. The function of keypad in this project is to unlock the door of the Smart Parcel Box.

Fig. 4. GSM Module

GSM module is applied for the communication between a computer and a GSM-GPRS system. This device can be performed by connecting the GSM module to the AVR and connecting a SIM card to a server's GPRS cover [9]. GSM module will send SMS as notification to the user of the Smart Parcel Box, as the parcel delivery process is successful.

Fig. 5. LCD

Liquid Crystal Display (LCD) is an electronic device that is used to display instruction that programmed in the form of output of software designing. It is made by liquid crystal. enveloped in a glass cover [10] LCD will be used to show every program executed and also instruction on the screen. This LCD usage might give a lot of help to the person in charge for delivery process.

2.2. System Architecture

Figure 6 below shows the System Architecture of Smart Parcel Box System (SPBS) Using Arduino. Basically, the system will be initialized when the power supply from the house or from other sources such as power bank to the system. The system require password from the couriers first before the system gives permission unlock the door of Smart Parcel Box. Password can be provided by sending the tracking number of the parcel as message to the user via apps. Next, the door will be open when the information on the message or tracking number is tallied with the message which is specified by user in the apps. Then, the couriers will get the password and can unlock the door when the message is the same. When the door is open, the parcel box will be insert into the parcel box and at the same time, the infrared sensor will detect the parcel. Once there is nothing at the door, LCD will display information to close the door and the door will be closed after a few seconds. After that, GSM module will send the notification in SMS form to notify the user on the parcel delivery status. LCD will display all information given by the Arduino Mega 2560 such as wrong password, door open or door close.

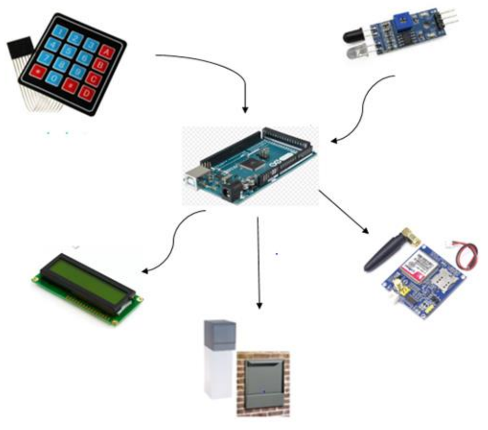


Fig. 6. System Architecture of Smart Parcel Box

2.3 Flowchart

Figure 7 shows the flowchart of the whole process for created Smart Parcel Box System. The system starts with courier send the tracking number of the parcel as a message to the recipient via Whatsapp Applications. Another application called AutoResponder WA is used to generate message whether the message is exactly the same as the recipient specified in the application. This application requires the recipient to fill out the parcel tracking number purchased before the delivery process. This is to ensure the courier can access the parcel box and to facilitate courier work. Then, courier need to send the message specifically to the recipient otherwise the courier could not receive the message from the recipient and need to send the tracking number again. If the message or tracking number is the same as the recipient's message, automatically the message is send to the courier. Means that the courier will get the password to unlock the door of the Smart Parcel Box. Following that, when the courier enters the password given, the Arduino will generate coding whether is the same as it programmed or not. If the password is correct, couriers can insert the parcel into the parcel box. The door will close if the IR sensor did not detect anything at the door. Once the door is close, GSM module will be instructed to send the message as notification to the recipient knowing that the parcel is already delivered to his/her home.

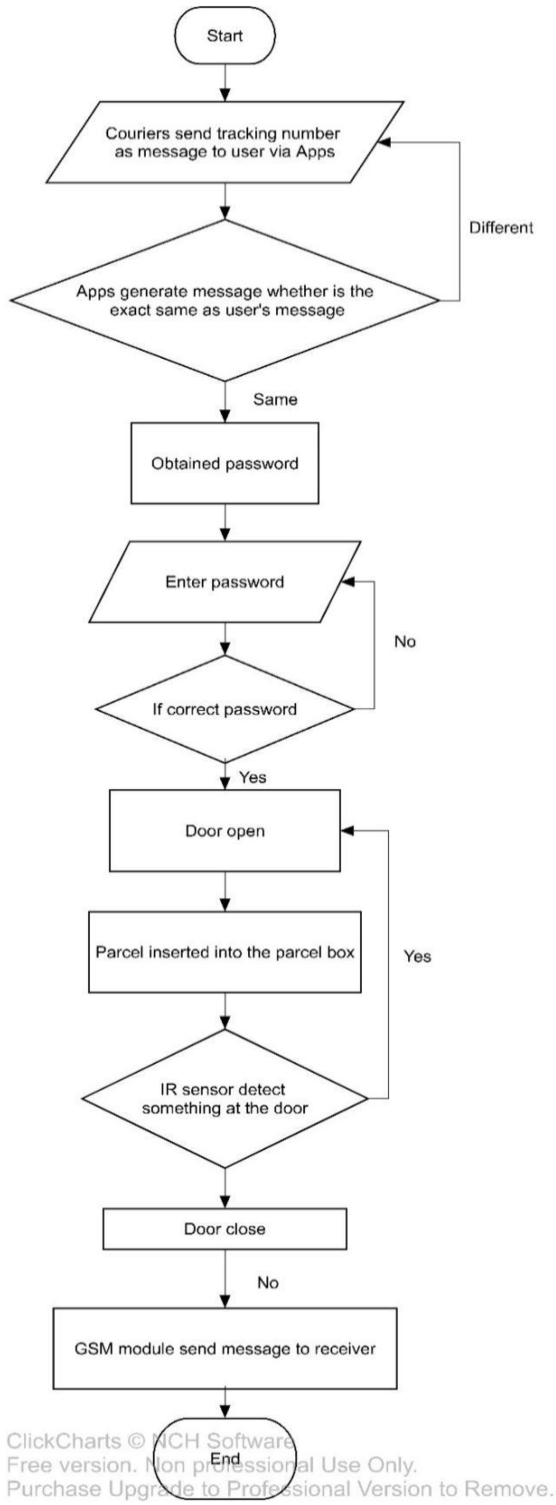


Fig. 7. Flowchart of the proposed Smart Parcel Delivery System

3 RESULTS AND DISCUSSIONS

Arduino Mega 2560 is the brain for Smart Parcel Box System. If the Arduino is damaged or the system failed, all components might not work properly and thus will cause the system failed to operate. Fig. 8 shows the prototype of the Smart Parcel Box System.

Fig. 8. Prototype of Smart Parcel Box System

This system will be initialized when voltage source is supplied from the direct current (DC) to the system. As soon as the whole system start up, all components used in the Smart Parcel Box will automatically turn on. First, the LCD displayed "Welcome, Please Enter Password" as shown in Fig. 9.

Fig. 9. Interface at starting

Before delivery process, the owner need to fill out the parcel tracking number purchased and the password for the parcel box through application called AutoResponder WA, which is shown in Fig. 10. In order for the courier to gain access to the parcel box, courier need to send the tracking number of the parcel to the owner of the parcel box specifically. Then, courier will get the password from the owner after the courier sent the message. It can be seen in the Fig. 11 for example, shows that the courier sent the message to the owner and following that the courier gets the password for the parcel box. Courier must send the tracking number correctly otherwise the couriers will not get the password.

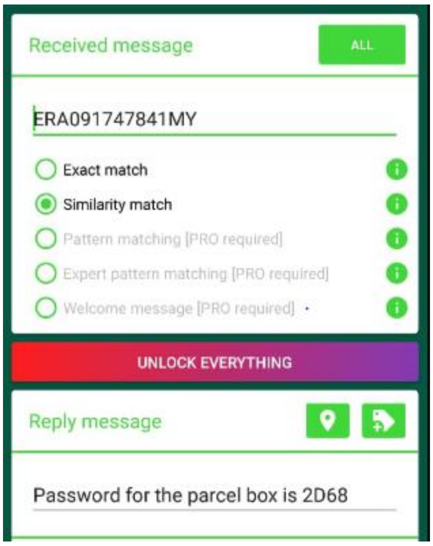


Fig. 10. AutoResponder WA

Fig. 11. Tracking number and password

Fig. 12 shows that the security of the parcel box has been checked by entering two different passwords. The first password is the correct password that will allow the door of the parcel box to open while the other password is the wrong password which cannot be used to open the door of the parcel box. As shown in Fig. 12, the LCD displayed “Access Granted, Door will open”. It indicates that the courier has entered the correct password for the parcel box. In Fig. 12, the LCD also displayed “Access Denied, Wrong Password”, shows that the courier must have entered the wrong password for the parcel box. The system used in this Smart Parcel Box System is very safe because every parcel box has its own password and need to have the parcel tracking number purchased first to get the password.

Fig. 12. Interface on the correct password and the wrong password

Next, the door of the parcel box operation depends on the system. As the system granted access to open the door, the door will open to receive the parcel which is as shown in Fig. 13. The LCD displayed “Please insert the parcel” when the courier inserting the parcel. The door will be opened until the courier finished inserted the parcel into the parcel box. The infrared sensor also had been installed in the system. To make sure that the door always open when the delivery attempt occurs, the sensor will detect the present of the courier or something in front of it. As the infrared sensor detect obstacle in front of it, it will cause the door to stay open until the courier leave the parcel box. Once there is no object detected by the infrared sensor, the system automatically closes the door as shown in Fig. 13 and lock to make sure that the parcel is safe in the parcel box. The LCD will display “DELIVERY ATTEMPT IS SUCCES”, as shown in Fig. 13, after the door of the parcel box is closed.

Fig. 13. Interface on the door operation, parcel and delivery attempt

Then, after delivery attempt completed by the courier, the door is closed. The owner of the smart parcel box will receive a message via SMS. GSM module will send a message to the box owner that the courier's attempt to deliver the parcel is successful. Fig. 14 shows the message sent by the GSM module to the owner of the Smart Parcel Box System.

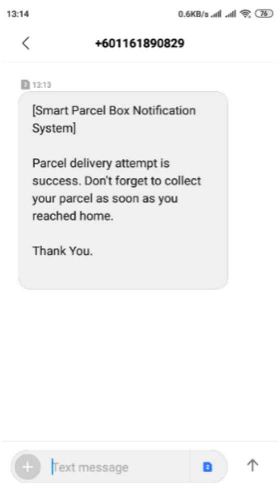


Fig. 14. Message sent by the Smart Parcel Box System to the user

4 CONCLUSION

The proposed Smart Parcel Box System provides user comfort and conveniences to buy item through online platform. With this system, customer become less worried about missing parcel, failed delivery of delay of delivery. Everything will be secured inside the smart parcel system which will be installed at customer's house. The system also provides user friendly apps for both parties where couriers only need to send the message which is the tracking number of the parcel to the recipient via app, and automatically get the password to gain access for the parcel box to deliver the parcel. This means that the Smart Parcel Box System ease the couriers work such as they do not need to knock on the door house or call the customer to wait for someone at their home to pick up their parcel.

For the recommendation and future improvement, it is suggested that instead of using AutoResponder WA application, a new application which typical for the courier company and the owner of the Smart Parcel Box System need to be invented. Nowadays, energy consumption is very important. So instead of using direct current, it is recommended to use solar panel or something that can harvest energy so that the energy used by Smart Parcel Box System is not from direct current.

5 ACKNOWLEDGEMENT

The authors would like to thank members School of Electrical Engineering, College Engineering, Universiti Teknologi MARA Shah Alam, Malaysia for the guidance and support during this project.

REFERENCES

[1] https://www.statista.com/statistics/1126355/philippinesimpact-on-online-purchase-behavior-covid-19/, June 2021

[2] Alias Abdullah, Mohd Shahrizal Azhari, Aimi Aqilah Mohamad Yusoff, "Coronavirus Impact on Online Purchases Behavior in Japan", International Journal of East Asian Studies, Vol 10, No 1, pp. 141-156, 2021.

[3] K. Hayashi, T. Nemoto, S. Nakaharai, "The Development of the Parcel Delivery Service and its Regulations in China," Procedia-Soc. Behav. Sci, vol 125, pp. 186-198, 2014.

[4] J. Castle, "Parcel delivery and postal problems" , 09 November 2016 [Online], Available : https://www.choice.com.au/shopping/shopping-for-services/services/articles/parcel-delivery-and-postal-issues

[5] POS MALAYSIA BERHAD, "Pos Malaysia Berhad - Annual Report 2016," 2016.

[6] A. Devi Pujari, P. Bansode, P. Girme, H. Mohite, and A. Pande, "Smart Letter Box System Using Obstacle Sensor For Notifies The User By Android Application," *Int. Res. J. Eng. Technol.*, vol. 10, pp. 2395–56, 2016.

[7] C. Alexander, "The Working Principle and Key Applications of Infrared Sensors," *Azo Sensors*, pp. 1–6, 2014.

[8] L. Louis, "Working Principle of Arduino and Using It As a Tool for Study and Research," *Int. J. Control. Autom.*

Commun. Syst. (IJCACS), Vol.1, No.2, April 2016, vol. 1, no. 2, pp. 21–29, 2016.

[9] M. Rahman, M. S. Ali, and S. Akther, “Password Protected Electronic Lock System for Smart Home Security,” vol. 7, no. 04, pp. 541–545, 2018.

[10] S. V. S. V. Prabhu Deva Kumar and S. Akashe, “Implementation of GSM Based Security System with IOT Applications,” *Int. J. Comput. Netw. Inf. Secur.*, vol. 9, no. 6, pp. 13–20, 2017.Y. Yorozu, M. Hirano,

Authors Introduction

Mr. Ahmad Luqmanulhakim bin Mohd Rusli

He received his B. Eng. Degree in Electrical Engineering from MARA University of Technology in 2019. He is currently a QC Engineer at Speedy Best Sdn. Bhd. His final year project about Development of Smart Parcel Box System.

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 interest is in the area of wireless communication (Physical and MAC cross layer optimization).

Mr. Meor Mohd Azreen Meor Hamzah

He received B.Eng. (Hons) degree from the University of Portsmouth, United Kingdom in 2001 and MSc in Telecommunication and Information Engineering in 2007 from Universiti Teknologi MARA, Malaysia. He is currently working as a lecturer at the School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA (UiTM) Shah Alam and member of BEM organization. His research interests include IoT applications and Embedded System.

Dr. Suzi Seroja Sarnin

She received her Ph.D. in Electrical Engineering in 2018 from the Universiti Teknologi MARA (UiTM). In 2001 she joined the UiTM where she currently leads a Centre of Communication Engineering Studies, Electrical Engineering Faculty. Her primary research topics include wireless communication, orthogonal frequency division multiplexing, internet of thing and long range wide area network. She is the author or co-author of more than 50 papers in international refereed journals and more than 50 conference contributions. The total awards that she has received through national competition is 3 gold, 3 silver and 6 bronze. This include Malaysia Technology Expo (MTE 2019), International Invention, Innovation and Technology Exhibition (IIDEX 2018) and The International Exposition of Research and Invention of Institution of Higher Learning (PECIPTA 2019).

Classification of Body Mass Index Based Face Images Using Facial Landmarks Approach and PCA plus LDA

Hasimah Ali, Ho Yong Kang, Wan Khairunizam Wan Ahmad
Faculty of Electrical Engineering Technology, University Malaysia Perlis, Malaysia

Mohamed Elshaikh
Faculty of Electronic Engineering Technology, University Malaysia Perlis, Malaysia

Norrima Mokhtar
Department of Electrical Engineering, Faculty of Engineering, University of Malaya, Malaysia

E-mail:hasimahali@unimap.edu.my

Abstract

Human faces contain rich information. Recent studies found that facial features have relation with human weight or body mass index (BMI). Decoding "facial information" from the face in predicting the BMI could be linked to the various health marker. This paper proposed the classification of body mass index (BMI) using facial landmark approach based on facial images. In this framework, Discriminative Response Map Fitting (DRMF) method has been used as feature extraction technique to detect and locate the facial landmark points on the facial images. About sixty-six (66) facial landmark points were identified. Only nineteen (19) of facial landmark points have been employed to extract the facial features in terms of distance and ratio features. A total of 221 facial landmark features were obtained and used as feature vector to classify the BMI classes. The rationale of using 221 facial landmark features is because these features were able to exhibit the unique characteristic of the BMI classes, which are normal, overweight and obese. Then, the extracted features were further reduced using Principal Component Analysis (PCA) plus Linear Discriminant Analysis (LDA) to map high dimension features into low dimensional feature with maximize between class scatter and minimize within class variations. Later, the reduced features were subjected to k-NN classifiers. A series of experiments has been conducted on MORPH II database using the reduced facial landmark features to classify the three BMI classes. Based on the experimental results, it shows that the reduced features using PCA plus LDA based on k-NN classifier has achieve the highest recognition rate with accuracy of 83.33 %. The obtained results show that the reduced facial landmark features were able to discriminate the three BMI classes of normal, overweight and obese, thus shows the promising results.

Keywords: Body mass index, facial images, facial landmarks, PCA plus LDA,

1 INTRODUCTION

Nowadays, humans have different way of communicating with other people, while the face is one of a human being's most impressive and special communication tools and plays an important role in visual communication [1-2]. Facial landmark is an abstraction and may have various definitions in previous studies. According to Galvánek et al. [1], the facial landmark is defined as a point that shares and has a clear biological significance for all faces. In computer science applications such as facial recognition, animation and speech analysis, facial landmarks play a prominent role. However, to automate facial landmark points such as artifacts created during 3D image acquisition, non-uniform facial expressions and positions and occlusions of facial features remains a challenging task and ongoing research[1]. In previous work, Coetzee et al.[2] have conducted the study in determining the relation between the facial cues with the body weight. They recruited two groups of Caussian and two groups of African subjects consist of both sexes with known BMI and measured their 2D facial images. They found that the width-to-height ratio, perimeter-to-area ratio, and cheek-to-jaw-width ratio are associated with BMI in males and width-to-height and cheek-to-jaw-width ratios are related to BMI in females.

In the other hand, Wen and Guo[3] developed a computational approach to predict the BMI from face images. They used an active shape model (ASM) to detect the facial fiducial points automatically. The correlation has been measured between the facial features and the BMI and it shows that all the correlation are significant based on the extremely small p-value obtained. The facial features consist of cheekbone width to jaw width ratio, width to upper facial height ratio, perimeter to area, average size of eyes, lower face to face height ratio, face width to the lower face height ratio and average distance between eyebrows and the upper edge of eyes have been extracted in predicting the BMI. The work of Wen and Guo were further extended by Jiang et al.[6] by proposing the fusion techniques by utilizing the psychology inspired geometric features (PIGF) and geometric facial representation-pointer feature (PF) defined by face shapes with a series of facial landmark points to extract a richer geometric representation. Deep learning feature representation has been used by evaluating the VGG-Face model for BMI prediction.

Recent years, a surprisingly phenomenon have been appeared in US, which is more than two-third of the adult are trend to be overweight or obese. In result, obesity have been confirming as a disease by American Medical Association due to the data show that nearly one out of every ten adult's death caused of obesity[3]. Due to the increase in weight and BMI value, in the end all this will come to a result which is getting personal health problem like diabetes, strokes and so on [4]. Thus, by take the advantage of BMI based facial images would be very crucial for us to aware the level of healthiness through BMI for each individual in order to encourage the awareness of health. Besides that, old way of detecting BMI is not convenience because of we need to calculate the BMI manually. So, detecting BMI using facial image not just another option for detecting BMI, but also a method that is non-invasive [3].

2 METHODS

Figure 1 shows the block diagram of the proposed method that consists of five main steps: original facial BMI database, facial landmark detection, feature extraction, feature reduction and BMI classification. Figure 1 shows the flowchart of the proposed method. Each of the blocks are explained in the following next subsections.

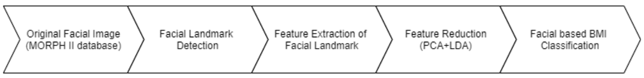


Fig 1. Block diagram of the proposed method

2.1. Image Database

In this study, a public database namely MORPH-II database has been used. It consists of 4206 facial images with measured BMI values. The database contains four categories of BMI which are underweight, normal, overweight and obese. Due to inadequate database for underweight category, we only considered three categories of facial BMI images such as normal, overweight and obese. A total number of 690 facial images (220 normal, 230 overweight and 240 obese) has been employed in the proposed method. Figure 2 shows the example of MORPH-II facial database with measured BMI for normal, overweight and obese category.

Fig 2. The example of MORPH II database for three BMI classes.

2.2. Facial Landmark Detection

In order to compute the features automatically, we used the Discriminative Response Map Fitting (DRMF) to detect facial landmark points first, and then calculate the distance and ratio features. The recently developed DRMF technique[5] can find 66 facial landmark points in a face image. We used 19 points out of 66 to produce the facial features that are related to BMI. Those points are marked as P$*$ in Figure 3 ($*$ indicates the $*$th points in the original list returned by the DRMF method). Note that even with the big number of 66 points detected by the DRMF, as shown in Figure 3 there are still some other points that are needed to compute the features.

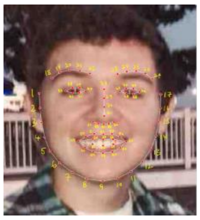


Fig 3. Illustration of facial landmark points for BMI classification

2.3. Feature Extraction

The 19 points out of 66 that define the face shape by a series of facial landmarks were extracted on facial images to produce the facial features. In this work, seven facial features were computed in terms of distance and ratio features. The rationale of using these 19 points (distance and ratio features) is because these features able to exhibits the characteristic of individual

BMI classes. They are visually shown in Figure 4. The meanings of these features are described below:

1) The landmark distance of point n to point m (n = 1, 2, 3, 4, 5, 6, 7, 8; m =10, 11, 12, 13, 14, 15, 16, 17)
2) The distance of point x to point 8 (x = 9, 10, 11, 12, 13, 14, 15, 16, 17)
3) The distance between point y to point 10 (y = 1, 2, 3, 4, 5, 6, 7, 8, 9)
4) The ratio of all of above (1, 2, 3)
5) The square root ratio of all of above (1, 2, 3)
6) Face width to height ratio (WHR)
7) Cheek to jaw width ratio (CJWR)

Fig 4. Illustration of the facial features used for BMI classification

2.4. Feature Reduction: PCA plus LDA

In this work, a total of 221 facial features was obtained and form a feature vector. Principal Component Analysis and Linear Discriminant Analysis were employed as feature reduction on the feature vector to transform the high features space into low dimensional space by maximizing between-class scatter as well as minimizing within-class scatter of the facial features. Due to these characteristic exhibit by PCA plus LDA, the intra-class and inter-class variations will be optimized.

2.5. BMI Facial Classification

In this study, k-Nearest Neighbour (k-NN) is used as classifier to classify the BMI facial classification such as normal, overweight and obese. The k-NN is one of the simplest algorithm that used in machine learning for regression and classification, but mainly used for classification. Compare to other classifier, k-NN will totally will save the time in training the data but more in predicting.

3 RESULTS AND DISCUSSIONS

3.1. Applied Facial Landmark Points on BMI Facial Images

Figure 5 shows the result of facial landmark points on facial images. A total of 66 facial landmark points has been identified. Although the DRMF technique has shown their advantages in the detecting facial landmarks, however there is also miss locating the facial landmark points that define the face shape. Figure 6 shows the example of miss locating the facial landmark points on the facial images. Thus, it may affect the accuracy of the BMI classification. Empirical studies show that only 19 points exhibit the unique features in which facial landmark points of point 1 to point 17, point 28 and point 52 are considered. From these points, distance points and distance ratio were computed to be used as facial features to classify the normal, overweight and obese. A total of 221 extracted facial landmarks features have been adopted for further investigation.

Fig 5. Results of the 66 facial landmark points

Fig 6. Miss locating the facial landmarks for obese category

3.2. Distribution of PCA plus LDA on Facial Features

In this experiment, the PCA plus LDA were conducted on original facial landmark features to map high dimension into lower dimensional and discriminate the significant features. Figure 7 presents the distribution of facial landmarks feature of three classes using PCA plus LDA in which class 1 is denoted as normal BMI (red dot), class 2 is overweight (green dot) and class 3 refers to obese (blue dot). It can be seen in Figure 7 that the distribution of three BMI classes are clearly discriminate. This can be inferred that the ratio between-class scatter of facial landmark features of the three class are maximized to that within-class scatter are minimized. However, small overlapping between the three classes still appear in the middle of the distribution.

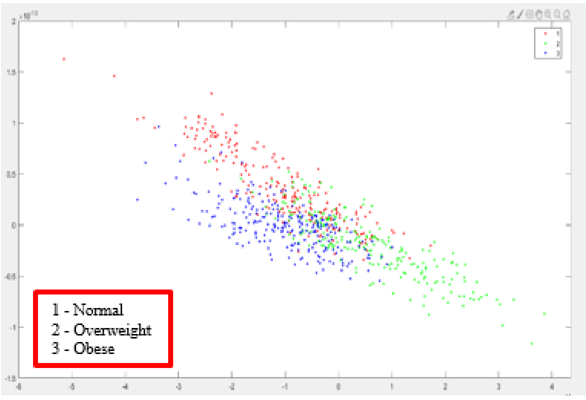


Fig 7. Distribution of reduced facial landmarks features using PCA plus LDA

3.3. Recognition Results and Discussion

To evaluate the effectiveness of the proposed method, a publicly available MORPH-II database of BMI Facial Images has been employed in this experiment. The DRMF technique has been used as feature extraction. There are 66 facial landmark points has been identified on the facial images. About 19 points out of 66 points which defines the facial shape were extracted and computed as facial features in term of distance and ratio features. A total of 221 facial features were extracted to be used as facial features. Further investigated on extracted facial features using PCA plus LDA as feature reduction before fed as input to the k-NN classifier. In this work, ten-fold cross validation has been used for training and testing the data. The nine folds of the data were used as training and the remaining ones used for testing. Each of the process was repeated for ten times at each fold and the final average accuracy was computed.

Figure 8 shows the classification rates of reduced facial feature using k-NN classifier for the BMI classification. The value of k was tuned from 1 to 12. Beyond this value, the accuracy tends to decrease.

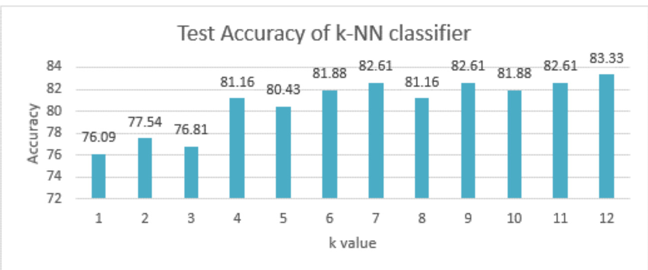


Fig 8. Classification rates of reduced facial features using k-NN classifier

As we can see from Figure 8 that the highest recognition rates was achieved when the value of k = 12 which is 83.33%. Table 1 shows the confusion matrix of using PCA plus LDA with k =12. Based on Table 1, it can be seen that the obese shows the highest recognition rate, which is 89.6%. Only 5 images were misclassified. On the other hand, the overweight gives the lowest recognition rate which is 71.7%. Thirteen out of 46 are misclassified. Seven of them were misclassified as normal and six were misclassified as obese. This can be inferred that, fiducial points of the overweight still mimicking the others BMI categories. Therefore, further study needs to be conducted to select and optimize the robust fiducial points in classifying the BMI facial images.

Table 1. Confusion matrix of reduced facial features by means PCA plus LDA using k-NN classifier (k = 12)

Status	Normal	Overweight	Obese	Average
Normal	**39**	7	2	88.6 %
Overweight	2	**33**	3	71.7 %
Obese	3	6	**43**	89.6 %
Total	44	46	48	**83.3 %**

4 CONCLUSIONS

This study has presented the classification of body-mass index based facial images using facial landmark points. It was observed that the DRMF able to identified the facial landmark points efficiently with 66 facial points. Throughout these facial points, only 19 facial points have been used to extract the facial features in terms of distance and ratio features producing 221 features. The rationale of using 221 facial features is because these features were able to exhibit the unique characteristic of the BMI classes, which are normal, overweight and obese. The used of PCA plus LDA in reducing the facial features helps to optimize the intra and inter class variation exist in BMI distribution. Based on the result obtained, it shows that the reduced features by means of PCA plus LDA with k-NN classifier has achieved the highest recognition rate with accuracy of 83.33 %. This results show that the reduced facial landmark features were able to discriminate the three BMI classes of normal, overweight and obese, thus shows the promising technique of the proposed method.

REFERENCES

1. Galvánek M., Furmanová K., Chalás I., and Sochor J. Automated facial landmark detection, comparison and visualization, in *Proceedings of the 31st Spring Conference on Computer Graphics - SCCG '15*, 2015.
2. Coetzee V., Perrett D.I., Stephen L.D., Deciphering faces: quantifiable visual cues to weight, Perception 39 (2010) 51–61.
3. Wen L. and Guo G., A computational approach to body mass index prediction from face images, *Image Vis. Comput.*, vol. 31, no. 5, pp. 392–400, 2013.
4. Hayford F. E. A., Appiah C. A., Al Hassan T., Asamoah-Boakye O., and Asante M., A pilot study comparing bioelectrical impedance analysis and body mass index in determining obesity among staff of a Ghanaian University, *Nutr. food sci.*, vol. 49, no. 2, pp. 240–248, 2019.
5. Asthana A., Zafeiriou S., Cheng S., and Pantic M., Robust discriminative response map fitting with constrained local models, in *2013 IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 3444–3451.
6. Jiang, M., Shang, Y. and Guo, G., 2020. *On Visual BMI Analysis From Facial Images*.
7. Gutin I. In BMI We Trust: Reframing the Body Mass Index as a Measure of Health. Soc Theory Health. 2018 Aug;16(3):256-271. doi: 10.1057/s41285-017-0055-0. Epub 2017 Oct 25. PMID: 31007613; PMCID: PMC6469873.

Authors Introduction

Dr Hasimah Ali

She is senior lecturer at Faculty of Electrical Engineering Technology, UniMAP. She received her PhD in Mechatronic Engineering from UniMAP (2015), MSc in Mechatronic Engineering from IIUM (2008) and BEng in Mechatronic Engineering from IIUM (2004). Her research interests are signal and image processing, facial expression, ground penetrating radar, robotics gripper and IoT.

Mr. Ho Yong Kang

He completed his BEng in Mechatronic Engineering from UniMAP (June 2021). He is currently working as a software engineer at Minebea Electronics Motor, Kedah. His research interests are image processing and face analysis.

Associate Prof Ir Dr. Wan Khairunizam Wan Ahmad

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.

Dr. Mohamed Elshaikh

He is a Senior Lecturer under Faculty of Electronic Engineering Technology (FTKEN), University Malaysia Perlis, Malaysia. He received his Ph.D. in Computer Engineering from University Malaysia Perlis, Malaysia, M.Sc Electrical and Electronics Engineering from University Technology Petronas (UTP), Malaysia and B.Sc Engineering Technology (Computer Engineering), University of Gezira, Sudan. His research mainly is on computer networking related.

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.

Story Units of the Types of Japanese Folktales and the Combination with a Noun Conceptual Dictionary

Jumpei Ono
Faculty of Software and Information Technology, Aomori University,
10-1 Seishincho, Edogawa, Tokyo, 134-0087, Japan

Motoki Kumagai
Faculty of Software and Information Science, Iwate Prefectural University,
152-52 Sugo, Takizawa, Iwate, 020-0693, 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

Story units in this study are the units described in Common Lisp based on the types of folktales Seki et al. analyzed. In particular, we have been developing story units to use in our narrative generation system, an integrated narrative generation system (INGS), as a type of narrative technique that generates a narrative structure based on the synthesis, transformation, expansion, etc., of a narrative structure. Story units function in INGS in combination with conceptual dictionaries. We attempted to combine story units with the verb conceptual dictionary. The objective of this study is combination with the noun conceptual dictionary. As the main constitutional elements of each story unit are verb and noun concepts, combining them with the noun conceptual dictionary enables the substantial function of story units based on the types of Japanese folktales as a group of narrative techniques in the INGS.

Keywords: Japanese Folktales, Narrative Generation, Noun Conceptual Dictionary, Story Generation.

1. Introduction

We have developed an integrated narrative generation system (INGS) [1, 2] that organically combines event generation using a conceptual dictionary [3] with narratology-based generation via a system [4, 5] based on Propp's theory [6] and Genette's narrative discourse [7].

For story generation using narrative structure, we targeted Japanese folktales. Ito organized folktales according to the Seki, Nomura, and Oshima's structure [8] and created folktale programs from the structure [9, 10] that we termed story units (Note: Volumes 1 to 10 contain the folktales Seki collected, and Volume 12 is a collection of contributed papers on folktales). Ito combined verb concepts with verbs in the created story units [11]. Ito analyzed nouns in noun phrases across the types of folktales. According to the results of the morphological analyzer MeCab [12], he classified noun phrases into 18 items and tried to combine nouns with noun concepts in the story units [13]. The goal of this study is to summarize Ito's results [13] and approach these issues and unresolved areas.

2. Our Study of Japanese Folktales and Story Units

A story unit is the knowledge used to generate a story, and it is a record of the story's structure. Fig 1 shows an example of the story unit. We created story units based on Seki, Nomura, and Ōshima's types of folktales [9, 10]. To make the story units available for story generation, we connected verbs and verb concepts in the created story units [11]. This made it possible to generate the structure of the events in the story. However, the story units did not bind to nouns.

```
(motif0669 (鬼の面[demon-mask]
   (被る 1[put-on] 見る 1[look] 逃げる 1[run-away]
    (or (取る 1[take] 帰る 1[go-home]) なる 1[become])))
         ((1 (event 被る 1[put-on] (agent (&sc 女[woman]))
         (object (&sc 鬼の面[demon-mask]))
         (location (&sc 山[mountain]))))
            (2 (event 見る 1[look]
          (agent (&sc 化け物[monster]))
          (counter-agent (&sc 女[woman]))))
            (3 (event 逃げる 1[run-away]
          (agent (&sc 化け物[monster]))))
            (or
             (4a
              (event 取る 1[take]
          (agent (&sc 女[woman]))
          (object (&sc 宝物[treasure])))
              (event 帰る 1[go-home]
           (agent (&sc 女[woman]))))
             (4b (event なる 1[become]
          (agent (&sc 女[woman]))
          (to (&sc 化け物屋敷の主人
              [host-of-monster-house])))))))
```

Fig. 1. Example of a story unit.

3. Content and Structure of the Noun Conceptual Dictionary

3.1. *Conceptual dictionaries*

The INGS has a conceptual dictionary [3]. Conceptual dictionaries are mainly referenced to generate events that comprise a story. There are noun, verb, and modifier conceptual dictionaries (adjective, adjective verb, welfare). This section explains the noun conceptual dictionary.

The noun conceptual dictionary has 5,809 intermediate concepts and 115,769 terminal concepts. The intermediate concepts indicate the classification of noun concepts, and the terminal concepts indicate the nouns that appear in the story. The hierarchical structure of intermediate concepts has 13 levels.

3.2. *Details about the noun conceptual dictionary*

Next, we provide a detailed description of the noun conceptual dictionary. The subordinate noun concepts that an individual intermediate noun concept can have are (1) intermediate noun concepts only, (2) terminal noun concepts only, and (3) both intermediate and terminal noun concepts. Thus, a hierarchical structure is formed. In addition, general nouns and proper nouns have different hierarchical structures. However, the description format is the same.

In addition to the symbol shown in Fig. 2, the noun concept corresponding to "terminal" is associated with the story unit nouns.

"Depth" indicates the position in the noun concept dictionary's hierarchical structure, which can contain values from 1 to 13.

"Hype" indicates an intermediate concept of parent.

"Hypo" refers to intermediate concepts of children.

"Terminal" shows intermediate concepts of children.

"Frame" does not currently have a meaning.

```
([Noun]
     (hierarchy
      (depth [Number])
      (hype [Intermediate noun concept])
      (hypo [Intermediate noun concept])
      (terminal [Terminal noun concepts]))
     (frame nil))
```

Fig. 2. The structure of an intermediate noun concept.

4. Combining Story Units and the Noun Conceptual Dictionary

Ito [13] attempted to combine nouns in the story unit with a noun conceptual dictionary. As a result, 3,695 nouns were combined with noun concepts.

The following is an explanation of the challenges of combining the types of folktales and noun concepts. There were 3,695 noun phrases across types of folktales. Of these, 2,641 noun phrases were mapped to the noun concepts. The rest are noun phrases that indicate the content of the characters' speech, such as "My horse is big," and here, the content of the speech is recorded in the story unit.

There are two types of noun phrases: single nouns, such as “snake” and “bull,” and nouns with modifying words, such as “good food.” The former can be examined directly in conjunction with the noun concept. In the latter case, we omit the elements that modify the word and examine noun concept binding. In any case, if the corresponding noun concept does not exist in the noun conceptual dictionary, we consider adding a new noun concept.

- In addition, it is important to consider the following. Use of language notation dictionaries: In Japanese, sentences are written using multiple types of characters. The Japanese language uses several kinds of characters to write sentences, so there are several ways to write a word. In Ito’s work [13], noun concepts were not checked using a language notation dictionary for multiple notation methods.
- In this paper, we would like to introduce the “noun concept” and the “proper noun concept.”

5. Conclusion

In this paper, we summarized the issues that arise when combining story units created from different types of folktales with noun conceptual dictionaries. In the future, we will address the story unit nouns that have not yet been combined with the noun conceptual dictionary.

References

1. T. Ogata, “Toward an Integrated Approach to Narrative Generation: Emerging Research and Opportunities”, IGI Global, (PA), 2020.
2. T. Ogata, “Internal and External Narrative Generation Based on Post-Narratology: Emerging Research and Opportunities”, IGI Global, (PA), 2020.
3. T. Ogata, “Building conceptual dictionaries for an integrated narrative generation system”, Journal of Robotics, Networking and Artificial Life, Vol. 1, No. 4, pp. 270-284, 2015.
4. S. Imabuchi, T. Ogata, A” Story Generation System Based on Propp Theory: As a Mechanism in an Integrated Narrative Generation System”, Lecture Note of Artificial Intelligence, Vol.7614, Springer, pp. 312-321, 2012.
5. T. Akimoto, T. Ogata, “Evaluation of a Narrative Discourse Generation System Based on the Concept of “Norm and Deviation””, Journal of Robotics, Networking and Artificial Life. Vol. 2, No. 1, pp. 50-53, 2015.
6. V. Propp, “Morphology of the Folktale”, University of Texas Press, 1968.
7. G. Genette, “Discours du Récit, Essai de Méthode, Figures III”, Paris: Seuil, 1972.
8. K. Seki, J. Nomura, H. Ōshima (Eds.), “Nihon mukashi-banashi taisei 11: shiryō-hen” [The complete collection of Japanese folktales, vol.11: data], Kadokawa Shoten, Tokyo (Japan), 1980.
9. J. Ono, T. Ito, T. Ogata, “Mukahi-banashi no motif no program ka to motif kozo no hikaku” [Programs of the motives of folktales and the comparison among motif structures], Proc. of the 61th Special Interest Group on Language Sense Processing Engineering, pp. 51-62, 2019.
10. T. Ito, J. Ono, T. Ogata, “Using Motifs of Folktales for Narrative Generation”, Proc. of the 59th Special Interest Group on Language Sense Processing Engineering, pp. 8-11, 2018. (SMC2018 Workshop on Informational and Cultural Narratology and Cognitive Content Generation.).
11. T. Ito, J. Ono, T. Ogata,” Implementing Story Units of Japanese Folktales with Conceptual Dictionaries.”, Proc. of 2021 the International Conference on Artificial Life and Robotics. pp. 191-196, 2021.
12. MeCab, MeCab: Yet Another Part-of-Speech and Morphological Analyzer. https://taku910.github.io/mecab/, 2006.
13. T. Ito, The Development and Generation Experiments of Story Units of Japanese Folktales Combined with Conceptual Dictionaries (Master’s thesis). Iwate Prefectural University, 2021.

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.

Mr. Motoki Kumagai

Motoki Kumagai has been an undergraduate student at the Faculty of Software and Information Science, Iwate Prefectural University since 2019.

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.

Visualization of the Unconscious in Quality Inspection in Manufacturing

Jun Nakamura*
Chuo University, Faculty of Global Management, 742-1 Higashinakano, Hachioji-City
Tokyo, 192-0393, Japan†

Nozomi Komiya
GInstitute of Policy and Cultural Studies, Chuo University, 742-1 Higashinakano, Hachioji-City
Tokyo, 192-0393, Japan
E-mail: jyulis.77f@g.chuo-u.ac.jp, nozokomisan@gmail.com
http://global.chuo-u.ac.jp/english/

Abstract

In quality inspection, which is the final stage of the manufacturing process, there is an operational manual on how the object should be inspected. In this study, we followed the gaze of the workers to search for the "way of looking" at the objects to be inspected, which is not described in the manual.

As a result of experiment, we found a difference between the reciprocating eye movements that occur in skilled worker and the static state of newcomers. It is suggested that this indicates a way of looking at inspection that has been unconscious until now.

Keywords: Skill Transfer, Experienced Worker, Reciprocal Motion, Quality Inspection, Manufacturing

1. Introduction

Contributions With the widespread use of IoT and AI, robotization has made remarkable progress in manufacturing processes, but the gap in labor productivity between small and medium-sized firms, which are inferior to large firms, is widening (SME Agency, 2018). As a matter of fact, there are still many handy manual operational processes in the sites of SMEs, and human resource development through improvement activities and creativity is required. For this reason, the authors have been conducting research on organizational learning in SMEs in terms of the skill transfer and improvement of work efficiency (Nagayoshi and Nakamura, 2017).

While daily improvement activities are being carried out at workplaces, the reality of efforts to improve work efficiency is that they are often focused on work procedures, which are merely defined in work instruction handbook. Therefore, the authors decided to focus on quality inspection, which is greatly influenced by human resources, and to analyze the skilled worker from the perspective of human resource development. At this time, we believe "work procedures" are insufficient for quality inspection, and how do they actually inspect the precautions written in the work instruction handbook? Therefore, the authors decided that gaze measurement would be effective against visual inspection, and conducted an experiment using eye tracking system in this paper.

The history of gaze measurement is very old. In the past, the direct observation method was used to observe the movement of the cornea and capillaries with the naked eye, and the afterimage method was used to project

*Refer to the first author's laboratory web page : https://j-nakamuralab.com/

the generated afterimage on the retina onto a screen and quantitatively measure the movement (Hashimura, et al., 2015). When it comes to the current direct observation method, for example, Tobii Technology's Tobii Pro Glass 3, has 16 LEDs and 4 eye cameras built into the lens to measure eye movements (Tobii Technology, 2021a), and we take this Glass as a tool to observe eye tracking for quality inspection worker in a manufacturing factory.

2. Quality Inspection of Fuel Cell Piping

The pipes used in household fuel cells are long and narrow. Since the pipes are bent at several points, the quality inspection has to be done manually (Fig. 1). In other words, inspection by machine is currently difficult. The fittings at the edge of the pipe and the exterior are inspected for scratches, dents, deformations, and other abnormalities. In this process, the bent parts are inspected for wrinkles, and the fittings of both edge sides of the pipe in particular are slightly protruding, making them difficult to see from some angles. For this reason, we look at the edges from various angles to see if there are burrs, scratches, dirt, etc., and if so, we polish them with a file or brush. Once the five products that have passed the appearance inspection are assembled, they are placed in a neat row and inspected again for comparison of appearance.

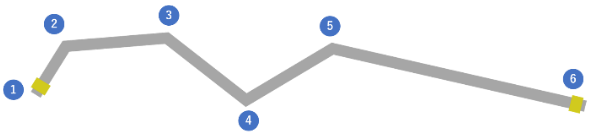


Figure 1: Image of piping (actually 3D instead of 2D)

In this paper, with the cooperation of Nihon Pipe System Co., Ltd. the workers of visual inspection of fuel cell piping were asked to wear glasses for eye measurement and to perform ordinary quality inspection work.

3. Methods

The gaze measurement experiment was carried out as follows.

- Location: The factory of Nihon Pipe System Co., Ltd.
- Date: Tuesday, August 3, 2021
- Workers: 1 veteran (the most skilled worker)

As mentioned above, the object of the quality inspection is the piping of the fuel cell, and the final product is the eco-cute. We paid attention to the edge where the notes on the test were written and observed carefully while playing back the gaze measured video. In order to capture changes in a particularly short period of time, the video was played back at 1/16th the normal speed.

4. Result

As a result of the eye-tracking of the experienced operator, the quality inspection of 15 pipes in 377.56 seconds was recorded. The average speed of the inspection was 20 seconds per pipe, although there were some inspections of five pipes side by side for comparison and re-inspection of one of the 15 pipes.

In eye-tracking, fixation is often extracted and the location of the gaze is observed. However, since the time required for the quality inspection of piping is so short that the extraction of fixation is limited, we decided to capture the movement of the eye and follow its trajectory. Among the inspection processes, we focused on the inspection of both side edges, which requires changing the angle of gaze, and found that the longest time was 0.521s while the shortest was 0.181s. The results are shown in Figures 2 and 3. We can see that both eye movements look like a reciprocating motion.

As for the pipes that were noticed during the visual inspection, those pipes are to be polished with a brush to remove burrs. Immediately afterwards, the edges are held with both hands to check its condition after polishing. Figure 4 shows the trajectory of this process. As in Figs. 2 and 3, we found that a reciprocating motion is performed.

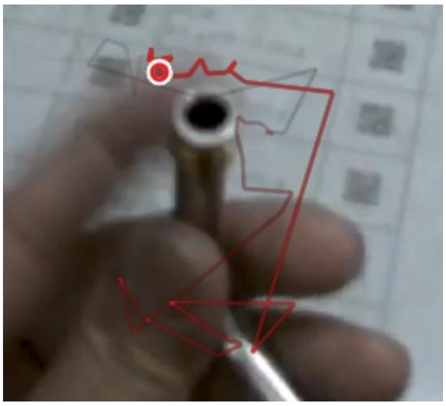


Fig. 2 Experiences worker's inspection of edges (The duration of the movement is 521ms)

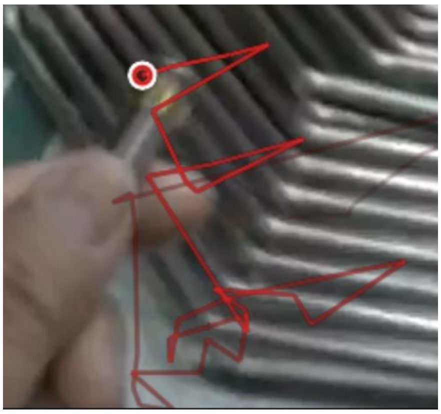


Figure 3: Edge inspection of experienced worker (The duration of the movement is 181ms)

Since the time required for the inspection was less than one second, it turned out that the workers themselves were so unconscious that they were not even aware that they were reciprocating.)

5. Discussion

Here, let us consider the reason for the reciprocating motion of the eyes in quality inspection. Denso Corporation is using eye tracking for inspection at some manufacturing sites called Nishio-Seisakujo, where the key point of the test is to let the gaze flow as if spelling out the letter "2" as shown in Figure 4 (Tobii Technology, 2021b).

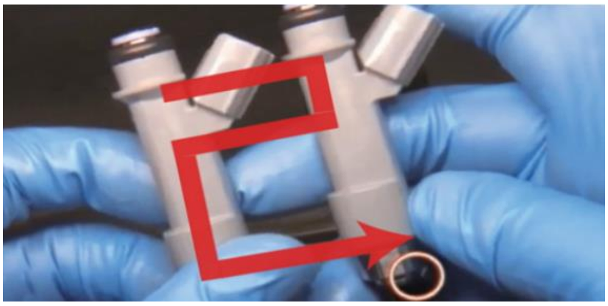


Figure45: Key points of quality inspection at Denso (citing inspected images from (Tobii Technology,2021b)

The Denso case is described as if spelling out the letter “2”, which can be broadly interpreted as a reciprocal motion of the eyes.

The above interpretation might support the following three points which could be inferred.

First, since the act of searching for a problem is a part of the normal inspection process, the pattern of concentrating on a fixed point may not be possible. Rather, the fixed pointed gazing pattern may occur after a flaw as an example is found in a quality inspection.

Second, the remarkable factor is the reciprocating motion. At first glance, it appears as if the subject is going through a trial-and-error process of looking at the target and then checking it again while slightly moving backward. However, since this is done in a very short period of time, less than one second, it can be inferred that the reciprocating motion of the eye itself is unconsciously performed.

Third, the question is whether the reciprocal motion of the eye is left-right, up-down, or diagonal. In Figs. 2 to 4, the eye movements are generally left-right, although there are some angular variations in the eye movements. On the other hand, in Figure 1, the gaze is up and down, and the reciprocating motion is quite short. This may be due to the difference in the shape of the object and the angle at which the gaze flows.

Fourth, with regard to inspecting the edge of the pipe viewing from the front of the edge as if you were looking at the cross section of the pipe, it takes more time than setting the pipe on its side looking across the edge, as mentioned in the chapter of result. The cross section is checked against the circumference of pipe, while the side is in a sense a straight line. This may have an effect on the speed. In addition, it is easy to understand that it takes more time to brush the same section after than before.

By the way, gaze can be further divided into three types: Tremor, Drift, and Microsaccade (Martinez-Conde, at al., 2004). Tremor is a small back-and-forth motion that makes observation difficult, Drift is the initial gaze that occurs simultaneously with Tremor or during Microsaccade, and Microsaccade is a short and fast viewpoint motion to correct Drift (Cornsweet, 1956).

The reciprocating motion of the eye in this paper may be Tremor, Microsaccade, or Drift, but it needs to be further explored in the future. In this paper, we focused on the edge of the pipe, but we also observed large reciprocal eye movements such as ①→②→③ and ③→②→① when we viewed the pipe from the side as shown in Fig. 1.

The most interesting feature of these reciprocating movements of the eye is that even in a very short time, decision-making could be executed to evaluate the test results. That means, even though the reciprocating movements are unconscious, judgments are made consciously by individual. This might be similar to driving a car, where instantaneous judgement, its evaluation and one's decision-making, are carried out simultaneously.

6. Conclusion

In this paper, we visualized and discussed the eye movements during the quality inspection of piping. Reciprocating eye movements were found during the quality inspection, although unconsciously, and the factors behind them were discussed. The analysis is based on the measurement results of only one experienced worker, and the limitation is unavoidable. The bottleneck is that it is not possible to obtain a large number of samples due to the limited number of workers for quality inspection, but this becomes a management issue in terms of skill transfer and human resource development.

In the future, we would like to conduct research on the visualization and analysis of the differences between experienced and new workers, as well as the view of piping from the side, which will lead to the skill transfer.

Acknowledgements

This paper was made possible with the understanding and cooperation of Mr. Tano, President of Nihon Pipe System Co., Ltd. We would like to express our sincere gratitude to all of them including the subjects.

This paper has been supported by Grant-in-Aid for Scientific Research (JSPS KAKENHI Grant Number 17K03872 and 19K03062) and Grant-in-Aid for Specially Promoted Research from Chuo University.

This paper has been translated into English and revised from a presentation made in Japanese at KST: Special Interest Group (Type2) on Knowledge and Skills Transfer.

References

1. Small and Medium Enterprise Agency, "SME White Paper", 2018.
2. S. Nagayoshi, J. Nakamura, "How accelerate Knowledge Acquisition and Information Distribution in the Organizational Learning from Failure", Proceedings of the 21st Pacific Asia Conference on Information Systems (PACIS), Langkawi, Malaysia, 2017.
3. M. Hashimura, H. Iizuka, J. Lee, "Special Issues No.3: Measurement Technique for Ergonomics, Section 4: Measurements and Analyses of Bioelectric Phenomena and Others (2) Measurement of Eyeball Movement", The Japanese journal of ergonomics, Vol. 51, No. 6, pp.406-410, 2015. (In Japanese)
4. https://www.tobiipro.com/product-listing/tobii-pro-glasses-3/
5. https://www.tobiipro.com/ja/applications/industry-and-human-performance/case-studies/case_DENSO/
6. S. Martinez-Conde, L. Stephen, D. Macknik, H. Hubel, "The role of fixational eye movements in visual perception", Nature Review Neuroscience, Vol. 5, No. 3, pp. 229-240, 2004.
7. T. W. Cornsweet, "Determination of the stimuli for involuntary drifts and saccadic eye movements", Journal of the Optical Society of America, Vol. 46, No. 11, pp. 987-993, 1956.

Authors Introduction

Dr. Jun Nakamura

He is a Professor of Faculty of Global Management at Chuo University in Japan. He graduated from the Department of Technology Management for Innovation, School of Engineering, University of Tokyo, with highest honors. His area of research is cognitive sciences.

Dr. Nozomi Komiya

She is a Guest Researcher of Institute of Policy and Cultural Studies, Chuo university. She graduated from University of Tsukuba(Doctoral Program in Risk and Resilience Engineering). The area of her research are Intelligent Information, Cognitive Science.

The Study on the Relationship Between the Comic Artists' Styles and the Visual Languages: From the Stylistic Changes in the Work of Japanese Comic Artists

Kaori Otsuru
The department of Comics, The Tainan University of Technology, Tainan City, 710302, Taiwan[*]

E-mail: t40128@gm.tut.edu.tw
https://www.tut.edu.tw

Abstract

Recently, attempts have been made to analyze manga using machine learning. In this research, a manga analysis method is proposed that combines iconography and text for the purpose of understanding the "style" peculiar to manga. Specifically, using the framework of linguistics and narrative story, the elements that form the impression of the work, the micro / macro level development of the story, and the synchronic and chronological transition of the author's drawing style are focused on.

Keywords: manga, comic, style, synchronic transition, chronological transition, machine learning

1. Artistic and linguistic perspectives on comics

Neil Cohn said in his book: *The disparaging viewpoint of this similarity would claim that the other artists were simply "knock-offs" or "clones" of the original. From a linguistic perspective, let's consider a different viewpoint: these drawers all use similar cognitive patterns in the way that they draw, and those similarities constitute a shared visual vocabulary that reflects and reinforces an identity of belonging to a particular group of people* (i.e. their studio and their subculture)[†]. Comics originally developed from two perspectives: the "Art Frame" (imitation is a suppression of creativity, and innate imagination should be developed naturally) and the "Language Frame" (it grows and develops by storing patterns in memory through the influence of others, appreciation and practice)[‡]. In a sense, comics is an art form that emphasizes its uniqueness, but also has a linguistic property that develops through interaction with others.

2. The imitative property of comics

In his essay *Manga Genron*, the "manga" scholar Inuhiko Yomota describes the differences in style that appear in manga as "stylistic", borrowing a phrase from poetry and novels. He notes that it is possible to create a sense of discomfort by mixing characters drawn in a gag style, as in the works of Fujio Akatsuka, for example, with characters and backgrounds drawn in a *gekiga* style, as in *Golgo 13*, for example. Indeed, a reader who is used to reading manga can see the incongruity of the designs at a glance, and take it as a meta-nonsense gag. Here, the "style" that Yomota refers to depends on the so-called genre, and he focuses on the heterogeneity that emerges when different "styles" of genre are mixed together. However, "genre" in "manga" is not self-evident, and its definition is vague. Rather, genres are simply the collective flow of styles that have been recognized and subsequently named by their predecessors. In this paper, I would like to consider the change of drawing as a part of the style of poetry and novels, which is called "individual style" (as Yomota also understands) and includes the atmosphere that the artist cannot control.

3. Regional differences

Cohn points to the existence of a visual vocabulary, idioms and syntax in comics, and then addresses regional differences. According to him, there are two streams of comics drawing in the United States: the *Kavian* style and the *Berkshian* style. The former is more nasal, more realistic in its depiction of muscles and emphasizes the distinction between female and male appearance. The latter is the typical cartoon style of drawing as it is known today, as seen in the early *Walt Disney films*. It is characterized by a small number of fingers on the hands, a lack of joints, and strange bodies.

On the other hand, the visual language of Japanese manga is characterized by large eyes representing goodness and narrow, pointed eyes representing seriousness and vice, pointed chins for women, diverse chins for men, and very large hair. The features depicted in this chart are certainly features of Japanese comics, but the actual drawings in the book are, at least in my opinion, more characteristic of the manga and anime of the 1990s. This shows that Cohn himself does not take into account the different styles of Japanese manga genres and their changes.

4. A style dependent on a particular genre

Natives of Japanese “manga” drawing, so to speak, may find it difficult to speak for themselves of a typical drawing stream in Japan. But that is not to say that there are none. The web animation *Senpai Club* (2014) by the Swedish unit “*makebabi.es”* emphasizes the characteristics of Japanese school romantic comedy and other genres of animation and manga. Despite the fact that the Japanese language is a mess and the drawings are rough, it is possible to discover the style of Japanese animation drawings. Another example is *Idol Densetsu Artemis*, an online manga by “Jojimura”, who has published several works on the Internet, in which the flow of panels and scene transitions are more than mature enough, but the drawing is like that of a girl who has just started to imitate manga and animation. But the drawings look as if they were done by a girl who has just started to imitate manga and animation. Neither “*makebabi.es”* nor “Jojimura” have the same style drawing works. However, viewers and readers recognize that they have seen similar works in the past, and they recognize a flow of atmosphere and style that integrates the characteristics of the drawings, so to speak.

5. Changes in personal drawing

For many professional cartoonists, it is almost impossible to say that the styles of their drawings have not changed over the years since their debut. It is thought that “manga” drawing begins with imitation of existing works, followed by deliberate changes such as drawing practice and the search for one's own style. At the same time, there are likely to be unconscious influences and changes, such as unknowingly changing one's drawing according to one's favorite works, or becoming more sophisticated as one continues to draw, or developing habits, good or bad. In this way, the style of many professional comic artist changes in ways that cannot be explained by "growth". It is not possible to make a general assessment of the changes as to whether they are good or bad. Some comics artists give the impression that their drawing has become more sophisticated in the years since their debut, while others receive comments from fans that their old drawing was better.

In this article, we will examine two manga artists who have been working in comics for more than 30 years to see how their drawings have changed. One is Kouga Yun, active mainly as a girl's manga artist and coterie artist, and the other is Obata Takeshi, a drawing artist active mainly in *Weekly Shonen Jump*.

Kouga celebrated her 30th anniversary in 2015 with the publication of a commemorative art book. In addition to her coterie magazine activities, she began publishing her work "*Wakakusa Monogatari*" e.g. in commercial magazines in the 1980s and is still an active artist today. "*Aashian*" started 1987 is likely a Girl’s manga at that time. You will find that a long time make character’s faces changed in "*Choju Densetsu Gestalt*", it spent 1992-2001. “*Sato-kun to Tanaka-san to”* is published 2007. That draw style has been renewed. Over the past thirty years, her drawing style has certainly changed, but in the atmosphere of her drawings from all eras, one can recognize the face of Kouga Yun in her works.

Takeshi Obata (1969-) was debut in 1989 with *CYBORG Grandpa G* (then under the name Shigeru Hijikata) in *Weekly Shonen Jump*. Perhaps we can’t realize it is his work because the character drawing is so different from now. However, it doesn’t mean that He was not good at drawing. The reason for this is that we find in his work a wonderful manga effects in "*Majin Bouken Lamp Lamp*". In fact, he has always been a man capable of drawing at the highest level of each period. In "*Otogi Zoshi Ayatsuri Sakon*", we can look the 90’s Japanese animation style character face. His work on the hit series *Hikaru no Go*, which ran from 1999 to 2003, is where his drawings changed the most. He continue to use that style in "*Death Note", and* remarkably changed in "*Bakuman*.", but at "*Platinum End*", he returned his style to "*Death Note".*

This change in the individual artist is likely to lead to a change in the larger trends of the manga industry in Japan, as we can see that the tendencies in the way characters are drawn in 1950 and 2020 are very different, in terms of head size, hairstyle, eyes and line thickness. It is easy to recognize that works of different genres are from different periods. In addition to the historical differences, the huge size of the industry has led to the branching out of characteristics from genre to genre and magazine to magazine, and it is thought that differences in style appear, so to speak, like dialects.

6. The style of comics: social influences and the singularity arising from personal nature

The phrase "the drawings are old" is a painful criticism for manga artists. Of course, in a commercial magazine, "it's all right if the drawings are old, as long as they are interesting", and this may not be a particular concern. However, "new drawing" is definitely a weapon for the comics industry in Japan to keep offering something new. What I would like to focus on in this article is the fact that authors and readers have a sense of what can be called "old/new drawing", and that the drawing itself changes in this way. In a sense, the drawing of comics has a social properties, propagating and changing like a language. If an artist's drawings, panel layout and atmosphere are to win the hearts of many fans, the stylistic elements must be considered not only in terms of personal factors (talent, training, knowledge, etc.), but also in terms of social factors (influence of other comics, period, magazine, field, social background, etc.).

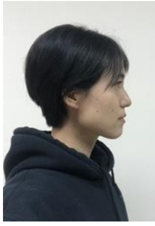
Time-changing comics artists change their style in order to incorporate fresh influences, while at the same time transmitting a cutting-edge style that influences others. It is not only the "Art Frame", i.e. the individual elements of individuality and talent, but also the "Linguistic frame", i.e. the multiple conscious and unconscious elements that influence and change each other in a language-like manner, that leads to the rich diversity and interest of manga styles. The study of comics as a visual language may not only advance the study of comics themselves, but may also help to clarify the nature of human language. In the future, I would like to study the existence of manga writing styles and the elements that make up these styles.

7. References

1. Cohn, Neil "*The Visual Language of Comics: Introduction to the structure and Cognition of sequential Images*" A&C Black, 2013.
2. Yomota, Inuhiko "*Manga Genron*", Chikuma Bungei Bunko, 1994.
3. makebabi.es "*Senpai Club*", 2014
4. Jojimura "*Idol Denshetsu Artemis*", 2018
5. Yun Kouga
   -"*Wakakusa Monogatari*" Gakushu Kenkyusha, 1985.
6. –"*Aashian*" "Monthly Wings" Shinshokan, 1987 -1988.
7. –"*Choju Densetsu Gestalt*" Enix, 1992-2001.
8. –"*Sato-kun to Tanaka-san*", "Comic Zero-Sum More", Ichijinsha, 2007.
9. Takeshi Obata
   -"*CYBORG Jichan G*", 1989.
10. –"*Majin Bouken Lamp Lamp*", Shueisha, 1991 -1992.
11. –"*Otogi Zoshi Ayatsuri Sakon*", Shueisha, 1995 -1996.
12. –"*Hikaru no Go*", Shueisha, 1998 -2003.
13. –"*Death Note",* Shueisha, 2003 -2006.
14. –"*Bakuman*.", Shueisha, 2008 –2012.
15. –"*Platinum End*", Shueisha, 2015 -2021.

Authors Introduction

Ms. Kaori Otsuru



She received her Doctor's degree from the Department of Social System Studies, The University of Kitakyushu, Japan in 2018. In 2020 she published the book "The Anthropology of Relationship Repair", based on her doctoral thesis, and in 2021 she was awarded the 20th Society Prize at the Japanese Society of Oceanic Studies.

Theoretical Backgrounds toward Text Mining for a Phenomenological Model of Taste Perception

Hiroki Fxyma
Department of Food and Beverage Servises, Tainan University of Technology,
No.529, Zhongzheng Rd., Yongkang District,
Tainan City 710302, Taiwan (R.O.C.)

Abstract

The experience of taste is constituted by the multilayered intersection of time aspects, such as the preceding eating experience, and the post-eating experience of looking back and re-sensemaking the tasting in a holistic way.This presentation will outline the key concepts in approaching the problem of how machines or humans perceive, recognize, and construct internal narratives of taste from the aspect of time.The experience of taste does not only unfold on a physical time axis in the mouth.

Keywords: Appreciation of Taste, Aesthetics, Phenomenology, Sake

1. Introduction

There are two genres of art: spatial art (e.g. painting) and art that develops over time (e.g. music). Leaving aside the classical argument that taste or odor are not an aesthetic subject, the appreciation of taste seems to have both a temporal aspect (referred to as "top notes" or "after-flavors") and a spatial aspect (think of the spread of whiskey when it is watered).In this article, I will discuss the temporal aspect of the act of tasting, including physiological time, cognitive time, and other concepts that may be relevant to understanding taste as a phenomenon.

2. Physiological and cognitive aspects of tasting time

2.1. *Food aversion learning and food preference learning*

When animals, including humans, are given drugs or radiation therapy that cause visceral discomfort, they learn to avoid eating the food they ate before the treatment, which is called taste (food) aversion learning (Garcia et al., 1955). Taste aversion is a type of learning in which a temporally preceding eating experience influences subsequent eating and preference patterns.

2.2. *Ortho-nasal / Retro-nasal*

There are two pathways for perceiving the aroma of food and drink: the ortho-nasal path is the air pathway through the nasal passages to the olfactory epithelium. On the other hand, the path of air from the oral cavity through the nasal cavity to the nostrils is called the retro-nasal path. The retro-nasal path is critical in the perception of flavors. In the case of sake, the aroma of the sake in the glass is ortho-nasal, while the aroma in the mouth and after swallowing is retro-nasal.

2.3. *Temporal changes in the mouth*

What changes occur in the mouth between the time food enters the mouth and when it is swallowed? When we look at a picture, we hardly see any changes in the picture itself. Still, there are significant physical and scientific changes taking place in the mouth during taste appreciation. The following is an overview of the changes that food undergoes in the mouth over time in four areas: retention, mastication, stirring, and the effects of saliva (dissolution and chemical changes).

2.4. *Temperature change due to retention*

The oral cavity is usually maintained at around 37 degrees Celsius. This is due to the high concentration of capillaries in the oral cavity, especially in the tongue.

After licking ice, the surface temperature returns to its original temperature within 10 to 20 seconds, waiting for the next food to arrive. The food we eat varies in temperature from below freezing to nearly 100 degrees Celsius. These foods undergo dramatic thermal changes during the few to ten seconds when they pass through the mouth. For example, cold ice cream is melted by the body heat of the oral cavity and changes its shape, while hot coffee is sipped and stays in the mouth, lowering its temperature to a level that is not harmful to the body. Because of this trust relationship, there are few temperature sensors in the esophagus and stomach. The function of this temperature change due to retention is to bring food closer to body temperature and render it harmless. Another critical role is to change the taste of the food by changing the temperature.

2.5. *Shape changes due to mastication (chewing)*

Just as chewing grapes dramatically change the sense of taste, shape changes caused by chewing have a significant effect on taste. There are some foods and hedonic items such as chewing gum and chewing tobacco that are flavored only by chewing.

Mastication is a semi-voluntary movement, and it is known that there are large individual differences in mastication patterns (Chen and Stokes, 2012). Mastication changes the volume of the oral cavity. This change in volume causes air to be pumped into the throat like an accordion or bellows, which then passes into the nose, creating a retro-nasal flavor.

2.6. *Dissolution by saliva*

Foods change their shape when they are chewed, and at the same time they are dissolved by saliva and transformed into different substances by chemical reactions. The most common example is the breakdown of glucose by saliva (glycolysis). Therefore, the taste in the mouth and the taste when swallowed are different in quality. This is partly due to cognitive factors, but it is also due to scientific changes in food.

2.7. *Stirring by the tongue*

Stirring by the tongue takes place in the oral cavity in synchronization with mastication. By taking enough time to chew and stir properly, a complete picture of the food in the mouth can be revealed.

Other physiological issues related to time include the following topics:

- Taste masking effect (tequila with a bit of lemon)
- Flavor enhancing effect (*zenzai* with *shio-konbu*, salty-kelp snack)
- Synchronization of flavors (the spice flavor of curry synchronizes with the savory flavor of aged sake)
- Acclimatization (adaptation to the smell)

When animals, including humans, are given drugs or radiation therapy that cause visceral discomfort, they learn to avoid eating the food they ate before the treatment, which is called food aversion learning (Garcia et al., 1955). This is called taste aversion learning, and food aversion learning is called food aversion learning. This is a type of learning in which a temporally preceding eating experience influences subsequent eating and preference patterns.

3. Time as Phenomenon

In this section, we will deal with time as a phenomenon. Time as phenomenon means considering the aspect of time that differs from so-called objective time (physical time, natural time, world time, etc.). It is an attempt to discuss the flavor of time as it is told in the personal, first-person narrative, rather than objective or measurable time.

Phenomenology is not the only discipline that deals with first-person phenomena, but it is certainly the one that should be given maximum consideration.

In phenomenological terms, first-person time means "reducing physical time to phenomenal time as it is experienced," and this reduction leads to Consciousness of Internal Time (Husserl, 1928).

In this study, I emphasize the active aspect of the appreciation of taste, and time in the attitude of subjectively constructing the taste. The theory of time that goes well with such a theory is the argument of Derrida, Ricoeur, and Merleau-Ponty, and in Eastern philosophy (especially as interpreted by Toshihiko Izutsu), the theory of time in *Kegon* and Islam. In "*Time and Narrative* (*Temps et Récit*)," Ricoeur criticizes Husserl's monotonous and unidirectional "flow" of time, and proposes "narrated time" as a different structure of time. In "Phenomenology of Perception," Merleau-Ponty argues for the significance of time as a subjective construct, as follows:

> It is of the essence of time to be in process of self-production, and not to be; never, that is, to be completely constituted. Constituted time, the series of possible relations in terms of before and after, is not time itself, but the ultimate recording of time, the result of its *passage*,

> which objective thinking always presupposes yet never manages to fasten on to.
>
> (Merleau-Ponty, 1962, p.482)

In this way, I believe that we should not have an attitude of sensing sake that exists as a third person, but an attitude of how sake – that has appeared as a first person – is interpreted, and composed (generated) a narrative.

3.1. *The emergence of Sake*

I have been proposing a new approach to define sake taste by focusing on adjectivals, as an alternative to a dominantly used method that focuses on nouns (see also, Fxyma, 2022). I will call this latter way of verbalization an "object-motivated event construction", where the experiencer primarily uses nouns to describe the event of tasting. This can be commonly found in an English tasting comment by a wine sommelier as in: "I feel a note of black cherry, cassis, and the rich flavor of the oak." where the sommelier detects the elements of the flavor, verbalizing them, perhaps selecting the terms from his or her list of tasting words.

This is analogous to an "audio" or a "visual event construction", where an event is reported objectively. For instance, if someone witnesses a traffic accident, the witness might construct the event as follows:

(1) "there are two cars" "two cars collide"

(1) is illustrative of an "object-motivated event construction" where the focus is placed on identifying the event participants (i.e., cars), just like the sommelier identifies the flavors.

There is an alternative way of reporting the same scene, as in (2), which I tentatively call an "emergence-motivated event construction."

(2) something happened! something crashed! Oh, two cars crashed.

This sequence might be thought of as merely a play of things being flipped around. However, the portrayal of the scenes in (1) and (2) are epistemologically distinct. The object-motivated event construction (1) and the emergence-motivated event construction (2) show the opposite approaches. If language interacts with cognition, or if language forms our thoughts, then to have the correct cognition theory is indispensable in the performance of proper language analysis.

I proposed to adopt the approach by the emergence-motivated event construction to define sake taste. Modeling (2), the sequence of cognizing tasting experience can be presented as in (3):

(3) taste emerges what is this taste? I feel the taste of X and Y.

The proposed analysis of event construction of tasting is not the same as that of the visual event construction. When we taste something—that is, when we have an event construction of tasting or when we conceptualize what we taste in our mouth—what we feel first is not the element of taste, such as sweetness, acidity, apple flavor, or other flavors in sake (as expressed by nouns), but the emergence of the tasting event itself.

Supporting emergence-motivated event construction means that adjectives, adjectival nouns, and verbs (but not nouns) take the leading role in the tasting description. The recognition of the emergence of an event is primarily expressed by adjectives and adjectival nouns. They are no longer merely modifiers but play a critical role, enabling us to encode the inceptive stage of our tasting experience.

3.2. *Generating Flavor*

After its emergence, sake is not brought to our consciousness in a constant and homogeneous manner. If we take the tongue as the subject, the tongue does not swim in a pool of homogeneous water quality, but swims in a river as a dynamic environment with currents and stagnation. If you think that sake in a glass is homogeneous, please look back at the previous section. Sake undergoes many actions in the mouth, including retention, stirring, and reaction to saliva. The information brought to the tongue is always changing and dynamic.Here, sake appears with shades of information (what Gibson calls "texture"). We find the gradation, intensity and singular points of the information through the interaction with the environment colored by the texture, and compose the figure and the ground.

4. Future Themes

There are many issues that could not be dealt with in this paper. One of them is the lack of discussion on the separation of the concepts of time and memory, although we have dealt with cognitive time. The senses also have short-term and long-term memories, but there is not enough discussion on whether the discussion of time can be integrated into the discussion of memory, or whether a different explanatory principle is needed than that of memory. We also missed the issue of how to compromise the phenomenological theory of time with the representations in cognitive science. This is an issue that cannot be avoided when considering appreciation through linguistic expression, and will be an issue for future study.

References

1. Chen, J. and Stokes, J. R. （2012）, Rheology and tribology: Two distinctive regimes of food texture sensation, Trends Food Sci. Tech., 25, 4-12
2. Fxyma, Hiroki. Clear is sweet: Defining aesthetic sake taste terms with a usage-based approach. *The Language of Food in Japanese: Cognitive perspectives and beyond*, Converging Evidence in Language and Communication Research, Kiyoko Toratani, ed.. John Benjamins.
3. Merleau-Ponty, Maurice. 2013. *Phenomenology of perception*. Routledge.

Authors Introduction

Hiroki Fxyma Ph.D.

Hiroki Fxyma (also spelled Hiroki Fukushima, or 福島宙輝) holds a Ph.D. in Philosophy from Keio University (2018). His main research interests are semiotics, cognitive linguistics, representation of taste. He worked as an Assistant Professor of Keio University (2015-2018) and Senior Lecturer in Kyushu Women's University. He is now an Assistant Professor of Tainan University of Technology.

Why is the Early Detection of Dementia Failed?

Yuki Hayashi
Chiba University,
1-33, Yayoi-cho, Inage-ku, Chiba-shi, Chiba, 263-8522, Japan
E-mail: yu@chiba-u.jp

Abstract

To prevent dementia, the early detection is important. However, it's often failed because it's difficult to articulate "dementia" as "dementia" verbally. In this paper, I analyzed semi-structured interview data stored in DIPEx-Japan, a database of personal experiences of health and illness, and discussed the difficulty of early detection of dementia.

Keywords: early detection of dementia, analysis of semi-structured interview data, illness narrative.

1. Introduction

It is said that early detection of dementia is important. For example, Honma[1] pointed out the importance of the early detection as follows.

- In case of Alzheimer's dementia, its progress will be delayed by Donepezil hydrochloride, which is a type of medicine. Therefore, a period of in-home care will be able to be extended.
- The right of self-determination will be respected. For example, the elderly people will be able to communicate with their family in regard to wealth management or nursing before they will suffer from several symptoms of dementia.
- QOL (Quality of Life) will be preserved. If care-givers (including family members) recognize the state of dementia in advance, they will be able to keep communicating with a person with dementia.

To detect dementia, several tools have been recommended. For example, fundamental check list is shown in Manual for preventing nursing[2]. According to this list, if an elderly person is often pointed out that he/she is forgetful, his/her cognitive functions will be suspected to decline. However, it is difficult to determine whether an elderly person is dementia or not. Generally, dementia is defined as ``the sustainable decline of intelligent abilities and the state which interfere with daily life and social life due to several types of cognitive disorders. We think early detection of dementia is difficult due to the ``sustainable" decline of intelligent abilities. In this paper, we will discuss the difficulty of early detection of dementia based on narrative data.

2. Dementia problem

Alzheimer's disease, Lewy body disease or cerebrovascular disease are known as the factor of dementia. However, an appearance of these disease is not always equal to the beginning of dementia. Deguchi[3] interpreted dementia ``not just as physical and somatic problem for the elderly people but as interactive problem due to communicative disorder between the elderly people and people concerned. So we regard dementia as communicative problem which is gradually articulated through communication with people concerned. Through an interview data, Deguchi illustrated the phases of dementia troubles as follows: Zero point of an unspecialized trouble, Variation of an interpretation or definition about the trouble, Confusing how to deal with the trouble, Regarding the trouble as a claim, Discussing how to deal with the trouble among people concerned, Determining how to deal with the trouble among people concerned.

To discover the beginning point of dementia trouble (``Zero point of an unspecialized trouble"), Deguchi interviewed the elderly people with dementia and their family. However, there were not so many samples so we'll analyze the other type of interview data in the next chapter.

Furthermore, in articulation and cognitive anomie hypothesis, Nakagawa[4] examined the process of dementia. According to his hypothesis, ``a claim on unspecialized trouble will refuse a process of normalization. Then the claim will gradually be articulated and be classified into deviation, social problem or the other type of problems. Finally, the factor and solution for the claim will also be articulated. In the following chapter, we will analyze communicative aspect of dementia and examine the difficulties of early detection of dementia.

3. Analysis

3.1. *Method*

We have analyzed semi-structured interview data stored in DIPEx-Japan[5]. DIPEx is a database of personal experiences of health and illness. It contains patient interviews data and the database is opened to public through the Internet. So patients, their family members and medical professionals can access the narrative. DIPEx-Japan has various theme of narrative. In each section, there are voice data (some interviewees hide their face) and text data like Fig. 1. In this chapter, we will focus on narrative about ``beginning of dementia".

3.2. *Results*

We examined how caregivers (including family members) noticed the beginning of dementia before some type of troubles are clearly classified into dementia. In the following analysis, characteristic phrases will be shown in several categories.

The first type of narrative is the narrative in which people didn't regard as dementia. There are some examples as follows (with Japanese translations).

- I didn't imagine the possibility of disease. (病気の予想なんかしてなかった。)
- I think our family had never considered the possibility of dementia. (認知症かどうかっていうような風にはみんな捉えてなかったんじゃないかと思う。)

In several cases, we observed the reason why people had not regarded as dementia.

- My mother had depression and saw her doctor regularly. So I was not sure whether she was dementia or not. (うつになって以来神経科とかにかかってて、それでずーっと来たので、その異変ていうのが、どこから異変ていうのがわからなくて。)
- I thought she looked strange due to her age. (年をとって、そういうふうになっているみたいな。)
- On that day, he might be not feeling very well. (その日は、体調が悪いのかなとか思って。)

Secondly, there are the narrative in which people pointed out some specific events they felt wrong or strange as follows (with Japanese translations).

Fig. 1. The construction of interview data. Upper side is voice data and lower side is text data (in Japanese).

- I think she often left her belongings. (置き忘れとかはね、結構あったと思う。)
- It took so long time for my husband to search his belongings. (あれどこいったんだろうって、探す時間が長くなって。)
- He sometimes forgot a password for ATM. (たまに暗証番号忘れてお金が下ろせなかったりとか。)
- He made a phone call to a certain person and did soon again to the same person. (さっき電話をかけたところにまたかけ直す、みたいなことがあって。)

These behaviours are nearly equal to the dementia symptoms which are known as typical ones. In addition, some people focused on the other type of behaviour.

- Her bahaviours, especially in the way of walking, looked wrong for me. (やっぱり挙動ですかねえ。歩き方がおかしいとか。)
- Her sleeping hours gradually increased. (睡眠時間がだんだん長くなってきたようなのを覚えています。)

As we mentioned, the beginning point of dementia is usually not clear. In fact, the following examples shows the aspect of dementia and a type of confusion for family members.

- I'm not sure when is the beginning of dementia. (いつごろ始まったかっていうのは、はっきりしない。)
- In those days, I was not sure whether my husband is dementia or not. (そのときは、やっぱり、分からなかった。)

However, there are the narrative in which people pointed out something wrong or strange as follows (with Japanese translations).

- A few years ago before my husband was diagnosed with dementia two years ago, I think something strange happened around him. (その診断の 2, 3 年前には、何かいろいろあったような気がする。)
- My mother was mature for her age. However, there were something strange with her. (しっかりした母親やったんですよ。ところが、何か様子がどうも。)
- When I visited my mother, I felt something wrong with her. (訪ねて行ったときに、どうもいつもと様子が違う感じを受けた。)

3.3. *Discussion*

Usually people didn't determine whether their family member was dementia or not because they compared their family with common people. In addition, it may be difficult for them to detect the differences between normal state and wrong state.

In the second type of narrative, people pointed out some specific events they felt wrong or strange. However, these behaviours are nearly equal to the dementia symptoms which are known as typical ones. So it is not so suitable for early detection of dementia. On the other hand, in the last type of narrative, people told they felt something wrong or strange. These narrative seem not to have significant meanings, but feelings which family members as caregiver may feel are important. We think these narrative are less influenced by cognitive bias.

4. Conclusion

For early detection of dementia, we regarded dementia as a type of communicative problem which was gradually articulated through communication with people concerned. From the viewpoint of that, we examined narrative data stored in DIPEx-Japan and classified characteristic phrases into several categories.

References

1. A. Honma, Early detection of senile dementia and future prospect in local cities, Gerontology New Horizon, vol. 15, pp. 20--23, 2003.
2. Revision committee on manual for preventing nursing, On preventing nursing, in Manual for preventing nursing, pp. 1--37, March 2012.
3. Y. Deguchi, Micro politics of troubles about senile dementia---retrospective narrative on family care-givers' trouble experiences---, Sociologist, faculty of sociology, Musashi University, vol. 1, pp. 39--75, 1999.
4. N. Nakagawa, Social problem game and researcher's game, the Bulletin of faculty of liberal arts, Toyama University, vol. 25, pp. 57--81, 1995.
5. DIPEx-Japan, Narrative on health and illness, referred on 27th April 2017.

Authors Introduction

Mr. Yuki Hayashi

He received his Master's degree from Graduate School of Humanities and Social Sciences, Chiba University, Japan in 2017. He is currently a Doctoral course student in Graduate School of Science and Engineering, Chiba University, Japan.

Relationship Between Worldview and Advertising Techniques

Yoji Kawamura
Faculty of Business Administration, Kindai University, 3-4-1 Kowakae
Higashi-osaka, Osaka 577-8502, Japan
E-mail: kawamura@bus.kindai.ac.jp
www.kindai.ac.jp

Abstract

This study clarifies the different worldview domains of advertising and the techniques used to construct them. First, the keywords related to worldview are extracted by analyzing case studies conveying the worldview of a brand. Next, based on the questionnaire survey on keywords, classification and systematization of the worldview domains factor analysis are attempted. Additionally, based on the questionnaire survey on advertising techniques that constitute the worldview, the relationship between the worldview domain and the advertising techniques is quantified by regression analysis. The 66 keywords were classified into eight domains by factor analysis. The total regression coefficients for the eight domains were found that "Characters and Product Movements," "Atmosphere," and "Product functions, Naming, and Logos" were influential in constructing the eight worldview domains. Regression analysis was used to quantify the relationship between the eight worldview domains and 11 advertising techniques. These relationships serve as rules when planning advertising creatives and contribute greatly to marketing automation.

Keywords: Worldview, Advertising Techniques, Relationship, Factor Analysis, Regression Analysis

1. Introduction

Worldview has been discussed in various academic fields such as philosophy, literature, political science, sociology, and business administration. In this study, the term worldview is considered from the advertising perspective in business administration. Worldview is described as "a comprehensive viewpoint of the world and human life." However, from the advertising perspective, worldview is often regarded as the viewpoint and style of the advertisement provider or the opinion and style of the receiver. Therefore, there is no exact definition of the term worldview.

This study clarifies the different worldview domains of advertising and the techniques used to construct them.

First, the keywords related to worldview are extracted by analyzing case studies conveying the worldview of brands. Next, based on the questionnaire survey on keywords, classification and systematization of the worldview domains factor analysis are attempted. Additionally, based on the questionnaire survey on advertising techniques that constitute the worldview, the relationship between the worldview domain and the advertising techniques is quantified by regression analysis.

2. Related Literature

Although many studies mention worldview, few clearly indicate its meaning and definition. Advertising practitioners highlight that it is important to consider worldview as "the power to present the story and the future" and to envision the worldview[1]. Moreover, the visual expression of art and design is a powerful tool to convey the worldview to others. Regarding the viewpoints and styles of those who provide advertisements, research is being conducted on the differences in the cultural plot patterns of creative teams[2], the portrayal of traditional cultural values in China[3], the analysis of characteristic elements of the expression styles of luxury brands[4,5], and the methods of corporate advertising and cultural branding[6]. Additionally, regarding the viewpoints and styles of the recipients of advertisements, research is being conducted on the gender differences in moral worldviews[7], cultural codes shaping lifestyle behaviors[8], the effectiveness of humor through the joint humor process of advertising and the cultural orientation of viewers[9], the impact of social context on consumers' neurophysiological responses to advertising messages[10], comparison of the persuasive power of narrative versus non-narrative advertising[11], and changes in the likelihood of co-creation with consumers according to the content creation style[12]. These studies do not systematically discuss the elements and domains of a unified worldview. However, they discuss the relation of the message, story, domain (culture, tradition, social context), atmosphere (luxury, humor), and method of expression to the views and styles of advertising providers and recipients.

In this study, the worldview domains of advertising techniques are extracted and classified to quantify the relationship between them.

3. Analysis Method

3.1. *Extraction of Keywords Related to Worldview*

Using introductory texts of the branding design samples[13] and the texts of essential design points as analysis materials, keywords related to worldview described in the texts such as "conveying ~," "~ness," "making ~ feel," and "expressing ~" were extracted.

3.2. *Questionnaire Survey on Worldview and Advertising Techniques*

As it is difficult to make a conclusion when a large number of keywords are factor-analyzed, keywords with similar meanings were aggregated into keyword groups. Questionnaire surveys were conducted on the keyword groups conveying a sense of the worldview and on the advertising techniques that constitute the worldview.

Based on the favorable factors of the Advertising Favorability Survey set by the CM Soken Consulting[14] and the analysis of the advertising techniques[15,16], the advertising techniques set included: advertising message, products, corporate attitude, story development, episode, characters and product movements, performers or character, background color, cutting tempo, music or sound, and atmosphere.

3.3. *Classification of Worldview Domains*

Factor analysis was conducted based on the response data concerning worldview keyword groups in the questionnaire survey. The statistical analysis software SPSS[17] was used for factor analysis, and factors (domains) were extracted by the Promax method with Kaiser normalization based on the principal axis factoring. Additionally, worldview domains were analyzed by gender based on the extracted factors.

3.4. *Relationship Between Worldview Domains and Advertising Techniques*

Based on the aggregated data on worldview domains and advertising techniques in the questionnaire survey, regression analysis was conducted with worldview domains as the dependent variable and advertising techniques as the independent variable. Regression analysis was performed using SPSS statistical analysis software, and the regression coefficient was calculated by selecting variables using the stepwise method.

4. Results and Analysis

4.1. *Extraction of Keywords Related to Worldview*

Sixty-six keywords were extracted based on 101 examples of branding design. Keywords with similar meanings were aggregated into 27 keyword groups, and these were used as questionnaire items. The keyword groups are shown in Table 1 below.

4.2. *Questionnaire Survey*

The questionnaire was administered to 903 participants (641 males and 262 females) aged 19–24 years. They were asked to respond on their worldview using a survey covering 27 keyword groups and the 11 advertising techniques.

4.3. *Classification of Worldview Domains*

The 27 keyword groups were classified into eight domains by factor analysis. The classification results are shown in Table 1. Although these do not cover all the domains, the main worldview domains related to advertising and store design were set and classified. The three worldview domains with the largest percentage indicated by the respondents were "Play/Novel" (36%), "Individuality/Surprise" (33%), and "Tradition/Foreign" (32%). Meanwhile, “Rural/Carefree” (24%), "Sincerity/Eternal" (19%), "Urban/Sophisticated" (16%), "Advanced/Future" (15%), and "Festival/Regional" (9%) were smaller. Fig. 1 shows the aggregated results by gender for each domain. “Play/Novel" was significant at the 5% level and was preferred more by females, while "Rural/Carefree" and "Festival/Region" were significant at the 5% level and was more preferred by males.

Table 1. The domains and keyword groups of worldview.

Domains	Keyword Groups
Play/Novel	Playful, Fun, Interesting
	Live feeling, Uplifting feeling, Tension feeling
	New sensations, Creation, Creativity
	Various, Mixed, Miscellaneous, Wide variety
Individuality/ Surprise	Enthusiasm, Commitment, Only yourself, Individuality
	Sharp identity, Upset common sense
	Hidden charm
Urban/ Sophisticated	Fashionable, Urban
	Extreme, Premium, Luxury
	Cool, Beautiful silhouette
	Sophisticated, The classic, Gentleman, Authentic
Advanced/ Future	Technology, Advancement, Intelligence
	Future, Possibilities
	Feeling unfinished, Feeling of effort
	Specialist, Professional
Tradition/ Foreign	Tradition, History, Not influenced by fashion
	Fusion of newness and history
	Make you feel exotic
Festival/ Regional	Festival, Celebrate life
	Contribution to the community
Sincerity/ Eternal	Sincerity, Honesty
	Eternal, Be loved forever
	Benefit for all three sides, Solidarity, Connection, Ties
Rural/ Carefree	Countryside scenery, Relaxing time, Natural
	Seasons, Color the life
	Casual, Comfortable, Carefree
	Living, Way of living

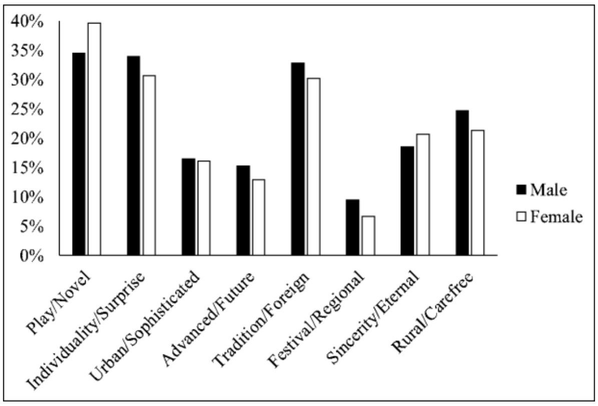


Fig. 1. The aggregated results by gender for each domain.

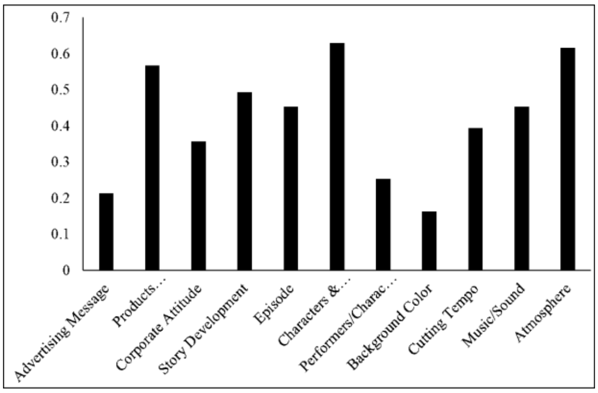


Fig. 2. The total regression coefficients for the eight domains.

4.4. *Relationship Between Worldview Domains and Advertising Techniques*

Regression analysis was used to quantify the relationship between the eight worldview domains and 11 advertising techniques. Table 2 shows the analysis results (regression coefficient at 5% significance). Fig. 2 shows the total regression coefficients for the eight domains. It was found that "Characters and Product Movements," "Atmosphere," and "Product functions, Naming, and Logos" were influential in constructing the eight worldview domains. However, for the trends of the main advertising techniques in each domain, "Music/Sound" and "Performers/Characters" affected "Play/Novel"; Atmosphere" and "Story Development" affected "Individuality/Surprise" and "Tradition/Foreign"; and "Background Color" and "Cutting Tempo" affected "Rural/Carefree".

Table 2. Relationship (regression coefficient) between worldview domains and advertising techniques.

Advertising Techniques	Worldview Domains: Play/ Novel	Individual ity/ Surprise	Urban/ Sophistica ted	Advanced /Future	Tradition /Foreign	Festival/ Regional	Sincerity /Eternal	Rural/ Carefree	Total
Advertising Message					0.073		0.071	0.069	0.213
Product func-tions, Naming, & Logos	0.149		0.181	0.153	0.084				0.567
Corporate Attitude		0.119				0.085	0.071	0.082	0.357
Story Development	0.104	0.136	0.07	0.094	0.088				0.492
Episode	0.071	0.093			0.071		0.138	0.08	0.453
Characters & Product Movements		0.099	0.115	0.117	0.075	0.105	0.118		0.629
Performers/ Character	0.151						0.102		0.253
Background Color	0.065							0.097	0.162
Cutting Tempo	0.074	0.091		0.134				0.095	0.394
Music/Sound	0.159	0.121	0.088		0.084				0.452
Atmosphere	0.088	0.156	0.106	0.071	0.1			0.094	0.615

5. Conclusion

In recent years, attempts have been made to automate marketing. Accordingly, effective communication is based on a worldview that is easily relatable to individual customers. In this study, classification of the worldview domain was attempted. These classifications can be used as indicators to determine individual customers' characteristics.

Additionally, a quantification of the relationship between worldview and advertising techniques was also attempted. These relationships serve as rules when planning advertising creatives and contribute greatly to marketing automation.

This research analyzed the 19–24 age group. However, an expanded age group and a comprehensive survey will become an effective indicator for determining the characteristics of individual customers. In the future, we can conduct surveys that cover a wider range of age groups to further substantiate advertising techniques.

Acknowledgements

This work was supported by Research Grants of Yoshida Hideo Memorial Foundation.

References

1. M. Mizuno, S. Yamaguchi, “Creating the World View: Business Skill of Sentience and Intelligence”, Asahi Shimbun Publications, Tokyo, Japan, 2020. (In Japanese)
2. G. V. Johar, M. B. Holbrook, B. B. Stern, “The Role of Myth in Creative Advertising Design: Theory, Process and Outcome”, Journal of Advertising, Vol.30, Issue 2, pp.1-25, 2001.
3. C. A. Lin, “Cultural Values Reflected in Chinese and American Television Advertising”, Journal of Advertising, Vol. 30, Issue 4, pp.83-94, 2001.
4. J. Kapferer, V. Bastien, “The Luxury Strategy: Break the Rules of Marketing to Build Luxury Brands (English 2nd Edition)”, Kogan Page, London, UK, 2012.
5. H. Gurzki, N. Schlatter, D. M. Woisetschläger, “Crafting Extraordinary Stories: Decoding Luxury Brand Communications”, Journal of Advertising, Vol. 48, Issue 4, pp.401-414, 2019.
6. B. Ertimur, G. Coskuner-Balli, “Brands Expressing Compassion and Care through Advertising”, Journal of Advertising, Vol. 50, Issue 3, pp.230-239, 2021.
7. F. F. Brunel, M. R. Nelson, “Explaining Gendered Responses to “Help-Self” and “Help-Others” Charity Ad Appeals: The Mediating Role of World-Views”, Journal of Advertising, Vol. 29, Issue 3, pp.15-28, 2000.

8. C. Rapaille, “The Culture Code: An Ingenious Way to Understand Why People Around the World Live and Buy as They Do”, Broadway Books, New York, NY, 2007.
9. Y. H. Lee, E. A. C. Lim, “What's Funny and What's Not: The Moderating Role of Cultural Orientation in Ad Humor”, Journal of Advertising, Vol. 37, Issue 2, pp.71-84, 2008.
10. R. Pozharliev, W. J.M.I. Verbeke, R. P. Bagozzi, “Social Consumer Neuroscience: Neurophysiological Measures of Advertising Effectiveness in a Social Context”, Journal of Advertising, Vol. 46, Issue 3, pp.351-362, 2017.
11. E. Kim, S. Ratneshwar, E. Thorson, “Why Narrative Ads Work: An Integrated Process Explanation”, Journal of Advertising, Vol. 46, Issue 2, pp.283-296, 2017.
12. I. Pentina, V. Guilloux, A. C. Micu, “Exploring Social Media Engagement Behaviors in the Context of Luxury Brands”, Journal of Advertising, Vol. 47, Issue 1, pp.55-69, 2018.
13. BNN Editorial Department, “New Logos and Tool Development: Design that Conveys the Brand's Worldview”, BNN, Tokyo, Japan, 2020. (In Japanese)
14. CM Soken Consulting, “Monthly Top 10 CM of New Works, CM INDEX, Vol. 36, No. 1, 2021. (In Japanese)
15. Y. Kawamura, “An Analysis on Commercial Film Techniques and Effects Using Positive Factors of Evaluation Database System”, Journal of Robotics, Networking and Artificial Life, Vol. 6, Issue 1, pp.1-6, June 2019.
16. Y. Kawamura, “An Analysis on Advertising Techniques to Gain the Consumer Engagement”, Research Abstracts of Research Grants of Yoshida Hideo Memorial Foundation, Vol. 54, pp.45-57, 2021. (In Japanese)
17. IBM, “IBM SPSS Statistics Base 25”, IBM, 2018.

Authors Introduction

Mr. Yoji Kawamura

He received his MBA from Graduate School of Business Science, University of Tsukuba, Japan in 1993. He is a Professor in the Faculty of Business Administration, Kindai University, Japan. His majoring is modeling the various social phenomena based on the modeling of consumer and the system building for simulation.

Adjective and Adjective Verb Conceptual Dictionaries in an Integrated Narrative Generation System

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, t-ogata@iwate-pu.ac.jp

Abstract

Although the main conceptual types in our previous narrative generation study using conceptual dictionaries were verb and noun concepts, other conceptual types are necessary for the implementation of more precise narrative generation functions. In this paper, we prepare frameworks for adjective and adjective verb concepts in our narrative generation system, an integrated narrative generation system (INGS). Furthermore, we provide antonyms for each adjective and adjective verb concept. This function will contribute to introducing contrasting rhetoric into narrative generation.

Keywords: Adjective Conceptual Dictionary, Adjective Verb Conceptual Dictionary, Antonym, Story Generation.

1. Introduction

We are developing a narrative generation system that organically integrates knowledge of narrative structures with artificial intelligence techniques. We call this an integrated narrative generation system (INGS)[1, 2]. Regarding knowledge of a narrative's structure, the INGS has a conceptual dictionary that is used to generate the narrative. Thus far, we have developed a noun conceptual dictionary and a verb conceptual dictionary for events that constitute a narrative [3]. However, in terms of word types, nouns and verbs are the minimum necessary components of a narrative. Therefore, we prepared a modifier conceptual dictionary containing adjectives, adjective verbs, and adverbial concepts as its components [4]. In this study, we focus on adjective and adjective verb concepts among the components of the modifier dictionary.

The purpose of this paper is to present the structure of the adjective and adjective verb conceptual dictionaries and study how to store and use antonym concepts.

2. Adjective and Adjective Verb Conceptual Dictionaries

This section describes the structure of the adjective and adjective verb conceptual dictionaries. Only a few tentative intermediate concepts are set up using intermediate concepts from the verb conceptual dictionary [4].

2.1. *Adjective conceptual dictionary*

There were 714 adjectives. Each adjective concept belongs to one of the following categories: attributes, possessions, relative relations, perceptual states, emotional states, thought states, physical states, and natural phenomena.

Individual adjectives follow the structure of verb concepts [3]. Each has a sentence pattern for single sentence generation, case structure, case constraints, and an intermediate concept to which the concept belongs. The case constraint expresses the concept that the adjective modifies, and one of the intermediate concepts in the noun conceptual dictionary is filled in. The example in Fig. 1 shows that “outdoor” is an adjective that indicates “bright.”

```
(set '明るい 1[bright] '((name 明るい 1[bright])
 (sentence-pattern “N1 が 明るい[N1 is bright]”)
 (case-cons-set
  ((case-frame ((agent nil) (counter-agent nil) (location nil)
(object N1) (instrument nil) (from nil) (to nil)))
    (constraint ((外[outdoor])))))
     (is-a (d 自然現象[natural phenomena]))))
```

Fig. 1. An example of an adjective concept.

2.1. *Adjective verb conceptual dictionary*

There were 1,191 adjectival concepts. Individual adjectival concepts belong to one of the following categories: existence, attributes, relative relations, perceptual states, emotional states, thought states, physical states, and natural phenomena.

Individual adjective concepts follow the structure of verb concepts [2] as well as that of adjective concepts. Each has a sentence pattern for single sentence generation, case structure, case constraints, and an intermediate concept to which the concept belongs. As with adjectives, the case constraint expresses the concept the adjective modifies, and one of the intermediate concepts in the noun conceptual dictionary is filled in. The example in Fig. 2 is an adjective that indicates that a “human,” an “animal,” or a “body part” is “robust.”

```
(set '頑健だ 1[robust] '((name 頑健だ 1[robust])
    (sentence-pattern “N1 が 頑健だ[N1 is robust]”)
    (case-cons-set
     ((case-frame ((agent nil) (counter-agent nil) (location
nil) (object nil) (instrument nil) (from nil) (to nil)))
     (constraint ((人[human] 動物[animal] 動物{部
分}[animal{part}]))))
       (is-a (d 身体状態[physical states]))))
```

Fig. 2. An example of an adjective verb concept.

3. Toward Defining Antonym Concepts

We propose a method of surveying antonym concepts for a given concept and incorporating their structures into conceptual dictionaries. To do so, we first investigate antonyms and then check the correspondence between the investigated antonyms and the concept dictionary.

1. **Investigating antonyms:** We consulted the *Weblio Antonyms / Synonyms Dictionary* (https://thesaurus.weblio.jp/antonym/), *Dictionary of Antonyms / Synonyms Online* (https ://taigigo.jitenon.jp/), *Large Dictionary of Antonyms* (https://hantaigo.com/), etc., to research and record the antonyms of the adjectives and adjective verbs.
2. **Comparison of synonyms and concept dictionaries:** Check if there are any adjectives or adjective concepts that match the investigated antonyms. If yes, create a relationship between the two concepts in the conceptual dictionaries. If there are no matches, register the concept of the investigated antonyms in the conceptual dictionaries and create a relationship between the two concepts.

The antonym concepts prepared using the above procedure were recorded on the list of adjective and adjective verb concepts, as shown in Fig. 3.

```
(set <Concepts> '((name <Name>)
 (sentence-pattern <Pattern>)
 (case-cons-set
  ((case-frame <Cases>) (constraint <Constraints>)))
    (is-a (<Concepts>))
    (antonym <Antonym concepts>)
   ))
```

Fig. 3. Adding antonym concepts to each adjective and adjective verb concept.

Next, we will provide an example of the utilization of the antonym concepts we have prepared here. We proposed a technique called coloring as a way to give a story a specific atmosphere [5]. This is a technique to change a story so that it has a specific atmosphere by adding words related to the desired atmosphere to the story.

One of the ways to realize coloring is to handle adjectives and adjective verbs. For example, to give the impression of blue, the words "refreshing" and "cool" were utilized. However, owing to the limitations of the color image scale, this technique only deals with words that have positive impressions. However, by using the antonymy relationship, it is possible to derive negative words from positive ones.

4. Conclusion

In this paper, we provided a detailed description of the adjective conceptual dictionary in the INGS. In addition, we discussed how to define antonym concepts corresponding to the two types of concepts, with the aim of applying them to story generation techniques.
In the future, we will continue to work on the definition of antonym concepts and attempt to conduct generation experiments using the expanded conceptual dictionary.

References

1. T. Ogata, "Toward an Integrated Approach to Narrative Generation: Emerging Research and Opportunities", IGI Global, 2020.
2. T. Ogata, "Internal and External Narrative Generation Based on Post-Narratology: Emerging Research and Opportunities", IGI Global, 2020.
3. T. Ogata, "Building Conceptual Dictionaries for an Integrated Narrative Generation System", Journal of Robotics, Networking and Artificial Life, Vol. 1, No. 4, pp. 270-284, 2015.
4. T. Ogata, J. Ono, Y. Herai, K. Igarashi, S. Sirai, "An extension plan of structuring and using of conceptual dictionaries and language notation dictionaries for an Integrated Narrative Generation System", Proceedings of the 48th Special Interest Group on Language Sense Processing Engineering, pp. 51-56, 2015. (In Japanese)
5. J. Ono, M. Kawai, T. Ogata, "Toward Love and Sex Narrative Generation Using a Noun Conceptual Dictionary", Journal of Future Robot Life, Vol. Pre-press, No. Pre-press, pp. 1-21, 2021.

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.

Prototyping Animation System that Combines a Kabuki Work and its Background Story: *Kyōganoko Musume Dōjōji* and the Legend of Dōjōji

Miku Kawai
Faculty of Liberal Arts, The Open University of Japan,
11 Wakaba nichome, Mihama-ku, Chiba, Chiba, 261-8586, Japan

Shunta Kudo
Faculty of Software and Information Science, Iwate Prefectural University,
152-52 Sugo, Takizawa, Iwate, 020-0693, Japan

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: mom.mxm39@gmail.com, g031q055@s.iwate-pu.ac.jp,
j.ono@aomori-u.ac.jp, t-ogata@iwate-pu.ac.jp

Abstract

Kyōganoko Musume Dōjōji, a famous *kabuki* dance work, is an adaptation of the legend of Dōjōji. A series of studies on *Kyōganoko Musume Dōjōji* included many themes, such as the analysis and simulation of the animation system of the stage performance structure, the survey and analysis of the relationship between *Kyōganoko Musume Dōjōji* and the legend of Dōjōji, and the design and experimental system development reflected in the above relationships. Based on these studies, we present an animation-based mechanism that flexibly associates the narrative flow of the stage performance structure with the story of the legend of Dōjōji.

Keywords: Animation System, *Kabuki* Dance, *Kyōganoko Musume Dōjōji*, Story Generation, The Legend of Dōjōji.

1. Introduction

Kyōganoko Musume Dōjōji is a *kabuki* dance with the legend of Dōjōji as its background. The legend of Dōjōji is a classic Japanese story about a male monk who lies to a woman, who ultimately burns him to death.

We used Watanabe's analysis of *Kyōganoko Musume Dōjōji* [1] as the basis for our study of *Kyōganoko Musume Dōjōji* and the legend of Dōjōji. First, by focusing on *Kyōganoko Musume Dōjōji,* we summarized basic knowledge (history and story content) of *Kyōganoko Musume Dōjōji*. In addition, we analyzed the stage performance structure of *Kyōganoko Musume Dōjōji* and the legend of Dōjōji by looking at actual video footage [2]. Moreover, based on the analysis table of the stage performance structure of *Kyōganoko Musume Dōjōji*, we created a two-dimensional (2D) animation system [3, 4].

On the other hand, as a presentation focusing on the legend of Dōjōji, we analyzed scenes from *Konjaku Monogatari* [5]. We classified each scene of *Kyōganoko*

Musume Dōjōji and the legend of Dōjōji into positive and negative and made associations. Furthermore, based on the results of the association, a three-dimensional (3D) animation was created and added to a previously performed 2D animation system [6, 7].

The goal of this study is to create a 3D animation that represents the legend of Dōjōji and implement an animation system that links the two, as proposed in the above study [3, 4, 6, 7].

2. Background: Combining *Kyōganoko Musume Dōjōji* and the Legend of Dōjōji

2.1 A method for combining the two works

Although the stories told in *Kyōganoko Musume Dōjōji* and the legend of Dōjōji are different, several scenes can be associated with both stories. For example, in the final scene of *Kyōganoko Musume Dōjōji*, Hanako climbs a bell and transforms into a snake. This is related to the scene from the legend of Dōjōji where Kiyohime burns Anchin to death. Therefore, we first correlated the scene in *Kyōganoko Musume Dōjōji* with the scene in the legend of Dōjōji [5].

We have previously analyzed the stage performance structure of *Kyōganoko Musume Dōjōji* [2]. In this study, we focused on three parts, namely 心(*kokoro*, heart), 歌詞(*kashi*, lyrics), and 振り(*furi*, performance), and found that *Kokoro* was a positive scene associated with “daughter,” while Kashi contained negative words such as “resentment.” Therefore, we classified each scene from *Kyōganoko Musume Dōjōji* and the legend of Dōjōji as either negative or positive. We then re-associated the negative and positive scenes based on these categorizations, as explained in the previous paragraph [7]. Based on the evaluation results, we propose a combination of *Kyōganoko Musume Dōjōji* and the legend of Dōjōji.

2.2 Making animations for the two works

While prototyping the 2D animation, we created 2D animations for all 11 scenes of *Kyōganoko Musume Dōjōji* [4]. Fig. 1 shows an example of a 2D animation. While prototyping the 3D animation, we created animations for 6 of the 33 events in the legend of Dōjōji [5]. Fig. 2 shows an example of 3D animation.

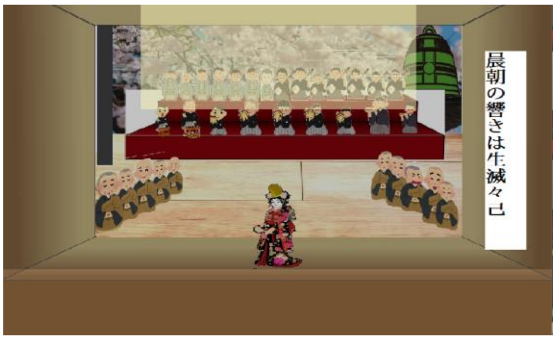


Fig. 1. An example of a 2D animation.

Fig. 2. An example of a 3D animation.

3. Implementing a Combined Animation System

In this study, we first added a 3D animation to the legend of Dōjōji. We created 3D animations of 27 events.

Next, we developed a system that combined 2D and 3D animations. We used *Kyōganoko Musume Dōjōji* as the main animation. As a result, the 2D animation of *Kyōganoko Musume Dōjōji* was replayed in chronological order, and the animation of the legend of Dōjōji was played in fragments.

The system algorithm is as follows: One time unit in a 2D animation was treated as one event. This procedure was repeated until all events were represented.

1. Play one 2D animation unit.
2. Refer to the evaluation result of the 2D animation unit.
3. Refer to the list of 3D animations corresponding to the evaluation result.
4. Play earlier 3D animations that have not yet been used in chronological order.

4. Conclusion

This paper proposed a system for combining animations of *Kyōganoko Musume Dōjōji* and the legend of Dōjōji, and examines the method for combining them. The system inserts a 3D animation of the legend of Dōjōji into a 2D animation of *Kyōganoko Musume Dōjōji*, thus providing the user with a story about the legend of Dōjōji, which is the background for *Kyōganoko Musume Dōjōji*. As a possible extension of this system, parameter-based selection of the explanatory content could be applied [8].

In the future, we plan to evaluate the system's generation results and investigate a more suitable coupling method.

References

1. T. Watanabe, "Musume Dōjōji", Shinshindō, Tokyo, 1986.
2. M. Kawai, T. Ogata, "An Analysis of Kyoganoko Musume Dojoji for the Generation of the Staging Structures of *Kabuki* from Narrative Plots", Proc. of the 63th Special Interest Group on Language Sense processing Engineering, No. 63, pp. 25-61, 2019.
3. M. Kawai, J. Ono, T. Ogata, "The Stage Structure of a *Kabuki*-dance: Survey, Analysis, and Representation of *Kyōganoko Musume Dōjōji*", The 37th Annual Meeting of Japanese Cognitive Science Society, P-117, pp.720-729, 2020.
4. M. Kawai, J. Ono, T. Ogata, "Analyzing the Relationship between the Legend of Dōjōji and the *Kabuki*-Dance *Kyōganoko Musume Dōjōji* to Develop Prototyping Systems", Journal of Robotics, Networking and Artificial Life, Vol. 8, No. 3, pp. 151-160, 2021.
5. M. Kawai, J. Ono, T. Ogata, "Analysis and Generation of the Narrative Elements in a "Dōjōji" Story", Proceedings of the 3rd International Conference on Computational Intelligence and Intelligent Systems, pp. 44-45, 2020.
6. M. Kawai, J. Ono, T. Ogata, "Towards a system that combines *Kyōganoko Musume Dōjōji* and the legend of Dōjōji", Proc. of the 65th Special Interest Group on Language Sense processing Engineering, No.65, pp. 13-24, 2021.
7. M. Kawai, J. Ono, T. Ogata, "Prototyping narrative representation system using a *Kabuki* dance and legendary story for the narration function of robots", Journal of Future Robot Life, Vol. Pre-press, No. Pre-press, pp. 1-35, 2021.
8. M. Kawai, J. Ono, T. Ogata, "*Kabuki* Explanation System Based on User's Knowledge and Interests", Proceedings of the 4th International Conference on Computational Intelligence and Intelligent Systems. (In printing)

Authors Introduction

Ms. Miku Kawai

Miku Kawai was an undergraduate student at the Faculty of Software and Information Science, Iwate Prefectural University from 2018 to 2020. She has been an undergraduate student at the Faculty of Liberal Arts, The Open University of Japan since 2021.

Mr. Shunta Kudo

Shunta Kudo has been an undergraduate student at the Faculty of Software and Information Science, Iwate Prefectural University since 2019.

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.

How Will Art Appreciations Change According to Information Change?

Akinori Abe*
Fuculty of Letters, Chiba University, Address
Chiba 263-8522, Japan†
E-mail: ave@ultimaVI.arc.net.my, ave@chiba-u.jp

Abstract

For the art appreciation in museums, usually a certain information will be provided as a caption. Visitors usually read the description to help his/her understanding. Thus, such a help will be necessary for ordinal person's understanding. Previously, we conducted an experiment to determine viewers' artwork understandings by gradually adding information in caption. In the previous experiment, we offered information only with the official information. In this paper, I will generate information with artists' own writings (artist's explanation). In addition, in this experiment, we used a rather abstract painting and representational but rather strange and difficult to understand paintings. In this paper, I will show how viewers' understanding or interpretation of artworks, and interest and preference of artworks change according to the changing information.

Keywords: art appreciation, information, change of bahaviour, understand.

1. Introduction

For the art appreciation in museums, usually a certain information will be provided as a caption. Visitors usually read the description to help his/her understanding. Thus, such a help will be necessary for ordinal person's understanding.

We have been studying the way of art appreciation ([1], [2], [3], [5]). For the study we have conducted several experiments. Several results could be obtained. For instance, in [5], we have obtained such results that museum visitors' impressions to artworks have changed according as the contents written in captions are changed. That is, frequent museum visitors' comments were roughly categorised into two types which were the report of no change ("No change on impression caused by the caption.") and the report about finding about the artwork ("I see a girl on the artwork after reading the caption."). Non-frequent museum visitors' answers were wider spread. Non-frequent museum visitors reported no change, findings of the artwork, impression changes and complaints about not enough information. This suggests that non-frequent museum visitors are highly dependent on the information provided by captions. For [3], we have obtained such results that for the sense of value, our hypothesis (For a positive information, value and preference of the artwork will increase. For a negative information, value and preference of the artwork will decrease.) is supported on a positive information. For a certain artwork, in fact, the value of the artwork increased. However participants' responses are normal. For the other artwork, the value of the artwork increased because of the technique. But the preference of the artwork did not increase. Because most of participants do not like a moth and so on.

In addition, in [1], we conducted an experiment to determine viewers' artwork understandings by gradually adding information in caption. In the previous experiments [1], [2], [3], we mainly offered information only with the official information.

In this paper, I will generate information with artists' own writings (artist's explanation). Then I conducted the similar experiments to obtain viewers' change of the behaviour to artworks.

2. Experiment

I conducted the experiment to obtain viewers' change of the behaviour to artworks according to the offered information.

In addition, I conducted an experiment in order to determine which factor (information) will change the viewers' preference, level of understanding, and interst of artworks in the art appreciation.

2.1. *Participants*

Participants were 22 (two are collected in the other day in my laboratory) adults including mainly university students and rather aged persons and their age were between 19 to 57 years old. Number of females was 8 and that of males was 14. Their unaided visions (visions without glasses) or corrected visions (visions wit glasses) were normal. All the participants were asked to answer about their art educational background.

Fig. 1. Experiment

2.2. *Stimuli*

The experiment was conducted in a lecture room in Chuo University on May 29, 2021. The room is usually used for the lecturing. So, the situation was rather different from those of the previous experiments. I used 3 artworks created in from 2018 to 2021 as stimuli. The artworks placed on the table or the chalkrail of the white board.

I selected 3 rather new artworks. The artists are in younger generation. One is an abstract artwork. The others seem representational artworks but rather difficult to understand. All artworks were numbered but were displayed without any captions and labels. The artworks are as followings;

1) Chikako Kai (甲斐千香子): Lunch (お昼ご飯) (2021) : seems representational artwork
2) Yuniko Kawamoto (川本悠肖子): Protect (守) (2020) : seems representational artwork
3) Yugo Kohrogi (興梠優護): \ 13 (2018): abstract artwork

2.3. *Method*

All participants answered the questions on the worksheet during appreciating the artwork. The questions were:

· Please write your thinking and feeling about this artwork, and stories in the artwork freely.
· 1.1 Do you like this artwork?
In a SD method style (5).
· 1.2 Why do you evaluate so?
· 2.1 How much can you understand the contents of the artwork?
In a SD method style (5).
· 2.2 Why do you evaluate so?
· 3.1 How deep are you interested in the artwork?
In a SD method style (5).
· 3.2 Why do you evaluate so?

In the next stage, I offered information gradually. The first information was the artist's name and biography, and a title of the artwork. For an artist I offered the artist's concept of drawing. The second information was the artist's own explanation about the artwork. After offering certain information, participants were asked the similar questions. In addition, I added the additional question.

· 4. If you have different feeling or thinking, please write down. For instance, about the contents of the artworks or story.

For the limitation of the time (about 30 minutes), I asked the differential question (such as Q4). And this was our main objective to know if the feeling or thinking changed according as the information change.

2.4. *Offered information*

In the previous experiments, we created information based on information in artists' homepage or information from galleries. In this experiment, for the artists' profile and concent, I created information based on information in artists' homepage or information from galleries. However, for the second information, I used information to interview the artist or artist's writing for the artwork (included in the artwork).

As a first information, I gave profiles of the artists.

The followings are the part of the second information[‡]

1) Chikako Kai (甲斐千香子): Lunch (お昼ご飯)
I felt the hidden repellence of lining up neatly and obeying the order when I was student. Accordingly I expressed the repellence as a porcelain bowl.

[‡] The original descriptions are written in Japanese. I translated them into English.

Where I put mayonnaise on the finger dolls then eat the rice with destroying them.

2) Yuniko Kawamoto (川本悠肖子): Protect (守)
   I grew vegetables with my friend in the share-field. But I have not be there because the coronavirus crisis. Accordingly I do not travel on a train to the field, and I do not firm there, so I have no time to blow up my image. [...]
   When I create Japanese a Japanese-style painting, I tried to search a subject matter from my memory to remind that I left a work glove on the vegetable's prop. [...]
   I drew the painting with my desire that the work glove guard the vegetables instead of me until when I will be able to go there.
3) Yugo Kohrogi (興梠優護): \13
   This painting belongs to the series of abstract paintings since 2018. [...]
   For this series of paintings, I focus on the abstract aspect and remove concrete aspects as possible as I can.
   For the question that what image did I have in my head when I drew the painting, I can say I thought of the pareidolia (シュミラクラ現象). I'm interested in such a brain's cognitive effect. I use the phenomena that the figure seems something intentionally to introduce the viewer's eyes to something (It is called movement in the school of painting.). I create paintings in the series with thinking such cognitive effects. [...]
   The pictorial movement has the most important meaning in my artworks. In my painting, if the drawn figure is deformed and shaking, viewers unconsciously follow the brushstrokes and colour in order to correct the deformation. Then into a painting which should be a still image we can embed moving element. [...]

Only for the first artist, in addition to the profile, I gave a concept of artist such that "I painted simething strange with things close to me.

3. Results and discussions

I have collected answers what participants wrote during the experiment. Several interesting results were obtained.

3.1. *Results on SD scores*

I review the result from the viewpoint of SD scores.

1) Artwork1:

Difference after first information

About the preference, very few participants added the score. About the interest, very few participants added and reduced the score. Very interesting phenomenon is that several participants added and reduced the score (from -3 to +3) about the understanding.

Difference after second information

For the all points, several participants added and reduced the score.

2) Artwork2:

Difference after first information

About the preference, very few participants added and reduced the score. About the interest, few participants added and reduced the score. About the understanding, most participants did not change their mind.

Difference after second information

About the understanding, several participants added the score. About the interest, few participants added and reduced the score. Very interesting phenomenon is that half of the participants added and reduced the score (from -3 to +3) about the preference.

3) Artwork3:

Difference after first information

About the preference, very few participants added and reduced the score. About the understanding, very few participants added and reduced the score. About the interest, very few participants added and reduced the score. Actually one participant reduced the score by 4.

Difference after second information

About the preference, very few participants added and reduced the score. About the understanding, half of the participants added the score. Actually three participants added the score by 4. About the interest, several participants added and reduced the score.

3.2. *Discussions*

In the previous experiments [2], we observed that for the preference, it was rather difficult to change the mind. However, if the artist's lovable comments were obtained, we tended to like the artworks more.

For the preference, we obtained rather different results than the previous experiments.

For the preference, in most cases, very few participants added or reduced the score. That is, they did not change their preferences. However, for the artwork2, after showing the second information, half of the participants added and reduced the score (from -3 to +3) about the preference. The comments from the participants who reduced the score were that "since artist's thought was rather different from mine..." and "since artist's image and mine was almost the same, and there was not bland new matter, the painting was rather boring." The reasons were extreme opposite. It seems that even the artist's thinking is the same as participants' thinking, they thougt that the image was not new. That is, they need the bland new concept. On the other hand, since the artist's thinking is different from participant's thinking, they tend to dislike the artwork. This may be also natural situation.

One of the comments from the participants who added the score was that "I felt the farmer's current situation, loneliness and forlornness under the COVID-19 pandemic..." It seems that the since participants felt the deep concept, they like the artwork more.

About the understanding, for the artwork3, half of the participants added the score. Actually three participants added the score by 4. Actually, this painting is a typical abstract painting and very difficult to understand what is painted. The artist's comment was an answer from my question. Accordingly it is very precise writing and included his concept and techniques. This may be the reason for the high score of understanding.

4. Conclusions

In this paper, I showed the result from the experiment conducted on May 29, 2021.

I analysed the result only from the viewpoint of SD scores. However, I could obtain a certain interesting phenomenon.

For the preference, in most cases, very few participants added or reduced the score. That is, they did not change their preferences. However, for the artwork2, after showing the second information, half of the participants added and reduced the score (from -3 to +3) about the preference. In this case, artist's image gives both good and bad influence to the viewers' mind. I can understand both minds and this situation is very interesting.

About the understanding, for the artwork3, half of the participants added the score. Actually three participants added the score by 4. The artist's comment was a very precise writing and included his concept and techniques. This may be the reason for the high score of understanding.

In the next paper, I will analyse participants' comment to understand the process of the art appreciation.

References

1. Akinori Abe, Kotone Tadaki, ``Captions with Several Levels of Explanation, Proc. of KES2019, 2335--2344, 2019.
2. Akinori Abe, Kotaro Fukushima, 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, 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. Helmut Leder, Claus-Christian Carbon, Ai-Leen Ripsas, ``Entitling art: Influence of title information on understanding and appreciation of paintings," Acta Psychologica, 121, pp. 176--198, 2006.
5. Kotone Tadaki, Akinori Abe, ``Museum Visitors' Behavioral Change Caused by Captions," Proc. of the 2nd. Int'l Workshop on Language Sense on Computer in IJCAI2017, pp. 53--58, 2017.

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.

Research on an AGV path planning method

Zhihui Chen[1], Xiaoyan Chen*, Shiming Wang

Tianjin University of Science and Technology, China;

E-mail: 1594838831@qq.com

www.tust.edu.cn

Abstract

AGV is an acronym for Automatic Guided Transport Vehicle. At present, the key technologies of AGV mainly include navigation and positioning technology, path planning technology, multi-AGV coordinated control technology and multi-sensor information fusion technology, etc. This paper studies A* algorithm in path planning. The A* algorithm can find the shortest path between two points. It mainly studies the principle of A* algorithm and simulates it in Matlab. By comparing the path length of ant colony algorithm and A* algorithm in Matlab grid graph, the excellent performance of A* algorithm is proved.

Keywords:AGV, Path planning, Ant colony algorithm, A algorithm*

1. AGV Introduction

AGV (Automated Guided Vehicle) refers to a transport vehicle equipped with electromagnetic or optical automatic navigation devices, capable of driving along a prescribed navigation path, with safety protection and various transfer functions.

AGV can be divided into two types according to the guidance method. One is the guidance of a fixed path, and the other is the guidance of a free path. AGV can also be divided into loading type, assembly, towing, fork and suspension types according to the applications. The earliest AGV was transformed from a tractor-type excavator in 1953 by Barrett Electronics Company of the United States. It can be automatically and conveniently connected to the logistics system, which improves production efficiency to a certain extent. Until the beginning of the 21st century, some Western countries have developed AGV systems with network guidance, laser guidance technology and composite guidance technology.

2. AGV Path Planning

Path planning means that the mobile robot searches for an optimal or suboptimal path from the starting position to the target position according to a certain performance index (such as distance, time, etc.)[1]. According to the degree of grasp of environmental information, path planning can be divided into global path planning based on prior information and local path planning based on sensor information.

Global path planning includes environmental modeling method, probabilistic road map method, Dijkstra algorithm, A* algorithm (A-star algorithm) and fast search random tree method.

Local path planning mainly includes artificial potential field method, genetic algorithm, and DWA (Dynamic Window Algorithm).

3. A* Algorithm

3.1. *A* algorithm formula*

The A* algorithm is essentially a heuristic algorithm[2]. The A* algorithm combines the strengths of the breadth first search algorithm and the Dijkstra algorithm. According to the evaluation function formula, evaluate each position that the robot can reach, and find the optimal position from it. Then the optimal path can be found. The A* algorithm is commonly used in npc moves in various games.

$$F(\mathrm{n}) = G(n) + H(n) \tag{1}$$

In Eq.(1), F(n) is a function to calculate the cost from the starting point to the ending point. G(n) represents the actual cost from the starting node to the node where it is. H(n) represents the estimated cost from the node where it is to the end. The key to the shortest path is to choose the cost function, so it is considered the most important part. In the A* algorithm, because the Manhattan distance is relatively simple, the Manhattan distance formula is usually used as the distance estimation algorithm.

3.2. *A* algorithm specific steps*

Simplify the search area into a set of quantifiable nodes.

STEP1: Put the starting point S into the Open list as the node to be viewed.

STEP2: Search for nodes in eight directions near the starting point S (except for obstacles), put these nodes in the Open list, and then make S node the parent node.

STEP3: Move the starting point S from the Open list to the Close list.

STEP4: Compare the node X with the lowest F value in the current Open list, and move the X node from the Open list to the Close list.

STEP5: Search for all reachable nodes around X node.

(1) If these nearby searched nodes are still not in the Open list, add these nodes to it, calculate their F value, and then set the parent node of these nodes.

(2) If these nearby searched nodes are in the Open list, calculate the path from S to the target node through X point. Determine whether F is higher or lower than before.

If the F value is higher than before, it means that the cost of the new path is higher, and there is no need to change it.

If the F value is lower than before, change the parent node to square X, and calculate the F value of the nodes around the X node again, and repeat the cycle

STEP6: Repeat the above search steps until the target node is found

STEP7: End judgment. Finally, if the target node appears in the Open list, it means that a suitable path has been found.The algorithm flow chart is shown in the Fig1.

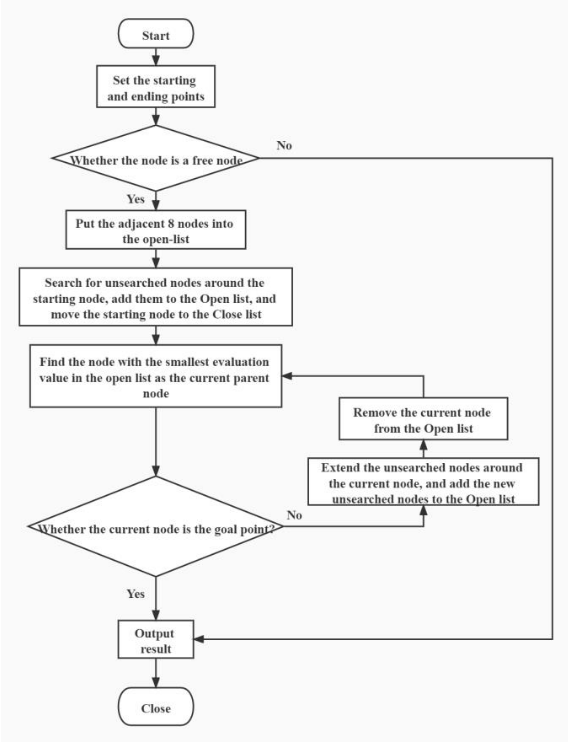


Fig.1. Flow chart of A* algorithm

4. Simulation on Matlab

In this paper, the A* algorithm is simulated on Matlab. Because all the researches in this paper do not need to consider the height of obstacles on a two-dimensional plane, the grid method is adopted for modeling. If there is an obstacle, it is "1" and if there is no obstacle, it is "0".As shown in Eq.(2).

$$\mathrm{cell}(x, y) = \begin{cases} 1 & \text{occupied grid} \\ 0 & \text{free grid} \end{cases} \tag{2}$$

According to the designed two-dimensional space environment, it is ensured that the robot can move freely in the preset environment map. Use Cartesian coordinate system to express the designed workspace in Matlab.

Construct a grid map with an environment of 20×20 and randomly generate 160 obstacles in the map, fix (1, 1) as the starting point, (19, 19) as the end point and use A* algorithm to simulate in Matlab. The simulation results are shown in Fig 2.

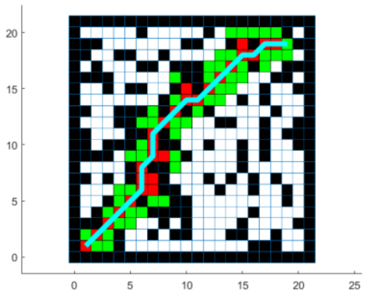


Fig.2. Simulation example of A* algorithm

Next, in order to compare the advantages of the A* algorithm. Compare the path length between A* algorithm and ant colony algorithm.

AG (Ant Colony Algorithm) is an optimization algorithm that simulates ants looking for food[3]. The principle of the ant colony algorithm is that each ant will secrete pheromone inversely proportional to the length of the path on the road on which it has crawled after exiting the hole to tell the subsequent ants. The pathfinding ends when they search for the target node. In the end, what the ant colony finds is the optimal foraging path.

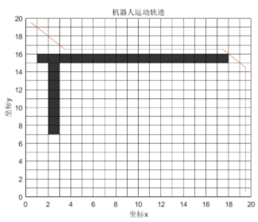
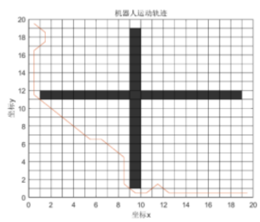
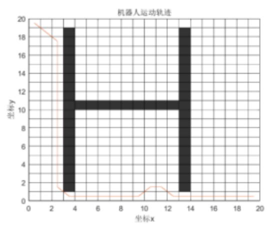
In the Matlab simulation, the basic parameters of the ant colony algorithm should be set first, including the pheromone intensity Q and the evaporation coefficient ρ. This paper adopts the ant colony algorithm to set the number of iterations as K=100; the number of ants M=50; the importance of pheromone Alpha=1; the importance of heuristic factor Beta=7; the pheromone evaporation coefficient Rho=0.3; the pheromone increase intensity coefficient Q =1. The parameter settings are shown in Fig 3.

```
Tau=ones(MM*MM,MM*MM);      % Tau initial pheromone matrix
Tau=8.*Tau;
K=100;                      %Number of iterations
M=50;                       %Number of ants
S=1;                        %Starting point of the shortest path
E=MM*MM;                    %Destination point of the shortest path
Alpha=1;                    % Alpha Parameter indicating the importance of pheromone
Beta=7;                     % Beta A parameter indicating the importance of the heuristic factor
Rho=0.3 ;                   % Rho Pheromone evaporation coefficient
Q=1;                        % Q Pheromone increase intensity coefficient
```

Fig.3. Ant Colony Algorithm Parameters

Establish two identical obstacle maps in the grid map and then use the A* algorithm and the ant colony algorithm to simulate the results as shown in the table 1and table 2 below.

Table 1 Simulation results

	Model 1	Model 2	Model 3
Ant Colony Algorithm			
A* Algorithm			

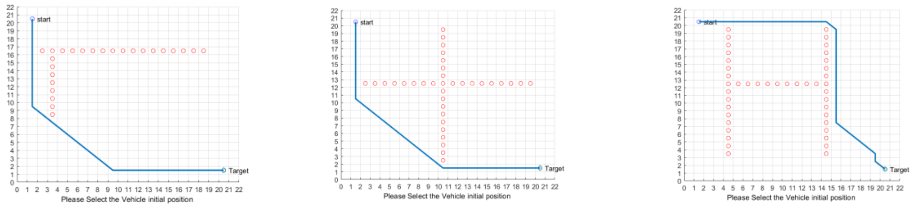


Table 2 Length comparison

Algorithm	Ant colony	A*	Short path algorithm
Model 1	37.272	33.3137	A*
Model 2	38.555	32.7279	A*
Model 3	38.772	34.4853	A*

The three sets of experimental data all show that the shortest path of the A* algorithm is shorter than that of the ant colony algorithm, so the A* algorithm is better than the ant colony algorithm.

5. Conclusion

The key issue of intelligent robot research all over the world is the path planning problem of mobile robot from beginning to end. The A* algorithm is already one of the commonly used algorithms for global path planning of mobile robots, and the A* algorithm is also effective in scientific and technological applications. But the A* algorithm still has many problems. Firstly, the current research of A* algorithm is mostly an ideal flat two-dimensional plane, and obstacles can only go around. In reality, this is not always the case[4]. When encountering obstacles that can overcome the past, the A* algorithm is not so good. Therefore, it is necessary to improve the A* algorithm so that the A* algorithm can be applied in more practical scenarios. Secondly, in order to prevent the A* algorithm from falling into the local optimal solution problem, a large number of search nodes will be carried out, and there will be some unrelated nodes among these nodes. Therefore, the balance between the optimality of the algorithm path and the number of irrelevant search nodes searched in the search process is a problem to be studied and solved. Thirdly, although the performance of the A* algorithm is relatively good, many scenarios have higher requirements for problem solving. Sometimes the A* algorithm alone cannot solve the problem particularly well[5], so it is necessary to integrate the A* algorithm with other algorithms. Solve more practical problems better.

References

1. C. Wang and J. Mao, "Summary of AGV Path Planning," 2019 3rd International Conference on Electronic Information Technology and Computer Engineering (EITCE), 2019, pp. 332-335.
2. C. Ju, Q. Luo and X. Yan, "Path Planning Using an Improved A-star Algorithm," 2020 11th International Conference on Prognostics and System Health Management (PHM-2020 Jinan), 2020, pp. 23-26. 2.
3. Q. Su, W. Yu and J. Liu, "Mobile Robot Path Planning Based on Improved Ant Colony Algorithm," 2021 Asia-Pacific Conference on Communications Technology and Computer Science (ACCTCS), 2021, pp. 220-224.
4. A. Candra, M. A. Budiman and K. Hartanto, "Dijkstra's and A-Star in Finding the Shortest Path: a Tutorial," 2020 International Conference on Data Science, Artificial Intelligence, and Business Analytics (DATABIA), 2020, pp. 28-32.
5. Xiang Liu and Daoxiong Gong, "A comparative study of A-star algorithms for search and rescue in perfect maze," 2011 International Conference on Electric Information and Control Engineering, 2011, pp. 24-27. : .

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. Dr. Xiaoyan Chen

She is 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. Shiming Wang

He graduated from Harbin Institute of Technology with a Ph.D in Control Science and Engineering, and is now the Director of the Automation Department of Tianjin University of Science and Technology.

Boiler level measurement and control system

Xiwen Liang, Shengli Sun, Xiaoyan Chen*

Tianjin University of Science and Technology ,300222， China;
E-mail:1540227453@qq.com

www.tust.edu.cn

Abstract

This paper develops a boiler level control system to measure and control boiler level easily. This level control system applies PID control algorithm in SIEMENS S7-300PLC, uses ladder diagram and statement table programming, and completes variable connection and screen editing in MCGS monitoring system. The PLC boiler level control system is completed by establishing the system mathematical model, designing the level control algorithm, and verifying the simulation. Through debugging, modifying parts of the parameters, the control effect of the control system achieve satisfied expectation.

Keywords: Automatic control, liquid level, PLC, MCGS

1. Introduction

Boiler as a kind of power energy equipment has been widely used since the 18th century. The main monitoring parameters of the boiler control system are liquid level, pressure and temperature, etc. For complex ones, it is also necessary to consider the air-to-coal ratio and water make-up capacity. These things have an important impact on boiler safety, stability and environmental protection. With the continuous development of intelligent control, the boiler liquid level control is becoming more and more perfect, but still faces the following two problems. 1) The control system is easy to diverge and oscillate. 2) The economy of the control system is very poor. The boiler liquid level control system designed in this paper adopts the PID control algorithm of Siemens S7-300PLC, and completes the variable connection and screen editing in the MCGS monitoring system. Through experimental verification, good results have been achieved.

2. Boiler Level Analysis

The control object of the boiler is a steam boiler, and its structure is similar to that of a single-capacity water tank in an ordinary boiler. The structural frame is shown in Fig. 1.

As shown in Fig 1, the input variable of this object is water supply, and the control variable is liquid level.

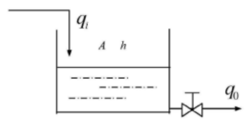


Fig.1. The composition of the single tank object

According to the material balance, the following relationship exists between input variables and control variables ($\frac{dV}{dt}$=$q_i - q_0$) .Since V=A*H gets A$\frac{dh}{dt}$=$q_i - q_0$. Since the outlet flow can be approximately expressed as

$q=\frac{h}{R}$ gets $A\frac{dh}{dt}=q_i-\frac{h}{R}$. The overall input and output model is Formula 1.

$$T\frac{dh}{dt}+h=Kq_i(T=AR,K=R) \quad (1)$$

Where $h=h_0+\Delta h$, $q_i=q_{i0}+\Delta q_i$. ($h0$, $qi0$ is the value of the balance state). Because of $h_0=0$. The input and output model of data changes is shown in formula 2.

$$T\frac{d\Delta h}{dt}+\Delta h=K\Delta q_i \quad (2)$$

In the control system, we often adopt the latter, and do not use special letters to indicate this amount of change. The result of Laplace transform[1] on the above formula is as formula 3.

$$TsH(s)+H(s)=KQ_i(s) \quad (3)$$

This transfer function itself is the transfer function of a step object. If an object input function is equal to the magnitude of the amplitude value and the function is the transfer step object input equal to a, then the transfer step input response of the step object is $Q_i(s)=\frac{a}{s}$. The step response function formula of a single tank is formula 4.

$$h(t)=L^{-1}\left[\frac{K}{s(Ts+1)}\right]=L^{-1}\left[\frac{Ka}{s}-\frac{KaT}{Ts+1}\right]=Ka\times L^{-1}\left[\frac{1}{s}-\frac{T}{Ts+1}\right]=Ka(1-e^{\frac{t}{T}}) \quad (4)$$

The step response of a single tank is shown in Fig. 2.

The response curve of the boiler liquid level during water supply is shown in Fig. 3.

If it is a simple water tank, we add water to the water tank. If the water supply is constant, the water level should rise linearly with a certain slope, as shown in Figure H1. However, the boiler is a heating system. When water enters the steam drum, the liquid level of the boiler will be lowered. When the water vapor encounters condensation, the density of the water increases as the temperature drops. We use pressure sensors, and the actual height of the high-density liquid level will be smaller than predicted. That is, the condensed water is H2, and then after superimposing it with H1, the actual liquid level value is obtained as shown in Figure H. Based on the above analysis, the water level cannot rise in a straight line with a constant slope, and there must be a certain lag. It is equivalent to a combination of an integral link and a delay link in the system. The transfer function is as formula 5.

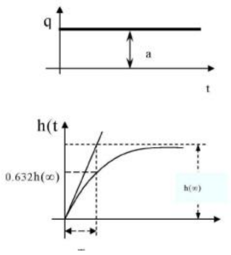


Fig.2.Step response of single tank

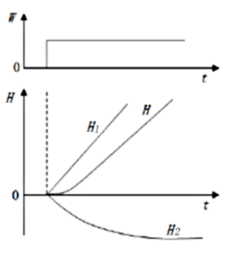


Fig.3.Variation curve of boiler drum liquid level during water supply

$$\frac{H_{(S)}}{W_{(S)}}=\frac{k_0}{s}e^{-s\tau} \quad (5)$$

Where k_0 represents the frequency and change speed of the water level when the feedwater flow under a certain amount changes as a unit; τ represents the delay.

3. PID Control

Through the analysis of the above characteristics, we use PID algorithm to achieve precise level control. For this reason, we choose PLC in the controller. Compared with a single chip, PLC adopts a modular design[4], which can greatly reduce external wiring, make it more convenient and quicker to use, and have higher control accuracy.

3.1. *PID*

PID can also be said to be a control algorithm based on closed-loop proportional, integral, and derivative. The control system continuously compares the set value with the actual process variable, and adjusts the control variable through the PID algorithm[1]. This control variable can adjust the process variable. Makes the set value and the process variable continuously reduce until it is 0. The control quantity is constantly updated, and the PID algorithm formula is used to determine the control quantity at the next moment. Its basic formula can be expressed as formula 6.

$$M_n=K_c*e_n+K_i*e_n+M_x+K_D*(e_n-e_{n-1}) \quad (6)$$

Where K_c is the proportional gain coefficient, which is multiplied by e_n to obtain the proportional control value. e_n is the deviation at the nth time of PLC sampling, e_{n-1} is the deviation at the n-1th time of PLC sampling; K_i is the integral term coefficient ; K_D is the coefficient of the differential term; M_x is the value before the integration; M_n is the calculated value at the nth time of the PID.

$$e_n = (sp_n - pv_n) \quad (7)$$

$$K_i = K_c * T_S / T_i * (sp_n - pv_n \quad (8)$$

$$K_D = K_c * T_D / T_S * (PV_{n-1} - pv_n) \quad (9)$$

Where sp_n represents the system given value at the nth time; pv_n represents the process variable at the nth sampling time; PV_{n-1} represents the process variable at the n-1 th sampling time; T_S represents the sampling time interval; T_i represents the integration time; T_D Represents the derivative time.

There are three PID adjustment modules in Siemens PID module, and different function blocks correspond to different scenarios. The function blocks FB41 "CONT_C", FB42 "CONT_S" and FB43 "PULSEGEN" correspond to different functions, we need to choose according to actual needs. In this research, what we need is a positional PID, so we use the FB41 function block.

3.2. *Control system algorithm*

The position-based PID adjustment adopted by the FB41 "CONT_C" module, and its PID control algorithm structure[5] is shown in Fig. 4.

CYCLE:TIME:PID sampling period, generally set to 200MS; SPINT:REAL:PID set value;

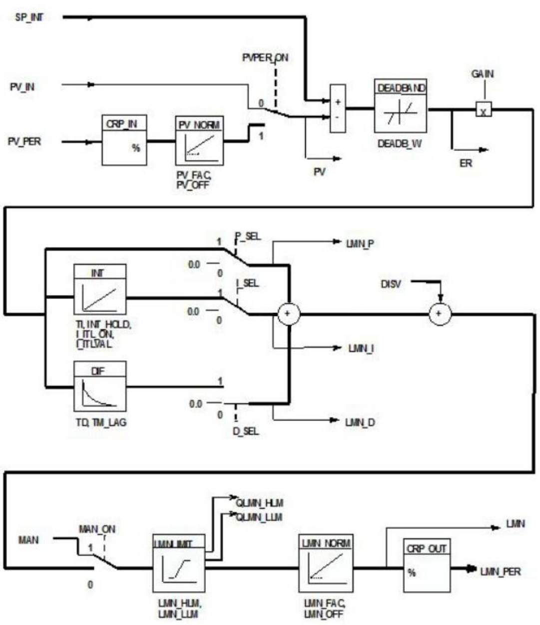


Fig.4.PID Function Block

PVIN: REAL: PID feedback value (also called process variable);

PV_PER:WORD: Unnormalized feedback value, valid by PEPER-ON selection; (not recommended) MAN: REAL: manual value, valid by MAN-ON selection: GAIN: REAL: proportional gain; TI: TIME: integral time ; TD:TIME: differential time;

TM LAG: TIME: I don't know. I haven't used it. It is related to differentiation:

DEADBW: REAL: Dead zone width: If the output oscillates with a slight amplitude near the balance point, you can consider using the dead zone to reduce sensitivity

3.3. *MCGS*

MCGS[5] configuration software is composed of several important parts, and their positions are different, and their functions are also different. The specific process is as follows: (1)) User window: Set the performance of other windows. (2) Main control window: Here you can configure the simulation screen and the parameter adjustment screen, which can define the parameters and reflect the process quantity in real time. (3) Equipment components: Here we build a channel to connect the real-time data converted from A/D and D/A in the monitoring system to the channel correspondingly. We read and write the real-time data in SF41 through the main control window, and display the set volume, process volume, manual volume, etc. in the corresponding read and write mode. The manual and automatic buttons are in the form of a combination of two labels, and the two buttons are reversed according to the button action to realize automatic conversion when pressed. The main control screen is shown in Fig. 5.

The parameter adjustment screen is shown in Fig. 6.

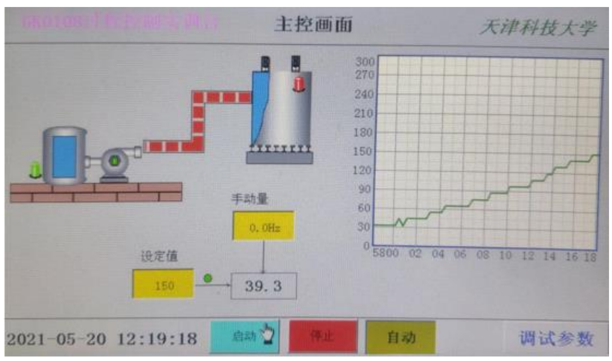


Fig. 5.Main control screen

The functions of the control system are as follows: Input

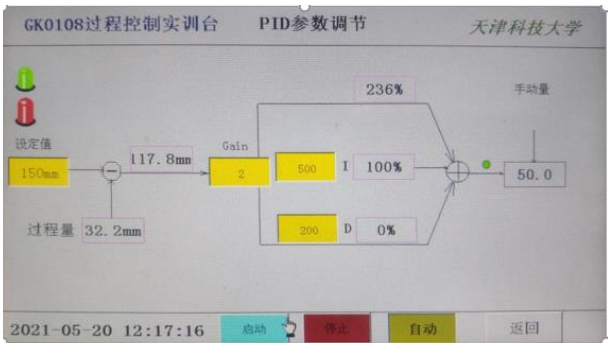


Fig.6.System parameter interface

the set value of the boiler, and the system will control the inverter according to the set value, and then control the liquid level. From Fig.4, we can clearly observe the boiler liquid level change. When the water level value of the water storage boiler is lower than the lower limit, the lower limit alarm light will light up and the motor will automatically stop. It can be seen from Fig.5 that when the liquid level reaches the set value, the total control quantity, namely the frequency of the inverter, will gradually decrease until it reaches zero. When the liquid level reaches the set value, the water level will be slightly higher than the set value due to the water in the pipe, but the liquid level will still be sent out. Therefore, the liquid level will drop. When the liquid level is lower than the set value, the PLC detects that the liquid level is too low. Then start the motor to feed water again, so that the boiler liquid level reaches a dynamic equilibrium.

4. Equations

The thesis introduces the research and theories of liquid level control technology based on Siemens, and focuses on the calculation algorithms and basic theories of various systems and functional blocks in the liquid level control software. Through software programming, hardware connection, sensors, frequency converters and motors, a complete set of liquid level monitoring system is combined. This paper can provide a certain theoretical basis for actual industrial production.

5. References

1. Ou Q , Zhao P , Wu Y . Control system of the boiler water level based on normalized PID[C]// Proceedings of 2013 3rd International Conference on Computer Science and Network Technology. IEEE, 2014.
2. Li Z , Xia S . The Study of Boiler Control System of Water Level of Steam Drum Based on New Immune PID Controller[C]// Second International Conference on Digital Manufacturing & Automation. IEEE, 2011.
3. Yongsheng Z , Shuqin Z . The Fuzzy Self-Tuning PID Control for Boiler Drum Water Level[J]. Applied Mechanics & Materials, 2012, 198-199:1231-1234.
4. Chow siu-fung. Design of boiler automatic control system based on Plc.. Journal of Baotou Polytechnic, 2020,21(03) : 8-10.
5. ZHENG Rongbin. Analysis on the Application of MCGS Configuration Technology in PLC training teaching in Secondary Vocational Schools. *Modern Vocational Education*,2021(46) pp:150-151.

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.

Prof. Dr. 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.

Mr.Xiwen Liang

He is an associate professor in the School of Electronic Information and Automation, Tianjin University of Science and Technology, engaged in the teaching of measurement and control technology and instrumentation. He is deputy director of the teaching and research section.

Recurrence quantification and time-frequency analysis of two-phase flow patterns *

Keying Ren, Xiaoyan Chen, Meng Du

College of Electronic Information and Automation,

Tianjin University of Science and Technology, China;

E-mail: renkeying@mail.tust.edu.cn
www.tust.edu.cn

Abstract

Two phase flow often occurs in industrial production. If it is not detected in real time, it will do great harm to industrial production. In this paper, a method combining recurrence quantitative analysis and time-frequency representation is proposed to identify the flow patterns of the Gas-liquid flow. From the construction of the experimental device to the collection of conductance fluctuation signal of two-phase flow, the recurrence plot and WVD distribution map are drawn by MATLAB which are used in the final flow pattern analysis of two-phase flow. Experimental results show that our method can accurately identify the flow pattern of two-phase flow.

Keywords: Two−phase, Recurrence quantification, time-frequency, flow-pattern

1. Introduction

Two-phase flow is a complex nonlinear system, when the two-phase flow is affected by the relative movement between the two phases caused by the difference in flow velocity between the two phases, and the fluid turbulence in the pipeline, the phase flow presents a high degree of uncertainty, irregularity and activity. In recent years, great progress has been made in the nonlinear analysis of two-phase flow patterns in China. Nonlinear characteristic analysis is applied to the signal processing of two-phase flow measurement, which provides strong support for the highly complex and uncertain flow pattern transformation mechanism of two-phase flow.

The concept of recursive graph was first proposed by JP Eckmann et al. in 1987[1]. After in-depth studies by Eckman, Casdagli, Marwan, Yan et al., recursive graph has been applied to the accurate identification of nonlinear dynamical system. And gradually transformed to the correlation analysis of oil-water flow, through the study found that the flow pattern of two-phase flow can be well reflected by recursive graph analysis method. The characteristics of fluid dynamics and the internal state of the kinematic system can be intuitively reflected through the recursive graph. But recursive graphs are difficult to measure dynamic systems. For this reason, Zbliut and Webber[2] proposed recursive quantitative analysis, which is mainly used to analyze the characteristics of recursive graphs.

In 1807, Fourier first talked about the method of understanding the intrinsic properties of the signal, but this method can only process a large section of the signal, and cannot process the local part of this section of the signal. In 1946, the short-time Fourier transform was proposed. In 1981, Y. Meyer and J. Morlet first studied the wavelet transform method[3]. After continuous development, it finally became the most effective method of time-frequency feature analysis. In 1932, Wigner proposed a method of analyzing unstable signals, which is Wigner's analysis method. However, this method will have cross terms when processing unstable signals, which will interfere with the experimental results. By analyzing this shortcoming, in 1966, Cohen further analyzed this method, which weakened the interference of the Wigner distribution cross term on the experimental results. This method is the Cohen-like time-frequency distribution analysis method. The time-frequency distribution of Cohen class is defined as follows:

$$P(t,f)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} z(u+\frac{1}{2}\tau)z^*(u-\frac{1}{2}\tau)\phi(\tau,v)e^{-j2\pi(vt+f\tau-vu)}dudvd\tau \qquad (1)$$

Many efficient time-frequency analysis methods have been obtained. There are still many shortcomings in the field of time-frequency analysis. The research on the analysis method of unstable signal is still a big problem that we need to overcome.

2. Hardware Design

In order to observe the flow pattern of the two-phase flow and the evolution of the flow pattern in the fluid flow process, a dynamic experimental device for the two-phase flow is designed to achieve uniform fusion of the gas and water in the pipeline, peristaltic metering pumps and gas flowmeter are installed below the water tank and air bottle to achieve real-time measurement of the two parameters.

1. Water tank and air compressor: These two containers mainly store water and air.

2. Peristaltic pump and gas flowmeter: The peristaltic pump can steadily transfer water from the tank to the experimental measurement pipe, which can detect the flow and speed of the liquid. Gas flowmeter is mainly used to detect the flow rate gas.

3. Experimental measurement pipeline: Through the experimental pipeline, we can fully mix the two substance, and can better allow the measurement device to measure the two-phase flow with real data.

4. Measuring device and sensor: The measuring device can measure the state of the two-phase flow in the experimental measurement pipeline in real time, and record the flow pattern of the two-phase flow and the evolution process of the flow pattern in time.

After the experiment began, the peristaltic pump and gas meter transported water and air stably to the experimental measurement pipe, forming gas and liquid two-phase flow. Real-time experimental data were recorded by measuring devices and sensors to provide basis for subsequent experimental analysis.

3. Data analysis

3. 1 *Time-frequency characteristic analysis*

The dynamic experiment of two-phase flow at low flow rate was carried out through the above-mentioned experimental device, and a large number of conductance fluctuation signals that can reflect the characteristics of Gas-liquid flow were obtained. Next, we will pass the signal obtained under the above experiment through Wigner-Vill. The Wigner-Vill transform method performs time-frequency characteristic analysis. After a large number of experimental verifications, we have summarized some flow pattern characteristics in each stage of the Gas-liquid flow: When the fluctuation frequency of the conductance fluctuation signal is low and the signal energy is relatively concentrated, the energy of a certain part is not very low, its Wigner-Vill transform map has the characteristic of slug flow. When the fluctuation frequency of the conductance fluctuation signal becomes higher and its energy is scattered and the local energy is small, its Wigner-Vill transform map has the characteristic of bubbly flow. In the Gas-liquid flow dynamic experiment, we can know that when the speed of water in the Gas-liquid flow becomes larger and larger, the energy of the conductance fluctuation signal becomes concentrated. When the flow rate of the water becomes faster, the conductance fluctuation signal becomes dispersed. Some experimental results are shown as Fig. 1.

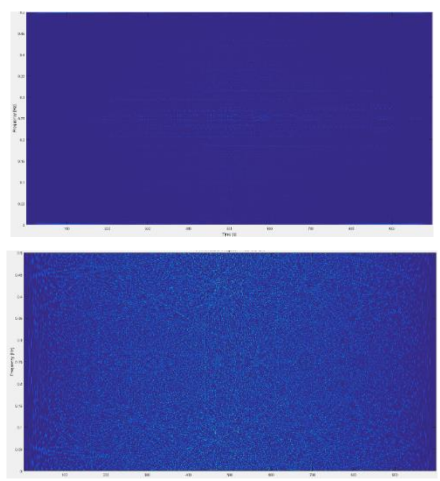


Fig.1.Time-frequency distribution atlas under different flow patterns

Through experiments, we can observe the different characteristics of the Gas-liquid flow under different flow patterns, however, the flow pattern of Gas-liquid flow cannot be accurately distinguished, so we also need to perform quantitative analysis of time-frequency data. Through quantitative analysis of time-frequency characteristics, we can accurately distinguish the flow pattern of Gas-liquid flow. We analyze these signals by summing the energy of each signal point. some experimental data are shown as Fig. 2.

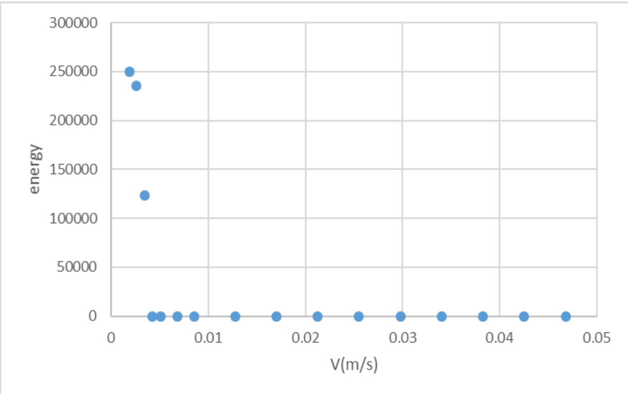


Fig. 2. Under different water content, the signal energy changes with the flow rate.

3. 2 *Recursive graph analysis of oil-water flow*

In the above experiment, we have obtained a large number of conductivity signals of gas-liquid flow, and now we make a series of recursive processing on these conductivity signals. After observing the signal graph and recursive graph, we can observe the recursive characteristics of different flow patterns. Some experimental results are shown as Fig. 3

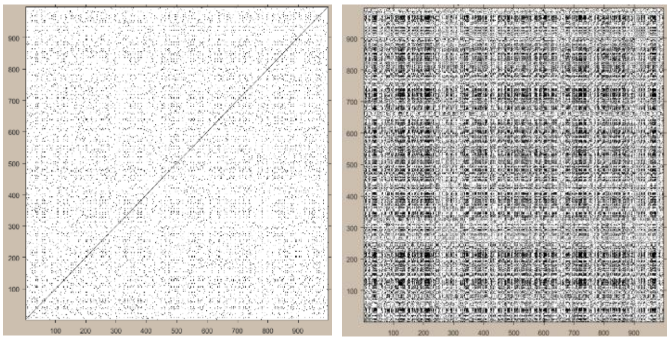


Fig. 3. Recursion diagrams for different flow patterns.

3. 3 *Quantitative analysis of recursive graph of oil-water flow*

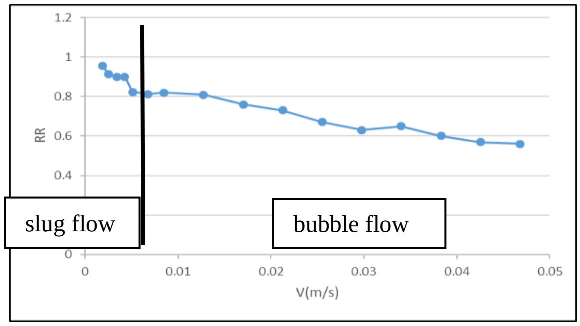
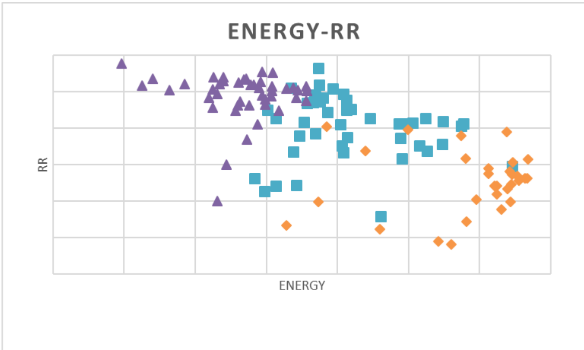


Fig. 4. Recursive rate of different flow.

By observing the above chart, we can know that when the gas flow velocity is fixed, the RR of the two-phase flow, namely the recursive rate, will decrease with the increase of the liquid flow velocity. experimental results are shown as Fig. 4.

3. 4 *Recurrence quantification and time-frequency analysis of two-phase flow patterns*

We have obtained some important conclusions from the recursive quantitative features and time-frequency features. We conduct a joint analysis of the flow pattern dynamics of two-phase flow based on recursive



quantification and time-frequency.

Fig. 5. Energy-recursive rate plot. In the figure, the purple point represents the Fine vesicular flow, the blue square represents the bubble flow, and the orange square represents the slug flow.

We can get the flow pattern of the two-phase flow through the above analysis and the position of the signal point in the figure. Recursive quantitative analysis and time-frequency characteristic analysis are combined to identify the flow pattern of two-phase flow. This method can distinguish flow patterns more efficiently and accurately.

4. Discussion

Through the above conclusions, we can know that recurrence quantification analysis and joint analysis of time-frequency characteristics can accurately identify flow patterns. As the requirements for industrial production and pipeline transportation are gradually increasing, This accelerates the research on the signal processing technology of two-phase flow. It is hoped that more efficient measuring instruments can be developed through this subject, so that industrial production becomes safer, more efficient and reliable.

References

1. Eckmann J P, Kamphorst S O, Ruelle D. Recurrence plots of dynamical systems. *World Scientific Series on Nonlinear Science Series A*, 1995, 16: pp.441-446.
2. Zbilut J P, Webber Jr C L. Embeddings and delays as derived from quantification of recurrence plots, *Physics letters A*, 1992, 171(3-4):pp.199-203.
3. Zhang, Dengsheng, Wavelet transform, *Fundamentals of Image Data Mining*, Springer, Cham. 2019. Pp.35-44

Authors Introduction

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. Dr. Xiaoyan Chen

She is 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.

Dr. Meng Du

He received the Ph.D. degree in control science and engineering from Tianjin University, China, in 2013. He is currently with the College of Electronic Information and Automation, Tianjin University of Science and Technology. His current interests include multi-phase flow parameter measurement, flow pattern identification, time series analysis, and machine learning.

Application of Deep Learning in Automatic Driving

Wei Su, Xiaoyan Chen*

Tianjin University of Science and Technology,300222, China;

E-mail: 965756192@qq.com

www.tust.edu.cn

Abstract

With the continuous development of science and technology, the field of artificial intelligence has become a research hotspot, especially deep learning, which has attracted much attention from all walks of life. Starting with the automatic driving solution, this paper mainly expounds the important role and technical route of the application of deep learning, and finally looks forward to the development direction and application of automatic driving technology based on deep learning.

Keywords: Automatic Driving, Deep Learning

1. Introduction

Automatic driving is developing rapidly in the direction of intelligence. As a cutting-edge branch of artificial intelligence, deep learning has promoted the explosive growth in the field of artificial intelligence, and has been widely used in the perception, decision-making, control and other related fields of automatic driving. The current research focus is computer vision. The research hotspot is the application of convolutional neural network to robot objects. Deep learning is also used to assist decision-making. The integration of deep learning and autopilot has a good prospect, which will promote the development of autopilot to a higher level of intelligence.

Deep learning is a kind of machine learning. Its algorithm is based on representation learning of data, which can discover the complex structure in big data, and use back propagation to guide the machine how to calculate the representation from the previous layer network, so as to change the internal parameters of each layer[1]. The purpose of deep learning is to make the unmanned vehicle have the same analysis and learning ability as people, because it can enhance the perception, decision-making and control ability of unmanned vehicle, it has a wide application prospect in image recognition, target recognition, path planning, human-computer interaction and so on.

2. Development status of automatic driving

2.1. *Automatic driving level division standard*

The automatic driving level division standard formulated by the SAE (Society of Automotive Engineers) is shown in Fig.1.
Level 1: system is mainly used to improve driving safety. Such as ABS(antilock brake system).
Level 2: partial automatic driving. The vehicle provides driving for multiple operations in steering wheel, acceleration and deceleration, and the human driver is responsible for other driving actions.

Level 3: conditional automatic driving. Most driving operations are completed by the vehicle, and human drivers need to pay attention in case of emergency.
Level 4: highly automatic driving. cars can drive automatically in special scenes such as highways, but cars still need human intervention under complex road conditions.
Level 5: fully automatic. human beings become passengers completely

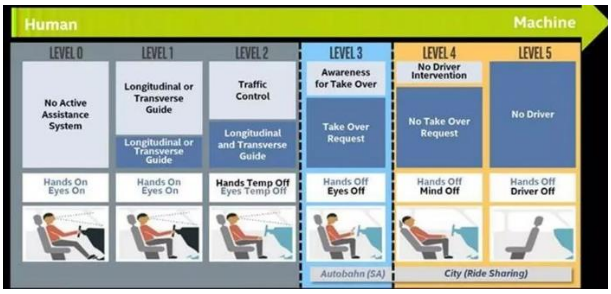


Fig. 1. The automatic driving level division standard

2.2. *Key technologies of automatic driving*

2.2.1 Environmental perception

Commonly people use sensors for vehicle environment sensing include lidar, millimeter wave radar, camera, etc. Cameras include monocular camera, binocular camera, and look around camera. Monocular camera has simple structure and mature algorithm, but its field of view depends on lens and ranging accuracy is low. The binocular camera has higher ranging accuracy. The look around camera can have a field of view of 360 degrees, but the sensing range is only 5 to 10 m, and the edge of the photo is seriously deformed.

In order to solve the problem of inaccurate camera ranging, engineers introduced lidar sensors. The lidar sensor emits a laser beam externally. After the laser beam meets the object, it returns to the laser receiver through diffuse reflection. The computer calculates the distance between the transmitter and the object according to the time difference between sending and receiving signals. The fast-rotating laser sensor can provide millions of data points per second, which can create 3D maps of surrounding objects and environment. Lidar is not affected by lighting conditions, but it is vulnerable to smoke or bad weather, and the cost of lidar is very high.

Millimeter wave radar is affordable, less affected by smoke, strong penetration and small volume. Its detection range can reach 200m. However, the data stability is poor and the algorithm is more complex.

2.2.2 High precision position

High precision position provides global path planning for autonomous vehicles, and obstacle avoidance planning based on perception results, also known as real-time navigation planning. At present, the two major challenges that positioning technology still faces are covering blind areas and high cost.

2.2.3 Collaborative decision making

The decision-making needs to integrate the current scene, path planning, user needs and other information, adapt to the real-time situation as much as possible, and make friendly and efficient decisions on the premise of ensuring driving safety[2].

3. Deep learning application on automatic driving

3.1. *Lane line detection*

Many scholars began to use the methods of deep learning to solve the problem of lane detection.

In 2018, Davy Neven[3] and others proposed a lane line detection neural network model LaneNet+H-Net, which can carry out end-to-end lane line detection. The model structure mainly including two network models LaneNet and H-Net, which is shown in Fig.2. LaneNet is a multi task model. One branch is used for binary semantic segmentation of lane lines to distinguish whether they are

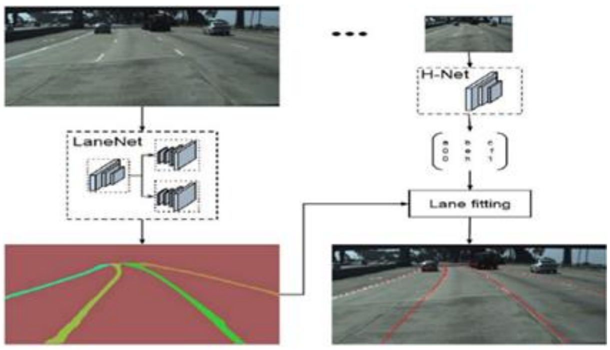


Fig. 2. LaneNet + H-Net network structure[3]

lane lines or background. The other branch uses the lane line division output of the first branch as input. During training, the segmented lane line mask is clustered by learning the designed clustering loss function, and the two branches are combined. LaneNet + H-Net network realizes end-to-end lane line instance segmentation.

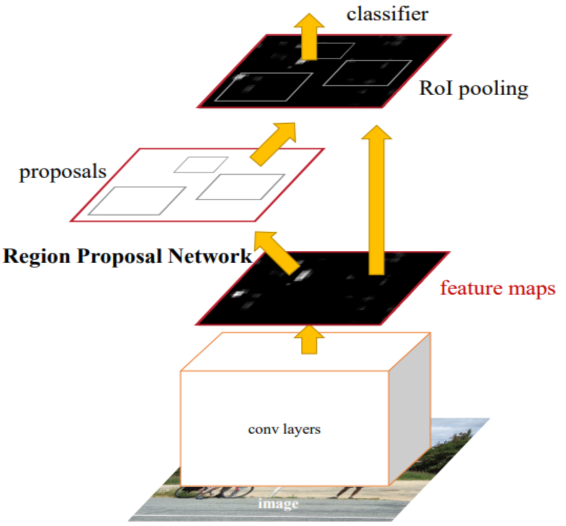


Fig. 3. Faster RCNN network structure[5]

3.2. *Pedestrian detection*

RCNN(Region CNN) algorithm is a pioneering work for target detection using deep learning. It is designed by Ross Girshick, a famous fair programmer under Facebook. It is also a great milestone in applying the traditional CNN method to target detection. Based on CNN's excellent feature extraction and performance classification, the transformation of target detection problem is realized by region proposal method.
However, RCNN has fatal shortcomings. Repeated feature extraction leads to a waste of computing resources. The later fast RCNN is improved on this basis to speed up the algorithm and no additional storage[4].

Faster RCNN network structure is shown in Fig.3. It integrates feature extraction, target extraction, candidate box and classifier into a network, which is completed by deep neural network and runs on GPU, which greatly improves the operation efficiency. Compared with fast RCNN, the comprehensive performance of RCNN is greatly improved. It is one of the mainstream algorithms of target detection. The faster RCNN network framework is shown in the Fig.3.

3.3. *Multisensor fusion*

Automatic driving based on deep learning has higher requirements for sensor data acquisition. As the data source of unmanned vehicle, various sensors are of great importance. However, as the input of the deep learning model, any sensor has its own shortcomings. So the

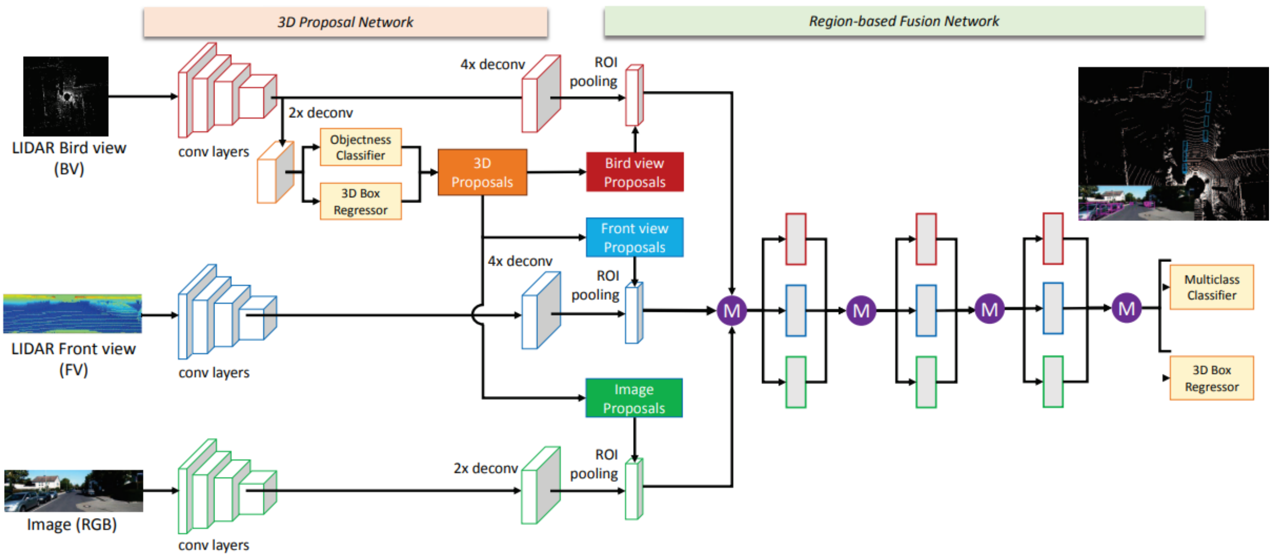


Fig. 4. MV3D object detection network structure[8]

multi-sensor fusion technology is of great significance for the accurate perception and cognition of autonomous vehicles.

At present, the target recognition scheme of multi-sensor fusion mostly integrates the traditional camera image and laser point cloud 2D processing scheme into one network, such as VeloFCN[6], Vote3D[7], etc. X Chen et al. proposed MV3D (multi view 3D object detection network) to fuse LIDAR point cloud data with RGB image information[8]. MV3D is used as the perceptual fusion framework. The 2D processing scheme of laser point cloud data represents the 3D point cloud information with the front view and aerial view of laser point cloud, and is fused with RGB image to predict the directional 3D boundary box[9]. The network structure is shown in Fig.4.

4. Conclusions

At present, deep learning shows a broad development prospect in the field of automatic driving. It can operate the original sensor data without manual assistance, infer the key features, and greatly shorten the preliminary engineering time. At the same time, it is also good at fusing multi-sensor data to capture the relationship between the original data, which is helpful for the autonomous vehicle to work in a complex environment. At present, the research focus of in-depth learning is computer vision, which will develop towards a higher level of cognition in the future. The purpose is to make autonomous vehicles reach the level of level 5 and realize real autonomous driving.

References

1. Yann LeCun, Yoshua Bengio , Geoffrey Hinton, Deep learning, *Nature*, 2015,521(7553):pp.436–444.
2. Peng Guo, Xueyi Wu, Hui Rong , et al, Local path planning of driverless cars based on cost function. *China Journal of Highway and Transport*, 2019,32(06):pp.79-85.
3. Davy Neven; Bert De Brabandere; Stamatios Georgoulis, et al.,Towards End-to-End Lane Detection: an Instance Segmentation Approach,*2018 IEEE Intelligent Vehicles Symposium (IV) IEEE*, 2018, pp. 286-291.
4. A. Ullah, H. Xie, M. O. Farooq, et al., Pedestrian Detection in Infrared Images Using Fast RCNN, *2018 Eighth International Conference on Image Processing Theory, Tools and Applications (IPTA)*,2018:pp.1-6.
5. S. Ren, K. He, R. Girshick, et al.,Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017,39(6):pp.1137-1149.
6. M. Engelcke, D. Rao, D. Z. Wang, et al. Vote3Deep: Fast Object Detection in 3D Point Clouds Using Efficient Convolutional Neural Networks, *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2017,pp. 1355-1361.
7. J. Lee, S. Walsh, A. Harakeh, et al.,Leveraging Pre-Trained 3D Object Detection Models For Fast Ground Truth Generation, *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*,2018,pp. 2504-2510.
8. X. Chen, H. Ma, J. Wan, et al., Multi-view 3D Object Detection Network for Autonomous Driving,*2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 6526-6534,
9. C. R. Qi, W. Liu, C. Wu, et al., Frustum PointNets for 3D Object Detection from RGB-D Data, *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 918-927.

Authors Introduction

Mr. Wei Su

He received his bachelor's degree from the college of computer and Communication Engineering of Zhengzhou University of Light Industry in 2021. He is acquiring for his master's degree at Tianjin University of science and technology.

Prof. Xiaoyan Chen

She is 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.

An intelligent home security system based on STM32

Wei Su[1], Xiaoyan Chen[*], Guangyong Xi[2]

[1]*Tianjin University of Science and Technology, 300222, China;*

[2]*Zhengzhou University of Light Industry, 450001, China;*

E-mail: 965756192@qq.com

www.tust.edu.cn

Abstract

Based on STM32F103C8T6 host control chip, this paper completes the design and development of an intelligent home security system combined with various hardware modules, cloud server and android app. The system integrates OLED display screen, buzzer, human infrared, Wifi and other modules. The serial port is connected with ZigBee coordinator. The terminal equipment with temperature, humidity and smoke sensors is adopted to monitor the environment, and the relays and stepping motors are designed to control the equipment at home by simulation. The access control system takes STC89C52 as the core, and integrates the functions of photographing module and steering gear simulation door opening. Android app uses socket technology to complete remote data communication through Alibaba cloud server and indoor security system. After testing, the system runs stably and has good performance.

Keywords: Smart home, Security, STM32,Zigbee, Wifi

1. Introduction

In recent years, smart home has developed rapidly. As an important part of smart home, smart home security is related to the safety of users' lives and property. Smart home security system can remotely monitor the home environment, accurately detect gas leakage, fire and other dangerous situations, and timely notify users through mobile app to strengthen the protection of users' lives and property. In addition, the access control system needs password to open the door, and the photographing function can take pictures when there are people in front of the door and save them to the memory card. The design of android app allows users to access the home security system by installing only one software, which greatly improves the convenience.

2. Overall system design

The overall design block diagram of smart home security system is shown in Fig. 1., which is composed of android app, server, indoor security system and access control security. The design of each part is as follows:

(1) This design uses Alibaba cloud server as the message transfer station of smart home security system to realize the connection between app and home security system in the wide area network. The main function is data forwarding.

(2) Users can remotely control their home devices and view their home conditions by operating the UI controls on the app. When an abnormal situation happens at home, app can send a warning notice to users.

(3) The indoor security system uses Zigbee to realize home networking. The kitchen node monitors the home environment through temperature and humidity sensors and smoke sensors, and sends the monitoring data to the coordinator; The bedroom node controls the stepper motor and relay. STM32 displays the relevant information on the OLED display screen and monitors whether there are people at home.

(4) The access control system is designed with password lock and automatic photographing function.

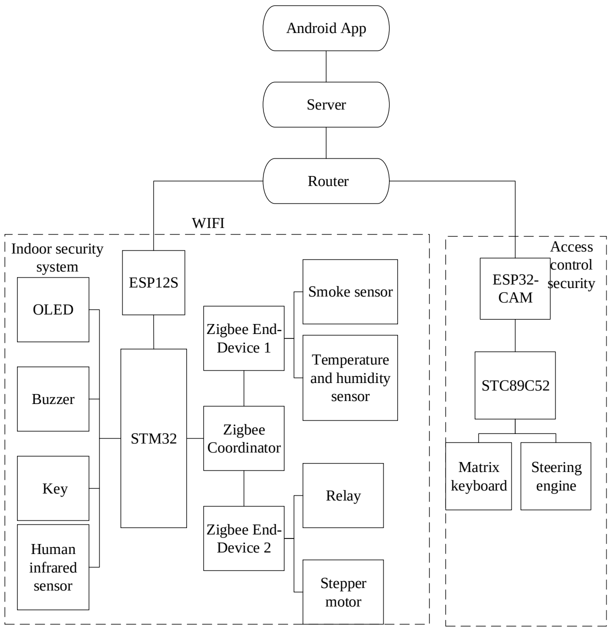


Fig. 1. The overall design block diagram of smart home security system

3. System hardware design

3.1. *Host Module*

The system needs a model that needs to be connected to human infrared sensor, OLED display, independent key, buzzer and other modules, and also to Zigbee coordinator. It needs to use multiple serial ports, IIC interfaces and multiple GPIO ports. Therefore, host module needs to be selected that its performance can meet the demand. In addition, low price and strong compatibility are also factors to be considered. Considering all aspects, the system uses STM32F103C8T6 as the host control chip.

3.2. *Communications module*

The wireless data communication scheme in this system is completed by Zigbee and Wifi. As the communication technology of indoor security system, Zigbee is mainly responsible for transmitting control commands and data collected by sensors. Wifi mainly provides external network access for the home security system, so that the system can transmit data with the cloud server.

3.3. *Various sensors*

The function of monitoring the home environment is realized through various sensors.

3.3.1. *Smoke monitoring module*

The system uses MQ-2 as the smoke monitoring module. The module converts the collected voltage analog signal into digital signal, and then converts it into combustible gas concentration according to formula[1] (1).

$$Rs/R0 = 11.5428 \times ppm^{(-0.6549)} \tag{1}$$

Where RS represents the current resistance value of the sensor, and R0 is the resistance value of the sensor in clean air.

3.3.2. *Intrusion monitoring module*

The system selects HC-SR501 to realize intrusion monitoring in home mode. The module can be placed at the porch to monitor intrusion. The working principle of HC-SR501 is to use infrared radiation. If the human body is detected, the sensor can generate a signal[2].

3.3.3. *water leakage monitoring module*

The system selects the raindrop sensor module to realize water leakage monitoring. The raindrop sensor is placed in places prone to water leakage such as kitchen and toilet. The sensor module has four pins, namely VCC, GND, DO and AC. DO is the digital signal output. The output high and low levels indicate whether there is water on the sensor surface.

3.3.4. *temperature and humidity monitoring module*

The system selects DHT11 temperature and humidity sensor to measure temperature and humidity. The sensor integrates temperature and humidity sensor elements, occupies less IO resources, and can read data with only one line. And the output is a digital signal with high accuracy[4].

3.4. *access control sub-system*

The access control system uses STC89C52 as the control core and uses the minimum system board to connect ESP32-CAM, HC-SR501, matrix keyboard and SG90 steering gear.

The system needs to capture the situation outside the door and save them. Therefore, the system selects ESP32-CAM as the access control camera module. ESP32-CAM uses serial port to receive data, connect OV2640 camera to take pictures, and insert SD card to store information. When the human body infrared sensor detects a person, it controls the camera to take pictures and save them.

4. System software design

4.1. *STM32*

The STM32 program flowchart is shown in Fig.2. After the STM32 is powered on, it first tests the relevant hardware such as OLED, keypad, LED and serial port, initializes the WIFI module, and then calls the function to send the AT instruction to the ESP-12SWIFI module through the serial port 3, so that it connects to the WIFI hotspot named Tom. After that, the single connection mode is set up to connect the Ali cloud server according to the server's external network address and port.Then

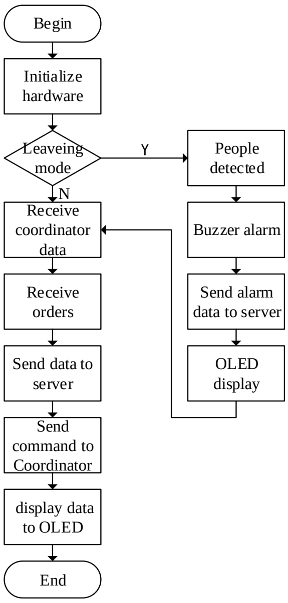


Fig. 2. The flow chart of STM32

judge whether it is home mode, and then process the data sent by the server and Zigbee. Since STM32 serial port 1 is connected to Zigbee coordinator and serial port 3 is connected to ESP-12S Wifi module, it needs to be handled separately.

4.2. *App*

App is initialized first and displayed on the interface after receiving the data. After receiving the alarm information, the app will be notified. You can also control home appliances through buttons. The App program flowchart is shown in Fig.3.

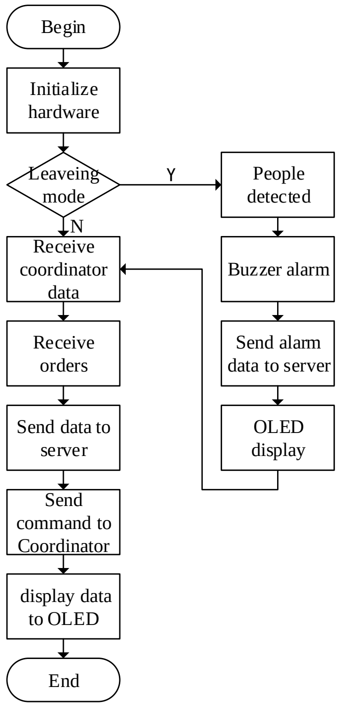


Fig. 3. The flow chart of App

4.3. *Zigbee*

After acquiring the environmental information data such as temperature, humidity and smoke, the terminal will call the data transmission function AF_ Datarequest(), you can set the sending mode to unicast, target address 0x00, target endpoint and cluster, and then send the data to the coordinator. The coordinator sends it to STM32 through the serial port.

4.4. *Server*

The server uses Python to write the socket server. In the initialization function init(), the socket object is first established in the function.Then the socket object is bound to the address that is the private address of the server. Every time a client connects to the server, a new thread will be opened to accept the new connection. When data is received, it will be forwarded to another terminal.

4.5. *Access control system*

The access control system monitors whether there are people outside the door through infrared sensors, and others control ESP32-CAM to take photos. Enter the password through the matrix keyboard, and if it is correct, System will control the steering gear to open the door. The access control system program flowchart is shown in Fig.4.

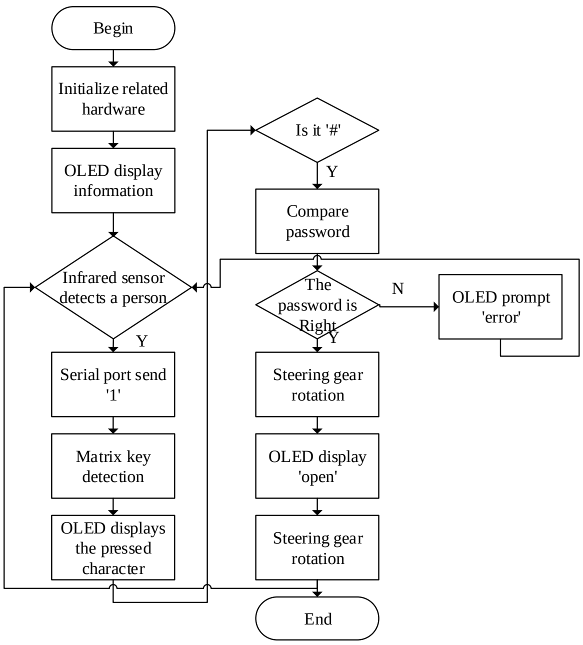


Fig. 4. The flow chart of access control system

5. System test

The appearance of the system is shown in the Fig.5. Start and build Zigbee wireless sensor network, preset Wifi hotspots, and connect open source hardware to the external network;

After that, open source hardware pushes various sensor data to Alibaba cloud platform through Wifi to realize real-time monitoring of home environment. Users can receive and process equipment management and alarm information through mobile app. For example, in case of flame, combustible gas detection or abnormal temperature and humidity, the system will push it to app for alarm.

App can display home environment information in real time, and can also control home devices through buttons. The app interface is shown in the Fig.6.

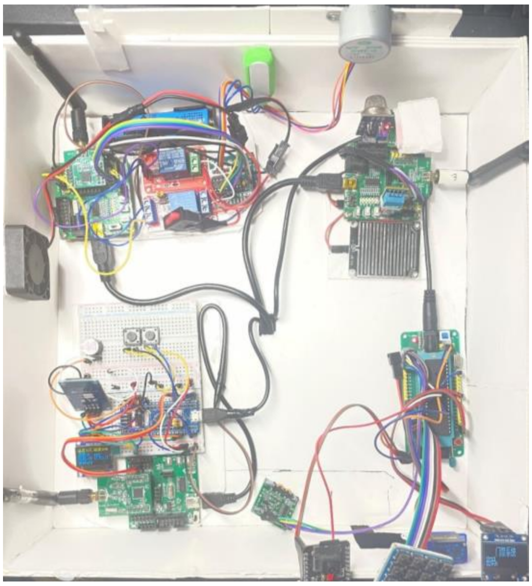


Fig. 5. The appearance of system

Fig. 6. The interface of App

The access control system has accurate password input and sensitive photography.

6. Conclusion

According to the long-term test,the intelligent home security system can monitor all kinds of sensor data in real time and upload them to the cloud platform of the Internet of things accurately and stably. Users can obtain home environment data through app to realize accurate monitoring and early warning. The system has the advantages of simple construction, convenient data query and strong scalability; It is not only applicable to the monitoring and management of home environment, but also can be used in office areas, venues and other fields.

The follow-up work focuses on the following two aspects: (1) strengthening the monitoring and early warning research on smart environment application scenarios;(2) studying artificial intelligence algorithm for the secondary development of the Internet of things cloud platform.

References

1. Trisnawan, I. K. N., Jati, A. N. , Istiqomah, N.et al., Detection of Gas Leaks Using The MQ-2 Gas Sensor on the Autonomous Mobile Sensor. *2019 International Conference on Computer, Control, Informatics and its Applications (IC3INA)*,2019,pp.177-180.
2. Y. Lihong,W. Jianxin, The Design of Intelligent Automatic-Door Based on AT89S52,*2016 International Conference on Robots & Intelligent System (ICRIS)*, 2016,pp. 428-431.
3. G. M. Debele and X. Qian, Automatic Room Temperature Control System Using Arduino UNO R3 and DHT11 Sensor, *2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, 2020, pp. 428-432.

Authors Introduction

Mr. Wei Su

He received his bachelor's degree from the college of computer and Communication Engineering of Zhengzhou University of Light Industry in 2021. He is acquiring for his master's degree at Tianjin University of science and technology.

Prof. Xiaoyan Chen

She is 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.

Mr. Guangyong Xi

He is an Associate Professor of Zhengzhou University of Light Industry, graduated from Hohai University with PH.D (2007).From June 2012 to now, he has been engaged in teaching and scientific research in the school of computer and communication engineering of Zhengzhou University of Light Industry.

Design of Smart Bracelet Based on STM32 Microcontroller *

Xia Miao, Xiaoyan Chen*, Jianliang Li

Tianjin University of Science and Technology,

No. 1038 Dagu Nanlu, Hexi District, Tianjin, China

E-mail: miaoxia@mail.tust.edu.cn
www.tust.edu.cn

Abstract

With the development of electronic information technology, the demand for high-precision and convenient electronic detection equipment for biomedical signals goes straightly high. Heart rate and steps counting are two important indicators of the human body. Based on this, this design studies an intelligent bracelet with health monitoring system, which can detect human movement state and steps counting, measure and analyze the heart rate, and connect wirelessly through Bluetooth module. A STM32 microcontroller is used to collect and analyze the information of motion state, heart rate and step quantity, and then send them to the APP, which is developed based on Android and displayed through Bluetooth module. The users can change their lifestyle by above parameters. Smart bracelets can play a role in reducing the risk of disease.

Keywords: STM32 microcontroller, smart bracelet, health detection, sensor

1. Introduction

Smart bracelet is a kind of wearable smart device. Today, with the rapid development of science and technology, smart wearable technology and Internet of Things technology are becoming more and more popular. Especially in the field of biomedicine, it promotes the progress and development of science and technology, and at the same time provides convenience for the medical treatment of people's health. Smart wearable devices are portable, accurate and functional, which brings a very good development prospect for smart bracelet in the future. Recently, the cooperation between smart wearable devices and the medical field has become increasingly mature. Using the Internet of Things technology and single-chip microcomputer for data collection and data analysis to improve the quality of life, network informatization is an inevitable development trend in today's society[1]. This paper designs and studies smart bracelets based on STM32 single-chip microcomputers by surveying the current social health status and general living habits, which are mainly used to detect the individual's exercise status and supervise the wearer's exercise.

2. Control system composition

The smart bracelet control system is composed of a microprocessor module, a heart pulse detection module, a motion state detection module and a Bluetooth module. The smart bracelet uses STM32 as the controller to receive the heart rate, exercise status and other data sent by the sensor, and process these data, analyze the human body's exercise status and physical health, and display the results on the OLED display. At the same time, the data is uploaded to the computer through the Bluetooth module for storage, which serves as the basis for the user's health dynamic management. The structure diagram of the smart bracelet control system is shown in Fig. 1

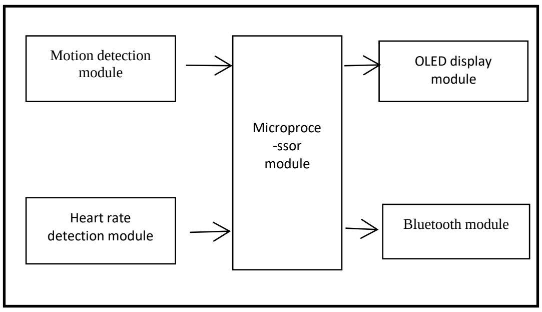


Fig. 1. The structure diagram of the smart bracelet control system

2.1. *Microprocessor module*

The STM32 series of single-chip microcomputers are a series of very powerful single-chip microcomputers with very high cost performance[2]. It is based on the ARM Cortex-M core designed for embedded applications that require high performance, low cost, and low power consumption. It has first-class peripherals: 1µs Dual-channel 12-bit ADC 4bit/s UART 18Mbit/s SPI, etc. In terms of power consumption and integration, STM32 microcontrollers also have a good performance, which is obviously slightly inferior to the power consumption of MSP430, but this does not affect the enthusiasm of engineers for it.

2.2. *Heart rate detection module*

The pulse measurement uses the heart rate pulse sensor module[3]. There are three main methods of pulse measurement: one is to extract from the ECG signal, the other is to calculate the pulse rate from the fluctuation measured by the pressure sensor, the third is the photovolume method. The basic principle of photovolumetric method is to use human tissues to make pulse measurement with different light transmittance caused by pulsation of blood vessels. The sensor used is composed of two parts: a light source and a photoelectric transducer, and is fixed on the patient's finger or earlobe through a band or a clip. The light source generally uses a light-emitting diode with a certain wavelength (500nm~700nm) that is selective to the oxygen and hemoglobin in the arterial blood. When the light beam penetrates the human peripheral blood vessels, the light transmittance of this light is changed due to the change of the arterial pulsation congestion volume. At this time, the photoelectric transducer receives the light reflected by the human tissue, converts it into an electrical signal and amplifies and outputs it. Since the pulse is a signal that changes periodically with the beating of the heart, and the volume of the arterial blood vessel also changes periodically, the change period of the electrical signal of the photoelectric transducer is the pulse rate. Most smart watches use photoelectric method to monitor heart rate. Their obvious feature is that the sensor part is equipped with a green LED light. There are many types of photoelectric sensors for this measurement principle. According to the different receiving positions of the light signal, the photoelectric method can be divided into two modes: transmission and reflection.

2.2.1. *Transmission photoelectric method*

The transmission photoelectric method refers to that the generator and photosensitive receiver on the wearable device are located on both sides of the measured part (usually fixed by a clip). The incident light passes through the skin and enters the deep tissues, except for the skin, muscle, blood, Outside of the absorption by bones, the remaining part of the light transmission is sensed by the photosensitive receiver. According to its principle, this method is suitable for measuring parts with relatively short distance between the two sides of the human body, such as earlobes, fingers, toes, etc., and representative smart wearable products are those ear clip heart rate monitors and nail oximeters Wait. Smart wearable products that adopt the transmissive photoelectric method are usually fixed with a clip. The product of this monitoring method usually adopts the structure of a sealed cassette in appearance, which can well reduce the external light interference, thereby improving the measurement accuracy and stability. Due to its high signal-to-noise ratio and stable signal, in addition to measuring heart rate, it can also analyze heartbeat function, blood flow and many other cardiovascular physiological information through waveforms. The disadvantage is that it is not suitable for use on smart bracelets and smart watches, and products used on earlobes, toes and other parts will feel uncomfortable to wear.

2.2.2. *Reflective photoelectric method*

Contrary to the transmissive photoelectric method, in the reflective photoelectric method, the emitter and the photosensitive receiver on the wearable device are located on the same side of the measured part, and the reflected light is mainly measured. The advantage of this method for measuring heart rate is that it is very simple and has low requirements for the measurement site. As long as the tissue is relatively smooth and there is less subcutaneous fat, it can almost be measured, such as the forehead and wrist. Therefore, most wearable devices such as smart bracelets and smart watches adopt this method to measure heart rate. Moreover, the appearance of smart bracelets or smart watches in the form of products also perfectly solves the dual requirements of transmission-type photoelectric method center rate monitoring and wearing comfort. However, although the reflective photoelectric method performs well in a stable state, when the device is worn on the end of the wrist, it will swing up and down like a pendulum as the user walks or moves irregularly. The centrifugal force will cause a large change in blood volume; The interaction of systolic pressure and centrifugal force in the blood makes it more difficult to distinguish the amount of blood in the blood vessels. Therefore, the accuracy of the heart rate data may be reduced. In addition, the tightness of the wearable device and the blood flow of the human skin will also affect the accuracy of the monitoring.

This article chooses the photoelectric reflection measurement method. The photoelectric reflection measurement method uses a photoelectric sensor for measurement. The photoelectric sensor is composed of a transmitting tube and a receiving tube. The transmitting tube emits a certain wavelength of light, which is transmitted back to the receiving tube through the refraction and reflection of the blood vessel. The cyclical contraction and relaxation of the heart cause regular contraction and dilation of venous blood vessels. The blood and hemoglobin in the blood vessels undergo regular changes in concentration and cause changes in blood transmittance. The light signal then fluctuates regularly and then returns. Receiving tube. The light signal received by the receiving tube is converted into an electrical signal and filtered and amplified to obtain the heart rate and pulse information. Among them, the absorption of light by the skin and muscle tissue remains constant throughout the blood circulation process. When the heart contracts, the blood content in the blood vessels is the highest at this time, the absorption of light is the highest, and the light intensity returning to the receiving tube is the smallest; On the contrary, when the heart relaxes, the blood content in the blood vessels is the lowest, the light absorbed is also the least, and the intensity of the light returning to the receiving tube is the largest. The intensity of the light received by the light receiver changes pulsatingly. The current method of measuring pulse wave with infrared sensors usually adopts photoplethysmography, which is a measurement method that uses optical means to indirectly obtain pulse information by detecting changes in blood volume. In the smart bracelet, the most widely used is the reflective measurement structure, so this article uses a photoelectric sensor for experiments. The analog information is converted into an analog voltage signal through the current-limiting resistor R4 and sent to the LM393 comparator, which can be divided by the pin 2 of the comparator chip. Comparing the analog voltages obtained by the voltage divider, a square wave can be obtained. The schematic diagram of the internal circuit of the pulse sensor and the LM393 comparator is shown in Fig.2.

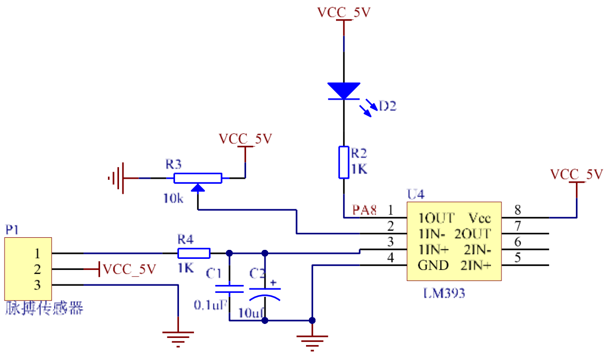


Fig. 2. The schematic diagram of the internal circuit of the pulse sensor and the LM393 comparator.

2.3. *Motion state detection module*

Devices such as accelerometers and gyroscopes are all inertial sensors. As the name suggests, accelerometers are devices that detect the acceleration of objects. The motion state of an object can be determined through accelerometers and gyroscopes. There are three types of acceleration sensors: piezoelectric, capacitive, and thermal. This article uses a three-axis capacitive acceleration sensor. The three-axis acceleration sensor has acceleration output in three directions of x-axis, y-axis and z-axis, and the output value forms a vector, which can detect the angle. A gyroscope is used to detect position information. The gyroscope is an angular motion detection device that uses a high-speed rotating body's momentum-sensitive shell relative to inertial space around one or two axes orthogonal to the rotation axis. The module has high accuracy and strong stability. The circuit corresponding to the gyroscope is shown in Fig.3.

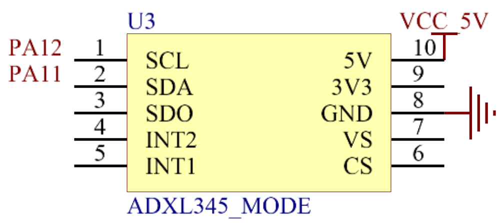


Fig. 3. The circuit corresponding to the gyroscope.

2.4. *Bluetooth module*

The embedded HC-05 Bluetooth serial communication module has two working modes: command response working mode and automatic connection working mode. AT commands can be executed when the module is in command response mode (or AT mode). Users can send various AT commands to the module, set control parameters for the module or issue control commands. In the automatic connection working mode, the module can be divided into three working roles: master, slave and loopback[4]. When the module is in the automatic connection mode, it will automatically connect according to the pre-set data transmission. Main mode: The module can actively search for and connect to other Bluetooth modules and receive and send data. Slave mode: It can only be searched and connected by other Bluetooth modules to receive and send data. Loopback: The Bluetooth module returns the received data to the remote master device as it is. The circuit connection schematic diagram of the Bluetooth module is shown in Figure 4.

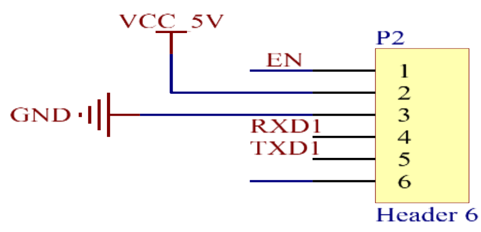


Fig. 4. The circuit connection schematic diagram of the Bluetooth module.

3. System total circuit design

This design is composed of STM32F103C8T6 microcontroller core board circuit + ADXL345 sensor circuit + heart rate sensor circuit + OLED circuit + Bluetooth module. The schematic diagram of the core board circuit is shown in Fig. 5.

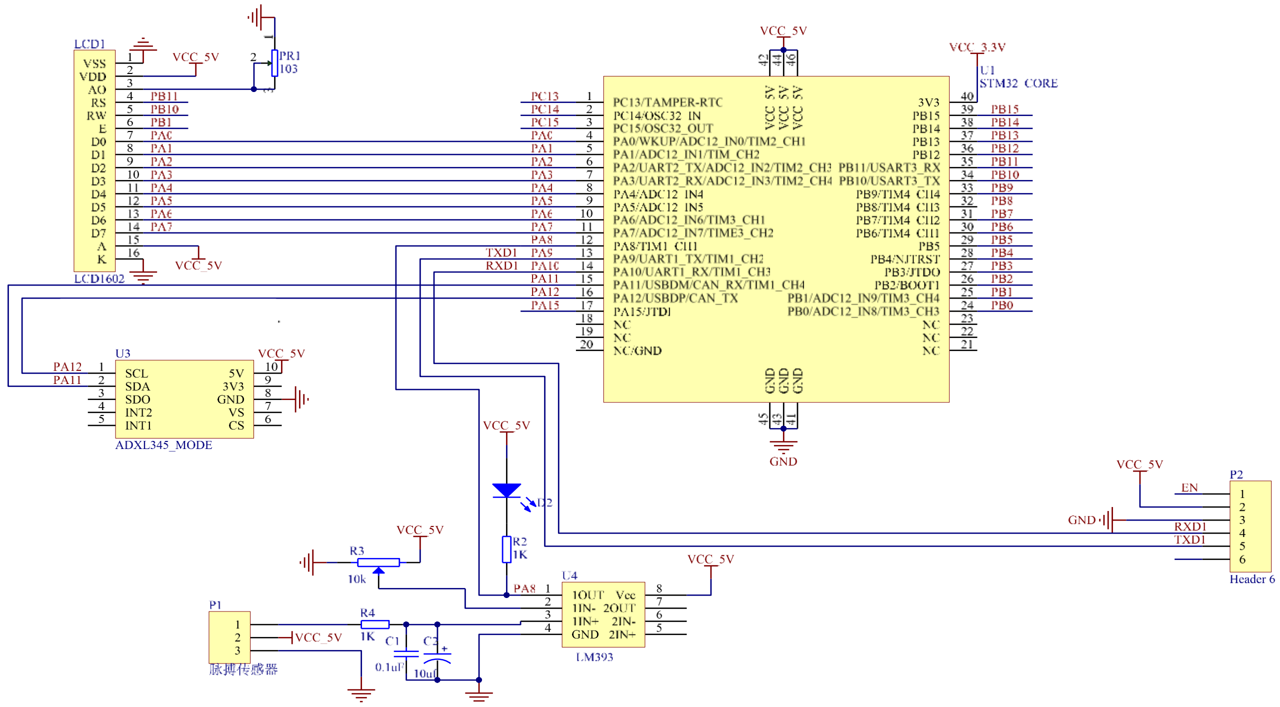


Fig. 5. The schematic diagram of the core board circuit

The actual product after circuit soldering is shown in Fig. 6.

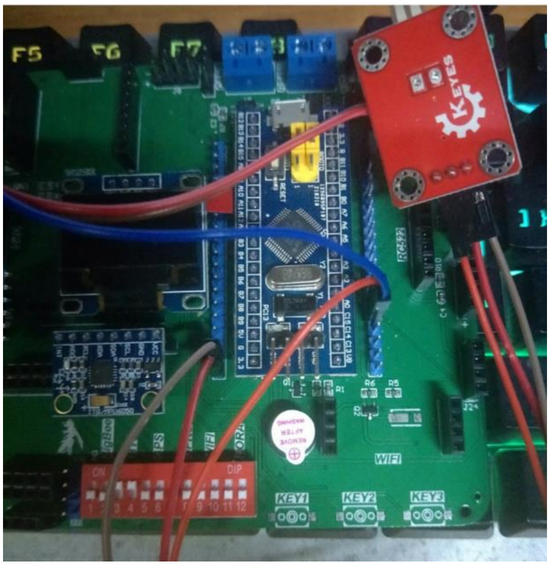


Fig. 6. The schematic diagram of the core board circuit.

4. Conclusion

The smart bracelet control system designed in this paper has been installed and debugged, and it has the characteristics of multiple functions, high data accuracy, and fast response speed, and has high use value.

References

1. Yang, T. , T. Huang , and L. Xie . "Design of Smart Bracelet Based on STM32 Micro-Controller."
2. Gu, Y., et al. "Design of Health Assistant Bracelet for The Elderly Based on STM32." 2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC) IEEE, 2019.
3. Yang, C. Y., and L. I. Qin. "Design and Implementation of Heart Rate Detection Based on STM32 Criminal Trial." Journal of Chongqing Technology and Business University(Natural Science Edition) (2016).
4. Petrovska, A.. "An embedded system for acces control with a Bluetooth module." (2015).

Authors Introduction

Ms. Xia Miao

She is currently studying for a postgraduate degree at Tianjin University of Science and Technology, majoring in electronic information.

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.

Dr. Jianliang Li

Jianliang Li, born in October 1973 Graduated from Analytical Instrumentation of Tianjin University in 1996, Associate Professor, Director of Teaching and Research Section.

The Design of Material Conveying and Automatic Sorting Control System Based on PLC

Qian Wang[1], Xiaoyan Chen[1*], Shengmin Cao[2]

[1] *Tianjin University of Science and Technology,Tianjin,China;*

[2] *TangShan College,TangShan,China*

*E-mail: 1248265505@qq.com, * cxywxr@tust.edu.cn, 1515513479@qq.com*

Abstract

The automatic sorting technology is widely used in all walks of life, such as logistics distribution centers, post offices, mining, port, etc. It can replace humans to do a lot of monotonous reciprocating or high-precision work. In order to liberate people's hands to do more challenging and innovative work and greatly improve work efficiency. This design uses programmable logic controller PLC and configuration software to design an automatic sorting control system. The PLC and MCGS present a friendly man-machine interface and powerful data management functions. The simulation results show that the system has high sorting efficiency and stable performance.

Keywords: automatic sorting, PLC, MCGS

1. Introduction

Before the emergence of automatic material sorting control system, the traditional logistics transportation basically relies on manual labor to carry goods. The traditional logistics transport of goods is slow, costs a lot of labor, and the accuracy of sorting goods is very low, but also brings huge labor costs to the enterprise. Now, with the advent of a variety of conveyor equipment, the entire conveyor system requires only a few people to ensure the normal operation of the entire system, which reduces the impact of manual sorting operations and effectively improves the sorting accuracy, so that enterprises reduce a lot of extra costs.

In this paper, a kind of automatic material sorting control system controlled by PLC is proposed. The automatic sorting control system uses Siemens PLC as the control device. A motor is used to drive the conveyor belt to rotate as a conveying device. Sensors are used to detect the material or color of the object, and send out signals. Receiving these signals, PLC controls the whole system based on those signals, then the system accomplishes the sorting of different materials or different colors of the block. Through running the system several times to detect objects of different materials, the results show that the automatic sorting system has the characteristics of high accuracy, high stability and high practicability. This system can also be replaced with corresponding sensors to sort objects of different colors.

2. System design

2.1. Hardware selection

The Siemens minicomputer S7-200CPU226 contains 2,048 storage units, 24 inputs and 16 outputs. In addition, it can be connected to 7 expansion modules, which can be extended to 248 digital I/O points or 35 analog I/O points.

The control requirements of the system include four sensors (one inductive proximity sensor, two photosensitive sensors and one color sensor), four in-place signals (materials 1, materials 2, materials 5, materials 6) and four switching signals, altogether 12 signals. The output signals include the movement of two conveyor belts, the movement of four pushrods, and a total of 18 signals are input and output. According to the I/O points, selecting a general minicomputer can meet the control requirements of the system, so the PLC of the system is S7-200CPU226. The address distribution table of PLC input and output in material sorting system is shown in Table 1 and Table 2.

Table 1 PLC input address distribution table

	address	function
input	I0.0	start
	I0.1	stop
	I0.2	inductive proximity sensor
	I0.3	light-sensitive sensor
	I0.4	material 5
	I0.5	material 6
	I2.0	start 1
	I2.1	stop 1
	I2.2	color sensor
	I2.3	light-sensitive sensor 1
	I2.4	material 1
	I2.5	material 2

Table 2 PLC output address distribution table

	address	function
output	Q0.1	pushrod 1
	Q0.2	pushrod 2
	Q0.0	belt
	Q0.3	pushrod 3
	Q0.4	pushrod 4
	Q0.5	belt 1

2.2. Introduction to MCGS

MCGS[1] is a kind of configuration software system that can quickly construct and generate upper computer monitoring system, which mainly completes field data collection and monitoring, front-end data processing and control. It can be run on a variety of platforms, and then presented to the user in a variety of ways (way: animation display, report display, etc.) by collection and processing of field data, to facilitate users clearly know where the problem is, timely treatment of the problem. It is widely used in the field of industrial control.

In this paper, PLC and MCGS are combined and communicated, and the configuration screen of automatic sorting system is simulated, so that users can understand the control process of the system more easily.

2.3. The overall structure of the system

The material sorting control system consists of two production lines, each of which includes a conveying module, a feeding module and a sorting module[2,3]. Production line 1 sorts objects according to their materials. The transmission module is composed of a linear conveyor belt driven by a DC motor. The blanking module is composed of a blanking trough. The sorting module consists of a push rod, a light-sensitive sensor, and an inductive proximity sensor. Production line 2 is to sort objects according to color, and the conveying module is composed of a straight conveyor belt driven by a DC motor. The blanking module is composed of a blanking trough. The sorting module consists of a push rod, a photosensitive sensor, and a color sensor. The structure diagram of the material sorting system is shown in Fig.1 and Fig.2.

The process of production line 1 in material conveying and automatic sorting system is as follows. The first step is to put objects of different materials into the hopper, and the conveyor belt motor drives the conveyor belt to rotate. The second step, the object is successively pushed onto the conveyor belt. When the object approaches sensor 1, the inductive proximity sensor quickly detects whether the material of the object block is iron. The third step, after the detection is completed, the conveyor belt carries the object to the corresponding position of the push rod. Finally, the corresponding putter moves quickly to push the block out of the conveyor belt. After the conveyor belt idles for 1min, the conveyor belt motor stops running. The flow chart of production line 1 is shown in Fig.3.

The process of production line 2 is as follows. The first step is to put objects of different colors into the hopper, and the conveyor belt motor drives the conveyor belt to rotate. In the second step, the objects are pushed onto the conveyor belt in turn. When the object approaches sensor 3, the color sensor quickly detects whether the block of color is red or not. The third step, after the detection is completed, the conveyor belt carries the object to the corresponding position of the push rod. Finally, the corresponding putter moves quickly to push the block out of the conveyor belt. After the conveyor belt idles for 1min, the conveyor belt motor stops running.

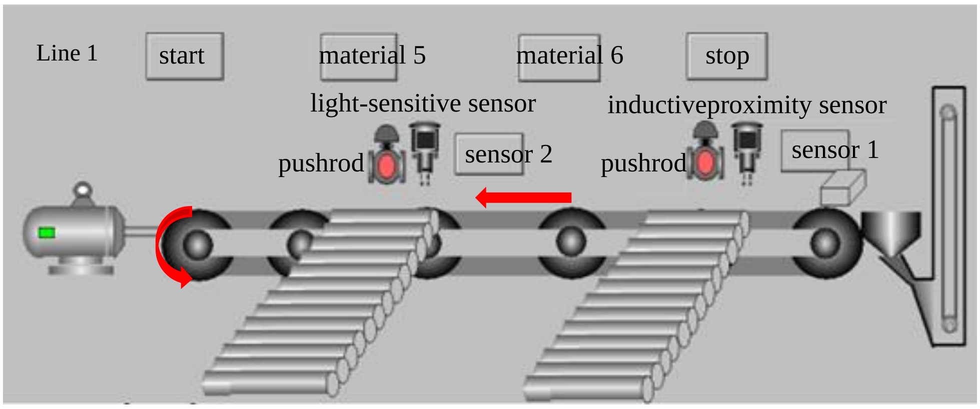


Fig.1. Simulation of sorting objects by material

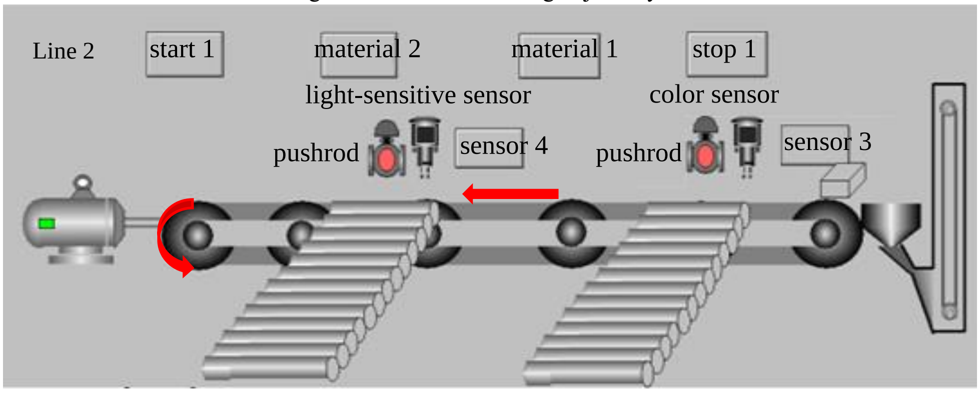


Fig.2. Simulation of sorting objects by colors

3. Results

To reflect the state change of simulation through the graphics in the monitoring screen, PLC and MCGS need to establish a communication interface through the device window. MCGS relies on the device window to establish a connection channel with the PLC, which can read external data and send data to the outside, and realize system monitoring. The following are the steps to design the connection between PLC and MCGS. The first step is to add "Siemens S7-200-PPI Universal Serial Port Parent Device" in the device toolbox. The second step is to edit the device properties under the universal serial port parent device. Acquisition optimization is 0-no optimization. The device name is device 0. The device annotation is Siemens_S7200PPI. The initial working state is 1-start. The minimum acquisition period is 1000ms. The device address is 2. The communication waiting time is 500ms. The number of fast acquisitions is 0. The data collection method is 0-block collection. The third step, by increasing the equipment channels of PLC and MCGS, and connecting the channels, the data between PLC and MCGS can be connected correspondingly.

In this paper, PLC is used to complete the control of the whole system module, and the MCGS software is used as the input control tool, to realize the automatic sorting of objects of different materials or different colors. The simulation results show that the system has high sorting accuracy, stable operation and high sorting efficiency.

4. Conclsion

The automatic material sorting system is based on Siemens S7-200CPU226 PLC as the control device, the system is mainly composed of a transmission module, a feeding module, and a sorting module. Each module is a small electromechanical system based on pneumatic devices as the execution unit, and a variety of sensors are installed throughout the system to provide signals such as the location, material and color of the objects. However, the automatic material sorting system has certain requirements on material packaging and material color, which is also a problem to be solved in future research work.

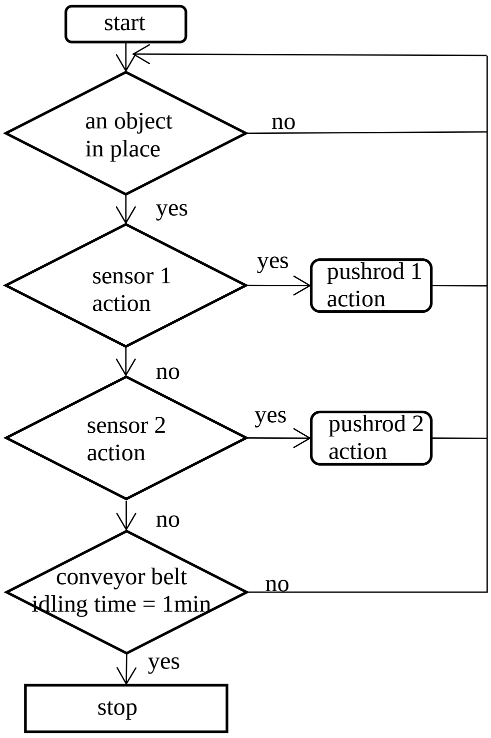


Fig.3. The flow chart of production line 1

References

1. Weidong Liang, et al. Design and application of MCGS configuration software. Guangzhou University. School of Physics and Electronic Engineering, vol. 1, no.26, pp. 33-35,2005.
2. Dai Jun. Design of Automatic Material Detection and Sorting System based on PLC control. Master of Engineering Dissertation, Dalian University of Technology, 2015.
3. Ying Lu. Design of Material Transfer and Sorting control system based on PLC. Coal Mine Machinery, vol. 5, no.37, pp. 187-189, 2016.

Authors Introduction

Ms. 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.

Prof. Xiaoyan Chen

She is 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.

Dr. Shengmin Cao

She(1515513479@qq.com) was born in 1965. She is now an associate professor in Tangshan University, mainly engaged in teaching and application of automation major.

Image Reconstruction Based on ResV-Net for Electrical Impedance Tomography

Qian Wang, Zichen Wang, Di Wang, Xiaoyan Chen*

Tianjin University of Science and Technology,Tianjin,China

E-mail: 1248265505@qq.com, wangzc@mail.tust.edu.cn, 20807936@ mail.tust.edu.cn,

** cxywxr@tust.edu.cn*

Abstract

Electrical impedance tomography (EIT) is a nonlinear and ill-posed inverse mathematical problem. Due to the above problem, the reconstruction image suffers from serious artifacts. To overcome shortcomings, we proposed a residual V-shaped deep convolutional neural network (ResV-Net). It consists of the feature extraction module and image reconstruction module which are optimized by ResBlock. The residual connection method can effectively increase the number of the forward information flow and reverse gradient flow in deep CNN and alleviate the problem of non-convergence caused by gradient vanishing. The simulation and experimental results show that the ResV-Net has a better visualization effect than the related imaging method.

Keywords: electrical impedance tomography, inverse problem, V convolutional neural network, image reconstruction

1. Introduction

Electrical impedance tomography (EIT) is a new visual measurement technology. EIT takes advantage of the electrical characteristics of different media with different conductivity in the area. The potential distribution on the observation area boundary are measured by applying an alternating current excitation to the measurement field, and the conductivity distribution are reconstructed by using appropriate imaging algorithm. However, due to the ill-posed and nonlinearity of image reconstruction problems, most current algorithms approximate nonlinear problems to linear problems based on sensitivity matrix theory, which results in the loss of part of boundary information and the interference of some artifacts in reconstructed images. The Landweber[1] iterative method and regularization method[2] proposed in recent years have relieved the influence of nonlinearity and ill-condition on image reconstruction to a certain extent, but the spatial resolution of reconstructed images still needs to be improved.

At present, a convolutional neural network has given many satisfactory results in image processing of inverse problems, indicating that the deep neural network model has a good performance in expressing complex nonlinear relations. To improve the quality of reconstructed images effectively, an electrical impedance tomography method based on a deep convolutional neural network is proposed in this paper. In this method, feature extraction and image reconstruction of the V-Net network are adopted, and a residual connection information transfer channel is added into each block. The deep learning model of this supervised mode is named ResV-Net. This strategy can effectively alleviate the gradient disappearance problem that may occur in the deep CNN model, as well as reduce the difficulty of network training and improve the robustness and generalization ability of the network. The ResV-Net is used to reconstruct the image with fewer artifacts, higher spatial resolution and more clear boundary of inclusions.

2. Method

2.1. Mathematical method model

The working model of current excitation by adjacent electrodes is used as the EIT data collection mode. A pair of adjacent electrodes are selected to inject current excitation, and another two adjacent electrodes are selected to measure the voltage as the measurement electrode. Sixteen pairs of adjacent electrodes are successively used as excitation electrodes, and a total of 208 measurements are collected. The EIT simulation model measurement system is shown in Fig.1.

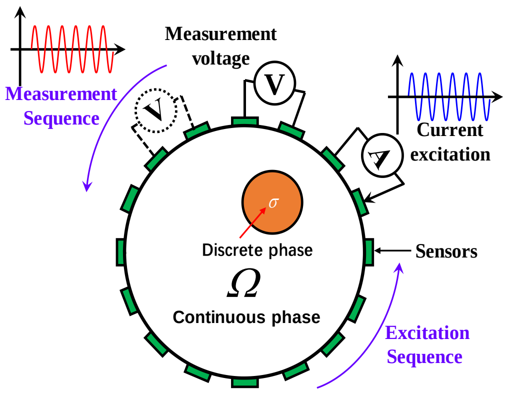


Fig.1. EIT simulation model measurement system

According to Maxwell equation[3], Neumann boundary conditions and complete electrode model[4], the mathematical model of EIT can be expressed as

$$\begin{cases} \nabla\cdot\sigma(\boldsymbol{r})\nabla\varphi=0, \boldsymbol{r}\in\Omega, \\ \varphi(x,y)=u(x,y),(x,y)\in\partial\Omega\setminus\bigcup_{L=1}^{16}e_L, \\ \varphi(x,y)+\sigma(r)\nabla_n\varphi(x,y)=U, \\ (x,y)\in e_L, L=1,2,\cdots,16. \end{cases} \tag{1}$$

Where $\boldsymbol{r}$ is the spatial position, $\sigma(\boldsymbol{r})$ is the conductivity distribution in the measurement field, $\varphi(x,y)$ is the potential distribution in the measurement field, $\boldsymbol{n}$ is the boundary unit normal vector, $\boldsymbol{\Omega}$ is the measurement field, e_L is the boundary attached electrode.

2.2. ResV-Net model structure

U-net has made a lot of satisfactory achievements in the field of computer vision in recent years, but the topological structure of the U-net cannot be directly applied to EIT imaging. Based on the above reasons, a V-shaped convolutional neural network with compression path as encoder and expansion path as the decoder isproposed in this paper. At the same time, in order to alleviate the problem of gradient disappearance caused by many hidden layers in the network, a residual connection is added in each module to avoid this problem. The network is named ResV-Net. It is composed of the feature extraction module and image reconstruction module. The feature transfer channel of residual connection is used between each block, and "skip connection" is used to connect the feature graph of the same dimension in the encoding module and decoding module. Residual information transmission mode is added in each block, and this connection method is made between of low-level features and high-level features. ResV-Net can improve effectively the training efficiency of the network and encourage the repeated use of features, as shown in Fig.2.

L1~L20 is a feature extraction module with the encoding function. The 16×16 voltage matrix is obtained by measurement sensors, and the input of ResV-Net is 32 × 32 pixel distribution which up-samples from the voltage matrix. Each block has two convolution layers with the 3×3 kernal and one residual connection layer realized by 1×1 convolution. The residual connection channel and the feature mapping layer are activated by 3 × 3 convolution and ReLU nonlinear function. The adjacent block are connected by 2×2 max-pooling layer, so that the complex image information can be expressed with less feature distribution. After feature extraction of 5 blocks, the 32×32 input features are mapped to a 2×2 feature map, and the number of feature channels is expanded from 1 to 256.

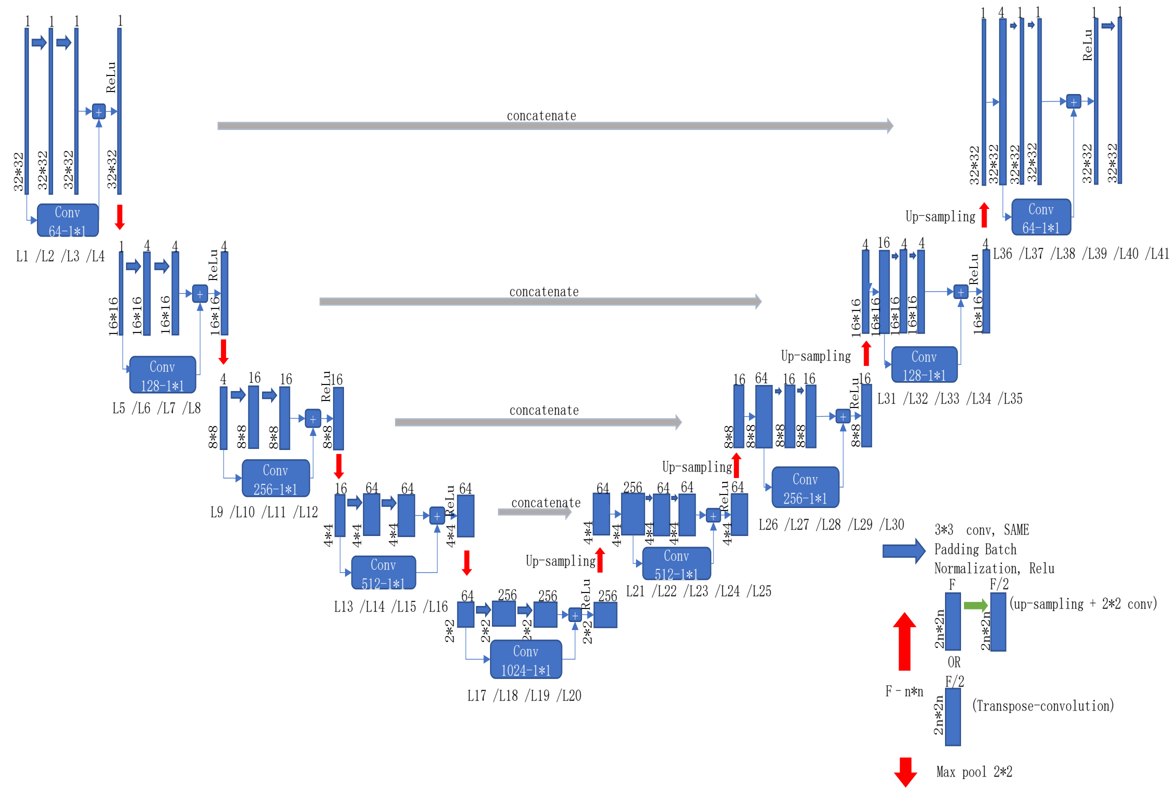


Fig.2. ResV-Net network topology

L21~L41 are image reconstruction module with the decoding function. Each block contains one feature compression layer, two convolution layers with the 3×3 kernal, one residual connection channel, and one feature mapping layer. The feature mapping layer is activated by 3×3 convolution and ReLU nonlinear function. An up-sampling layer with a convolution kernel of 2×2 was used to reconstruct a feature map of larger size between two adjacent blocks. After feature reconstruction of four blocks, the output feature is a feature map with the same size 32×32 as the input information, and the number of channels decreases gradually from 256 to 1. Finally, a convolution kernel with a convolution layer of 1×1 is used to eliminate the "checkerboard" effect caused by up-sampling, and the final result is taken as the output result of the network.

In ResV-Net, input features in each block are used to extract boundary features of parameter distribution by 1×1 convolution, and the constructed abstract features are fused with the output of a 3×3 convolution using a residual connection. The structure is shown in Fig.2. Meanwhile, the features of L4, L8, L12 L16 were spliced with L36, L31, L26 L21 respectively using a skip connection. Such connection mode, on the one hand, increases the repeated utilization of features and makes the deep network can learn more data distribution features. On the other hand, multi-channel feature connection can increase forward information flow and reverse gradient flow, avoid the problem of gradient disappearance in the deep network, make network training easier and increase the use efficiency of parameters.

2.3. Model training

The ResV-Net requires a large number of data samples, and the quantity and quality of data samples will affect the robustness and generalization ability of the network. During the simulation, a tank with a radius of 0.095m is set, and 16 electrodes are attached on the boundary of the tank with the same spacing and height. The water with the conducting of 0.6s /m is set as the homogenous background, and the conductivity of inclusions is set as the inhomogenous media. The finite element method in COMSOL Multiphysics is used to divide the observation domain into discrete triangular mesh, which solves the forward problem and obtains the boundary voltage. COMSOL Multiphysics and MATLAB are used to divide the measurement field into 32×32 mesh, and each pixel corresponded to the distribution of conductivity values in the field. The number of inclusion in the simulation database ranges from 1 to 4. The positions and sizes of inclusions do not overlap each other and are randomly set.

Due to the depth and complexity of the ResV-Net, a loss function is used to calculate the error between the predicted result $\hat{x}$ and the real result x. The loss function[7] is

$$Loss = \alpha \times l(\hat{x}, x) + \lambda \times \|\theta\|^2 \qquad (2)$$

Where, $l(\hat{x}, x)$ represents the cross entropy of $\hat{x}$ and x. $\|\theta\|^2$ is a regular term, which can be used to constrain the solution of parameters to avoid the over-fitting problem of training. $\hat{x}$ and x are important optimization methods in the network. $\|\theta\|^2$ is auxiliary constraint factor, and their weight coefficients are α = 1 and λ = 0.01 respectively. The ResV-Net employs the mini-batch gradient descent method in training samples, which is expressed as

$$\theta_{t+1} = \theta_t - \eta \bullet \nabla_{\theta_t} Loss(\theta_t; V_{n,m+n}, \sigma_{n,n+m}) \qquad (3)$$

The value of m is 120, and the initial learning rate is 0.01.

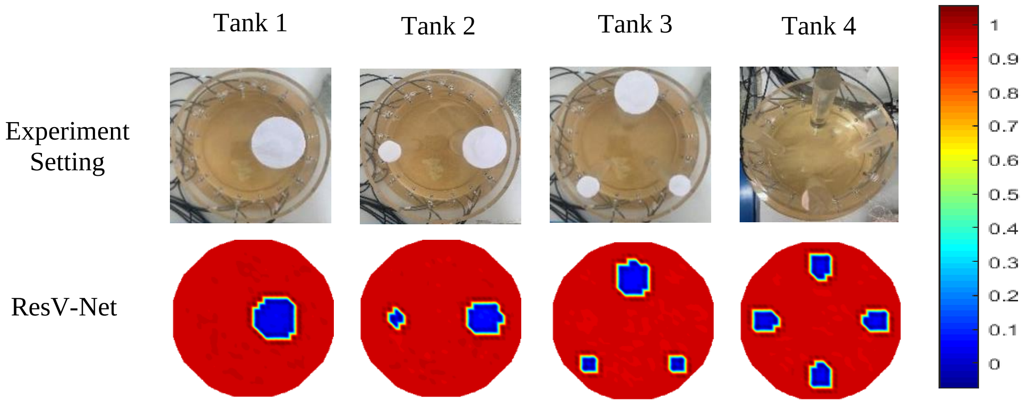


Fig.3. Simulation imaging results of RESV-NET network model

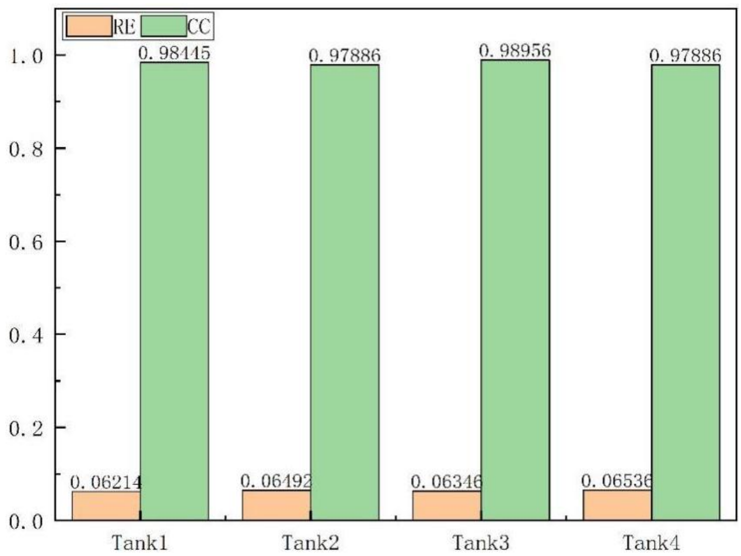


Fig.4. Evaluation index of simulation imaging of ResV-Net network model

The learning rate is reduced to 0.9 times before each update. This process is repeated until the ResV-Net converges. By using the mini-batch gradient descent method to update the parameters of the network, the loss function can reach the local optimum.

3. Results

The relative error (RE) and correlation coefficient (CC) are two general metrics for quantitative evaluation of reconstructed images[5,6].

In order to verify the performance of the proposed ResV-Net network model, the data in the test data set is used as the input of the network model for training. The training results are compared with the results of image reconstruction obtained by the commonly used algorithms such as NOSER, TR, CG and CNN. The comparison results verify the nonlinear ability of the ResV-Net network model to deal with EIT inverse problems. The ResV-Net model simulation imaging results are shown in Fig.3, and RE and CC are shown in Fig.4.

Fig.3 and fig.4 show that the ResV-Net has high RE and low CC, network in each image reconstruction using the characteristics of a residual connection between block transmission channels, at the same time use "skip connection" coding module and decoding module connection, the characteristics of the same dimension figure in the reconstruction image boundary clear, high spatial resolution and good visualization.

4. Conclusion

For EIT image reconstruction, the ResV-Net is proposed in this paper. The network consists of feature extraction module and image reconstruction module. The network adopts the feature transfer channel of residual connection, and uses the "skip connection" method to connect the feature maps of the same dimension in the feature extraction module and the image reconstruction module. It avoids the problem of gradient disappearance which may occur in the deep network, and increases the repeated utilization rate of features. The ResV-Net is superior to traditional imaging methods in terms of reconstructed image quality and quantitative evaluation indexes. And using this method in simulation and experiment can obtain good imaging results. However, the application of deep learning in image reconstruction still faces many challenges, such as incomplete simulation databases. This is also one of the problems to be solved in future research work.

References

1. Feng Li, Feng Dong and Chao Tan,"landweber iterative image reconstruction method incorporated deep learning for electrical resistance tomography," IEEE Transactions on Instrumentation and Measurement, vol.70, pp.1-11, 2020.
2. Yanbin Xu, Yang Pei, Feng Dong. "An adaptive Tikhonov regularization parameter choice method for electrical resistance tomography," Flow Measurement and Instrumentation, vol. 50, pp. 1-12, 2016.
3. Kuosheng Cheng, D. Isaacson, J. C. Newell, et al., "Electrode models for electric current computed tomography," in IEEE Transactions on Biomedical Engineering, vol. 36, pp. 918-924, 1989.
4. Erkki Somersalo, Margaret Cheney, David Isaacson, "Existence and Uniqueness for Electrode Models for Electric Current Computed Tomography," SIAM Journal on Applied Mathematics, vol. 52, pp. 1023-1040, 2016.
5. Feng Li, Chao Tan, Feng Dong, et al., "V-net deep imaging method for electrical resistance tomography," IEEE Sensors Journal, vol. 20, pp. 6460–6469, 2020.
6. Feng Li, Chao Tan, and Feng Dong, "Electrical resistance tomography image reconstruction with densely connected convolutional neural network," IEEE Transactions on Instrumentation and Measurement, vol. 70, pp.1–11, 2021.
7. Xiaoyan Chen, Zichen Wang, Di Wang, et al., "Deep Autoencoder Imaging Method for Electrical Impedance Tomography," in IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1-15, 2021.

Authors Introduction

Ms. 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.

Mr. Zichen Wang

He was born in Tianjin, China in 1997. He received the B.E. degrees from Tianjin University of Science and Technology, Tianjin, China, in 2019. He is currently pursuing the M.S. degree in Information and Automatic College at Tianjin University of Science and Technology, Tianjin, China.

His research interests include signal/image processing, digital signal/image processing, intelligent algorithms, electrical tomography (especially for industrial and clinical applications of electrical impedance tomography), numerical solutions of inverse problems, finite/boundary element methods, and deep learning.

Mr. Di Wang

He was born in Changchun, Jilin Province, China. He graduated from Changchun University of Science and Technology with a B.E. degree in engineering. He is currently pursuing the M.S. degree in Information and Automatic College at Tianjin University of Science and Technology, Tianjin, China.

His research interests include electrical impedance tomography (EIT), instrumentation and measurement, data acquisition systems, circuit and system design, numerical algorithms for inverse problems and other imaging sciences. Currently, he is working on a resolution-enhanced EIT imaging system for continuous human lung monitoring using field programmable gate array (FPGA) and a high-precision data acquisition system.

Prof. Xiaoyan Chen

She is 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.

An Estimation Method of Coastal Ocean Debris Using Aerial Drone

Kazuo Ishii
Dept. of Human Intelligence Systems, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka, 808-0196, Japan
E-mail: ishii@brain.kyutech.ac.jp

Kanako Shirahashi
Center for Socio-Robotic Synthesis, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka, 808-0196, Japan
E-mail: y-nishida@brain.kyutech.ac.jp

Yuya Nishida
Dept. of Human Intelligence Systems, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka, 808-0196, Japan
E-mail: ishii@brain.kyutech.ac.jp

Moeko Tominaga
Department of Integrated Systems Engineering, Nishinippon Institute of Technology
1-11 Aratsu, Kanda, Miyako, Fukuoka 800-0349, Japan

Yoshiki Tanaka
Dept. of Human Intelligence Systems, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka, 808-0196, Japan
E-mail: ishii@brain.kyutech.ac.jp

Dominic B. Solpico
Dept. of Human Intelligence Systems, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka, 808-0196, Japan
E-mail: ishii@brain.kyutech.ac.jp

Abstract

The actual situation of marine litter has not been measured accurately; however innumerable floating garbage are drifting in the ocean. Especially, non-perishable waste such as microplastics continues to grow and is damaging marine life, including endangered species, and some are washed ashore and causing pollution damage to coastal areas. Microplastics incorporated into marine organisms, Arctic Sea ice, and deep-sea seafloor sediments have also been detected. The Ellen MacArthur Foundation in the United Kingdom estimates that the total amount of marine debris exceeds 150 million tons, with more than 8 million tons of new inflow each year. We measured and compared the amount of ocean debris in coasts in Hirado and Matsuura cities, Nagasaki with manual count and an aerial drone observation.

Keywords: Ocean debris, Arial drone, Coastal cleaning

1. Introduction

United Nations had announced A/RES/70/1, Resolution adopted by the General Assembly on 2015 [1] on 17 Sustainable Development Goals (SDGs) with 169 associated targets which are integrated and indivisible to encourage world leaders pledged common actions into a broad and universal policy agenda. UN showed the path towards "sustainable development, devoting ourselves collectively to the pursuit of global development and of "win-win" cooperation which can bring huge gains to all countries and all parts of the world". In the 17 SDGs, the goal 14 describes on conserve and sustainably use of oceans, seas and marine resources for sustainable development. The goal 14.1 is set to prevent and significantly reduce marine pollution of all kinds, in particular from land-based activities, including marine debris and nutrient pollution by 2025, and the goal 14.2 is set to sustainably manage and protect marine and coastal ecosystems to avoid significant adverse impacts, including by strengthening their resilience, and take action for their restoration in order to achieve healthy and productive oceans by 2020.

In the G20 Osaka Leaders' Declaration [2], the meeting mentioned the importance of measures to address marine litter, especially marine plastic litter and microplastics, necessity to be taken nationally and internationally in order to encourages to swiftly take appropriate national actions for the prevention and significant reduction of discharges of plastic litter and microplastics to the oceans. as a common global vision, Japanese government proposed the "Osaka Blue Ocean Vision [3]" that aims to reduce additional pollution by marine plastic litter to zero by 2050 through a comprehensive life-cycle approach, which includes reducing the discharge of mismanaged plastic litter by improved waste management and innovative solutions while recognizing the important role of plastics for society.

Toward the "Osaka Bule Ocean Vision", Ministry of Foreign Affairs of Japan established an association, Japan's MARINE Initiative [4] to advance effective actions to combat marine plastic litter at a global scale focusing on **Ma**nagement of wastes, **R**ecovery of marine litter, **In**novation, and **E**mpowerment. Under the MARINE initiative, Japanese government started to support empowerment in developing countries to promote waste management, recovery of marine litter, and innovation.

As grass-roots movements, many researchers and groups have been working on the study of quantitative and qualitative observation of marine litters on coasts. Yamaguchi has investigated the marine litters quantitatively in a lot of Japanese coasts for long time [5]-[7]. In ref. [7], Yamaguchi discussed on the actual situation of coastal pollution by the drifted garbage in the Japan Island. They performed investigations of coastal environment at a large number of seashores in the Sakishima Is. (Okinawa), Honshu and Hokkaido districts, and counted the number of garbage drifted in seashore was counted, and the garbage was also categorized according to its type and nationality. It was pointed out that the problem of coastal pollution by the drifted garbage was a serious problem of environment in Japan and strongly requested both the prevention measure and the disposition method of the drifted garbage. And also they proposed a measurement method to quantitate the mount of microplastic on coast by sampling coastal soils.

We had started the investigations of marine litters in a coast of Matsuura city (Fig.1, Coast (A) and (B)) and Hirado city (Fig.1, Coast (C)) using aerial a drone and compared with sampling, and report the results of the survey in this paper.

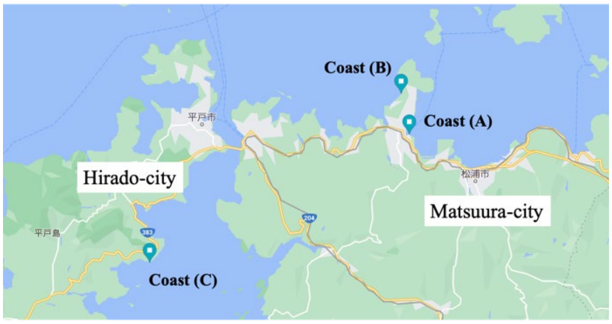


Fig. 1 The target coasts in Nagasaki, Kyushu Island. Coasts (A) and (B) are in Matsuura city and Coast (C) is in Hirado city. The coasts (A) and (B) direct for north direction, and coast (C) for south east. The basic map is from Google map.

2. Investigation of marine litter

2.1. *Coast of Matsuura city: Coast (A)*

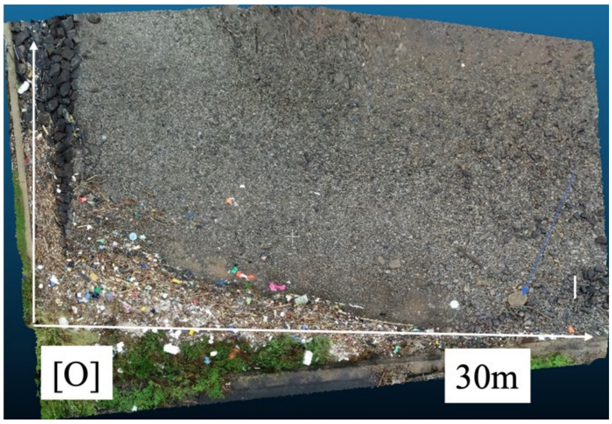
The first investigation of the target coast in Matsuura city: coast (A) in Fig.1 was performed on September 2020. The coast (A) is open for north direction and the garbage is drifting from Asian countries, especially a large amount of marine litter arrives in winter seasons. As shown in the photo of coast (A) in Fig. 2, the coast was covered by a large amount of garbage with many color and shape variations.

Firstly, the process to generate coastal map is introduced using Fig. 3. In this observation, a drone took a video from air, then the video is divided to 70 pictures so as to overlap 90% of each other of continuous pictures (Fig. 3(a)). The series of pictures are converged to a 3D-model (Fig. 3(b)) using Metashape v1.7 from Agisoft [8]. Agisoft Metashape [8] is a software product that performs photogrammetric processing of digital images and generates 3D spatial data to be used in GIS applications, cultural heritage documentation, and visual effects production as well as for indirect measurements of objects of various scales. The software calculates the feature points, generates a point cloud, make meshes and texture maps in 3D and its orthomosaic image. In the Fig. 3(b), we can see the stone area, garbage's area and walls with height in 3D. After the generation of 3D models, a free software CloudCompare [9] is used to evaluate and compare the maps. CloudCompare is a 3D point cloud

Fig. 2 The photo of coast (A) in Matsuura city, Sep./15th/2020. A large amount of marine litter arrived on the coast.

were used for the area of about 40m x 20m.

(b) The 3D-model of coast (A) using a software "Metashape" of Agisoft,.



(c) The obtained 3D model is adjusted using a free software CloudCompare based on reference points.

Fig. 3 The map of coast (A) is generated based on pictures obtain by a drone. To merge the pictures, a software Metashaps and CloudCompare are used.

(and triangular mesh) processing software, which has been originally designed to perform comparison between two dense 3D points clouds or between a point cloud and a triangular mesh. It relies on a specific octree structure dedicated to this task, and extended to a more generic point cloud processing software, including many advanced algorithms such as registration, resampling, color/normal/scalar fields handling, statistics computation, sensor management, interactive or automatic segmentation, display enhancement and so on.

Figure 4 shows an example of analysis using CloudCompare. The height from the plain surface which is based on stone area is visualized by colors. The marine litter drifted and mounted on the red-colored area. The 3D models are adjusted in the sizes, translation and angles of rotation transform matrix. Figure 5 is the 3D model after garbage clean up. By comparing the models of Fig. 3(c) and Fig. 5, the change of volume is estimated.

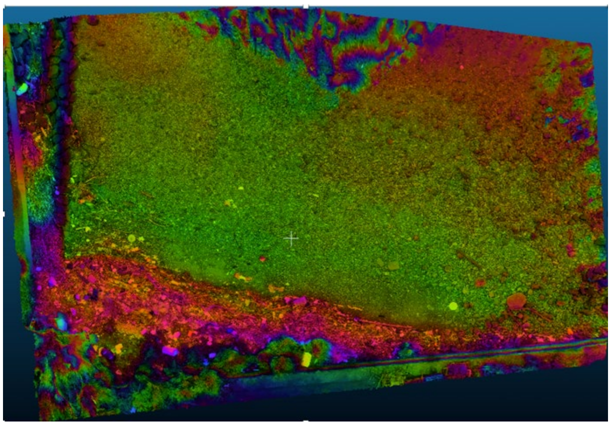


Fig. 4 An example of analysis using CloudCompare. The height from the plain surface is visualized by colors. The marine litter drifted and mounted on the red-colored area.

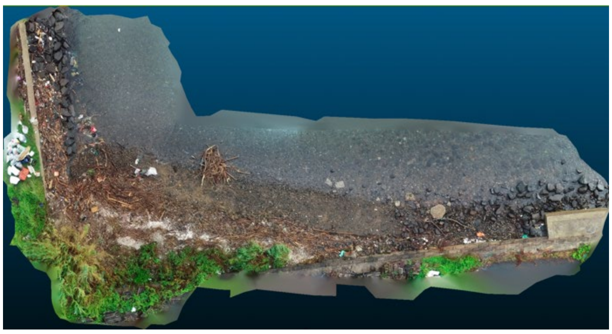


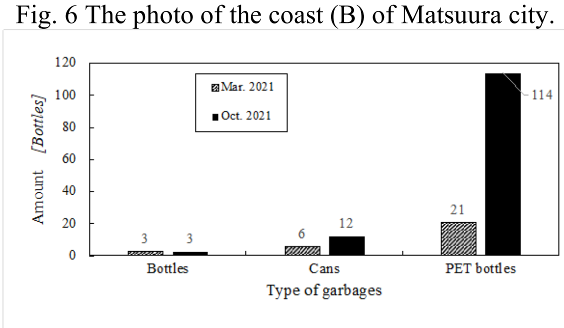
Fig. 5 The 3D model after clean up the coast. By comparing with the model in Fig. 3, the volume difference can be estimated.

This kind of analysis is utilized in the research of civil engineering to compare the soil volume and change after constructions [10].

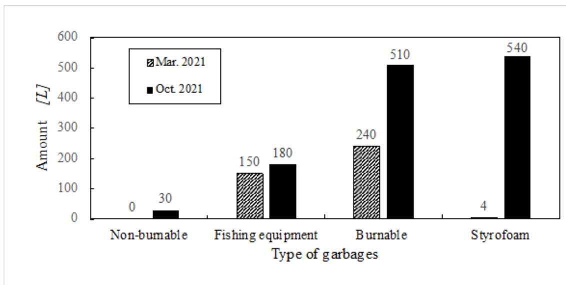
2.2. *Coast of Matsuura city: Coast (B)*

Figure 6 the photo of the target coast (B) in Matsuura city. We set 3 areas with 10m x 10m square and removed marine litters and count the number of bottles, plastics, fisher tools, Styrofoam. The left bars are results in the

Fig. 6 The photo of the coast (B) of Matsuura city.



(a) Amount of bottles



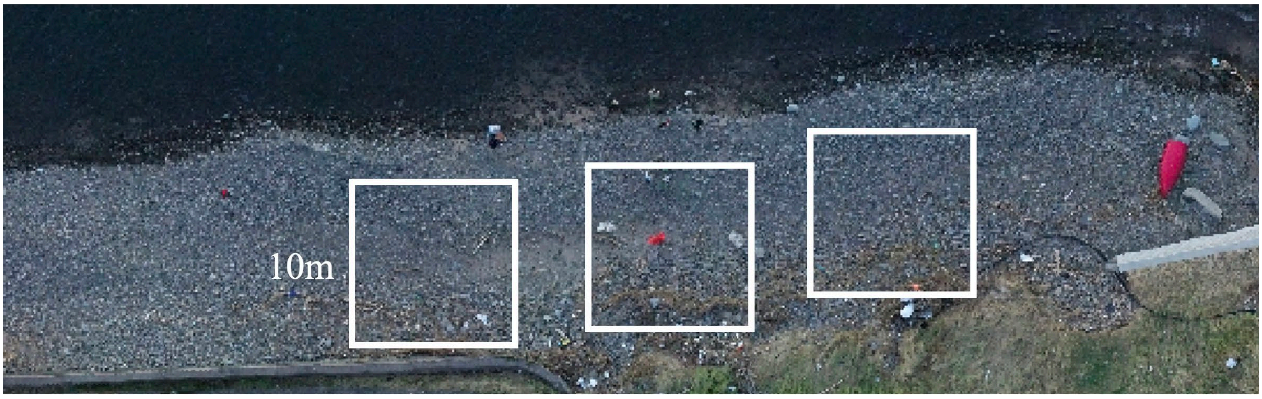
(b) Other garbage

Fig. 7 Types of bottled garbage washed ashore on the coast of Matsuura

March 2021 and the right October 2021. Figure 7(a) shows the numbers by categorizing bottles, cans and pet bottles and (b) are the non-burnable garbage, fishing equipment, burnable garbage, and Styrofoam measured in volume [L].

2.3. *Coast of Hirado city: Coast (C)*

The target coast (C) in Hirado city was also recorded by a drone and the 3D models before and after clean up are compared using the same software. The marine litter is measured in the 3 square areas by 10m x 10m in sizes. Comparing with the coasts of Matsuura city, the coast (C)



(a) Before clean up the coast

(b) After clean up the coast

Fig. 8 The photo of the coast (C) of Hirado city.

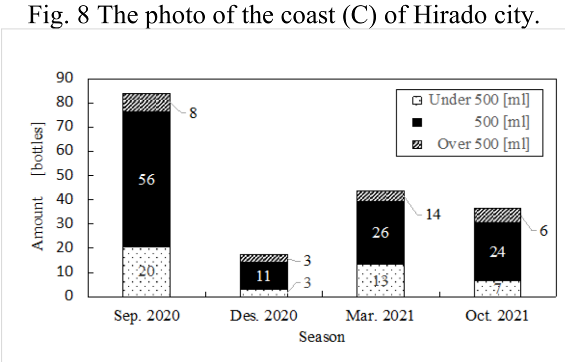


Fig. 9 Types of marine litter in the coast (C) Hirado city. The survey on Sep./2020 was after the typhoon, so that a large amount of garbage was on the coast.

directs in south east and the amount of marine litter is fewer than those of (A) and (B). Figure 8 shows the 3D models for comparison before and after clean up the coast.

3. Conclusions

We had started the investigations of marine litters in the coasts of Matsuura city and Hirado city using an aerial drone and compared with manual sampling. In order to make garbage maps of the coasts, we introduced the drone and measured the coast from air, and made 3D models using the software Metashape and CloudCompare. Japan is surrounded by ocean and huge amount of marine litter has arrived on the coasts. We should develop methods to visualize, digitize and quantilize the conditions of marine pollution.

.

4. References

1. Resolution adopted by the General Assembly on 25 September 2015, United Nations A/RES/70/1, 2015.
2. G20 Osaka Leaders' Declaration, on 28-29 June 2019.
3. Toward Osaka Blue Ocean Vision, https://g20mpl.org in Jan. 20th/2022.
4. Japan's "MARINE Initiative" toward Realization of the Osaka Blue Ocean Vision, https://www.mofa.go.jp/ic/ge/page25e_000317.html in Jan. 20th/2022
5. Hareyuki YAMAGUCHI, Coastal Pollution by Foreign Drifted Gabages, JSCE, Vol.83 (3), pp.60-62, 1998
6. Yamaguchi, H., Pongpo, Thunyawit, "Coastal Pollution by Drifted Garbage in Japan Island, Proceedings of the Symposium on Global Environment, 2000, Volume 8, Pages 111-120, 2011 (in Japanese)
7. Yamaguchi, H., "A large amount of microplastics verified in coastal areas where microplastic pollution is becoming more serious ",Proceedings of the 25th Symposium of Global Environment, pp.23-28, 2017 (in Japanese)
8. Metashape, https://www.agisoft.com
9. CloudCompare, http://www.cloudcompare.org
10. Lague, D., Brodu, N. and Leroux, J., "Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (N-Z)," ISPRE Journal of Photogrammetry and Remote Sensing, 82, pp. 10-26, 2013.

Acknowledgement

This research is partially supported by Grant-in-Aid for Scientific Research(B) 18H01643. Many Kyutech students in our lab helped as the volunteers to remove the garbage.

Authors Introduction

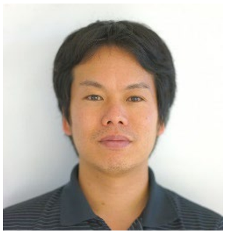
Prof. Kazuo Ishii

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.

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.

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.

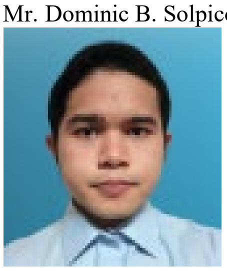
Dr. Moeko Tominaga

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.

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.

Mr. Dominic B. Solpico



Mr. Solpico obtained his M. S. degree in Electronics Engineering at the Ateneo de Manila University, Philippines in 2015. He is currently taking his Doctoral degree course at Kyushu Institute of Technology, Japan. His research interests are in the fields of Intelligent Aquaculture, Wireless Sensor Networks (WSNs), and Field Robotics.

Compact Ultra-Wideband Slotted Microstrip Patch Antenna for 5G, IoT and RFID Applications

Gurney Nga
Department of Electrical and Electronic Engineering
Faculty of Engineering, Technology & Built Environment
UCSI University, Kuala Lumpur, 56000, Malaysia

Mastaneh Moakyef
Department of Electrical and Electronic Engineering
Faculty of Engineering, Technology & Built Environment
UCSI University, Kuala Lumpur, 56000, Malaysia

Shahid Manzoor
Department of Electrical and Electronic Engineering
Faculty of Engineering, Technology & Built Environment
UCSI University, Kuala Lumpur, 56000, Malaysia

Manickam Ramasamy
Department of Electrical and Electronic Engineering
Faculty of Engineering, Technology & Built Environment
UCSI University, Kuala Lumpur, 56000, Malaysia

E-mail: mastaneh@ucsiuniversity.edu.my , 1001645290@ucsiuniversity.edu.my , shahid@ucsiuniversity.edu.my , ManickamRamasamy@ucsiuniversity.edu.my
www.ucsiuniversity.edu.my/

Abstract

With new emerging technologies in 5G era, the need of compact and cost-effective antenna holds significant promise for delivering contributions to society. The compact UWB omnidirectional microstrip patch antenna of 36mm × 32mm × 1.6mm for IoT, 5G and RFID applications is proposed and simulated. The proposed antenna covers frequency of 5.1256 GHz to 6.4391 GHz with UWB (bandwidth > 500MHz) characteristics. The proposed antenna has an FR4 substrate that made the antenna a cost-effective candidate in 5G era. The directivity of 5.3 dBi with gain of 3.51 dBi are obtained from the simulation results.

Keywords: Microstrip Patch antenna, UWB, ,5G, IOT, RFID

1. Introduction

Due to the rapid deployment of technology in people's lives and the users concern on their non-stop connections with different gadgets, the need of having a good antenna that can provide connections to the latest wireless technologies and providing a reliable connection between everything known as Internet of Things (IoT) under 5G's umbrella became at the center of attention of most of researchers [1].

5G or so called the fifth generation of wireless communication is the new revolution and evolution in a wireless field that speed up the advent of IoT. 5G is an essential infrastructure that allows all communications to be linked through the IoT at anytime and anywhere. The main advantages of 5G are providing larger capacity, higher data-rate, lower latency and higher application reliability. It is 1000 times more capacity, 100 times faster and 100 times more connectivity compared to 4G [1].

IoT has accelerated the development of many devices around us and has brought up possibility of many applications. Its vision is to build a fully connected environment to ensure more convenient and better human living by improving the issues like energy management, climate change, automation and transportation, healthcare and treatment, logistic management and other related fields [2].

Radio frequency identification (RFID) is a key technology for making the IoT to automatically detect objects by using wireless communication. In recent years, RFID has been widely used in many places, such as logistic management and production. RFID is one of the main IoT technologies and one of the ten most important technologies in the 21st century, as an automatic detection and data capture technology [3]. RFID uses frequency from 30 Hz to 5.8 GHz which depends on the applications [4].

According to the Federal Communications Commission (FCC) in February 2002, the specifications for UWB systems to use in the band of 3.1 GHz to 10.6 GHz and a large bandwidth of more than 500 MHz [5]. UWB provides higher data rates with very low radiation power. Hence, this technology is getting more significant and popular in wireless communication systems [6].

In this paper, the study and design of an UWB microstrip patch antenna is presented. Microstrip patch antenna is chosen over other types due to its compact size, light weight, high bandwidth, multiband properties, low cost and high gain [7]. The proposed antenna needs to cover three of the applications as well as meeting the requirements of UWB. The rise of the IoT has accelerated the development of many devices around us and has opened up the possibility of myriad applications. The IoT vision is to build a fully connected environment. This can be done by using smart objects and devices to produce data and transmit through the internet automatically for decision making purpose [7-9]. This is to ensure more convenient and better human living by improving the issues like energy management, climate change, automation and transportation, healthcare and treatment, logistic management, and other related fields. The IoT devices includes remote monitoring, tracking, collection of data, manufacturing and also for media applications. Nowadays, there are many IoT applications which are already identified globally i.e., smart city, smart home, smart logistics, smart transportation, smart healthcare and smart agriculture [8-10].

2. Antenna Design

The antenna is designed to cover frequency of around 5.8 GHz with UWB properties. The proposed antenna is fed by the microstrip line. The length and width of the antenna patch are calculated using the eq (1), eq (2) and eq (3) respectively.

$$\varepsilon_{eff} = \frac{\varepsilon_R+1}{2} + \frac{\varepsilon_R-1}{2}\left[\frac{1}{\sqrt{1+12(\frac{h}{W})}}\right] \quad (1)$$

$$Width = \frac{c}{2f_0\sqrt{\frac{\varepsilon_R+1}{2}}} \quad (2)$$

$$Length = \frac{c}{2f_0\sqrt{\varepsilon_{eff}}} - 0.824h\left[\frac{(\varepsilon_{eff}+0.3)(\frac{W}{h}+0.264)}{(\varepsilon_{eff}-0.258)(\frac{W}{h}+0.8)}\right] \quad (3)$$

The geometry of the microstrip patch antenna is shown in Figure 1. The FR-4 (dielectric constant = 4.4) is used as the substrate with a thickness of 1.6mm. The copper is used as the patch and its thickness is kept at 0.035mm. The rectangular patch has a size of 16mm × 12mm with a square slot of 2mm × 2mm in it. The antenna has two inset feeds with size of 1mm × 4mm. The feedline has a length of 10mm with 3mm width.

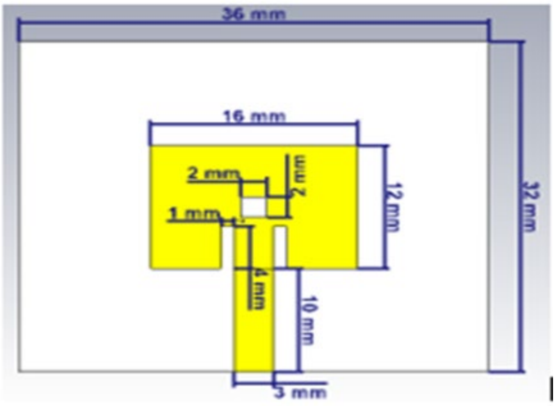


Fig. 1. Antenna dimensional parameters

3. Simulation Results

The antenna is simulated using Computer Simulation Technology (CST). A conventional rectangular patch is implemented on the top surface of FR-4 dielectric substrate in a half- length conductive ground. A rectangular $2 \times 2\,mm$ slot has been engraved at the center of the patch. The geometrical layout along with the dimensional parameters are depicted in Fig.1. The proposed slotted antenna exhibits an acceptable frequency response which is discussed in depth in the next section.

3.1. *Return Loss (S11)*

The S_{11} performance of the proposed antenna is simulated using CST Studio Suite 2019. Figure 2 illustrates the antenna resonates from 5.1256 GHz to 6.4391 GHz. The threshold value of -10 dB is taken as the base value which the mobile communication can operate perfectly [8]. At 5.8 GHz, the return loss with value of -17.927 dB has been obtained.

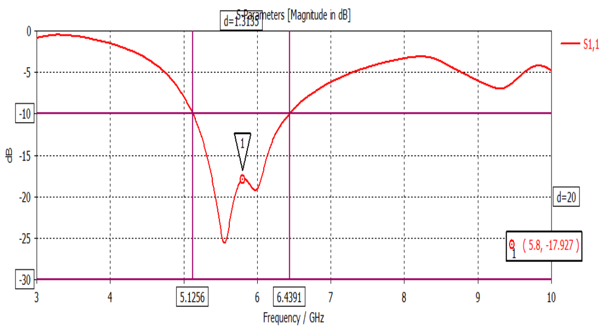


Fig. 2. Return loss at 6.4391 GHz.

3.2. *Voltage Standing Wave Ratio (VSWR)*

Figure 3 shows the VSWR of the proposed antenna. From the graph, it is observed that the VSWR value is below 2, from 5.1256 GHz to 6.4391 GHz. In fact, the VSWR value of below 2.5 is acceptable for most wireless applications [8]. If the VSWR increases, there will be more power reflected from the antenna, which is not considered as a good transmission.

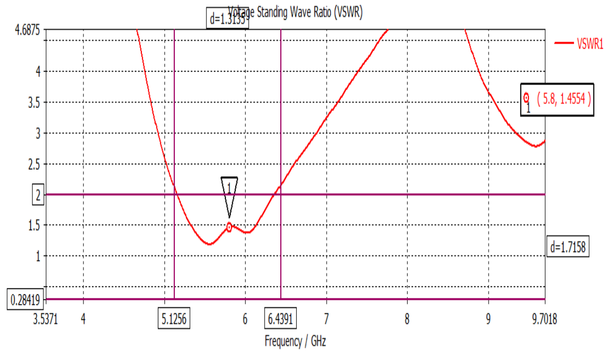


Fig. 3. VSWR value at 6.5 GHz

3.3. *Directivity of Antenna*

The directivity of the proposed antenna is 5.3 dBi at 5.8 GHz with a side lobe level of -8.1 dB, as depicted in Figure 4.

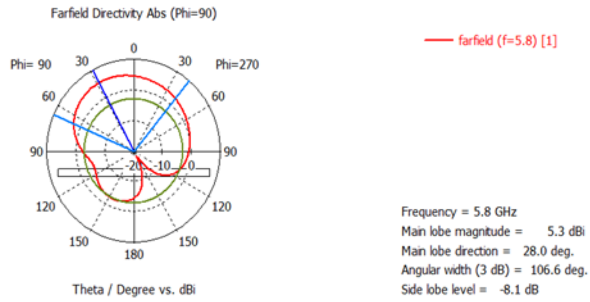


Fig. 4. Directivity of proposed antenna

3.4. *Antenna Gain*

Based on Figure 5, the gain of the proposed antenna is 3.51 dBi at 5.8 GHz, which is acceptable in terms of small size antenna. The gain is calculated by the ratio of the power emitted by it in a given direction with distance to the power emitted at the same distance by an isotropic antenna which is calculated as in (4).

$$gain = 4\pi \frac{radiation\ intensity}{total\ input\ power} \tag{4}$$

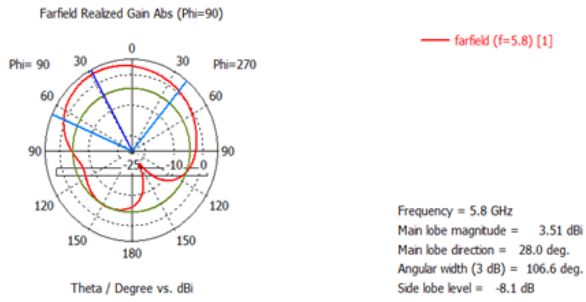


Fig. 5. Antenna gain at 6.4 GHz

3.5. *Radiation pattern*

Figure 6 shows the 3D radiation pattern of the proposed antenna. An omni-directional pattern is shown, which is good for mobile communication. The major lobe happens as the gain in the direction is high at its resonant frequency. The width of the main beam is between -3 dB. Low gain in the direction resulting in minor lobes.

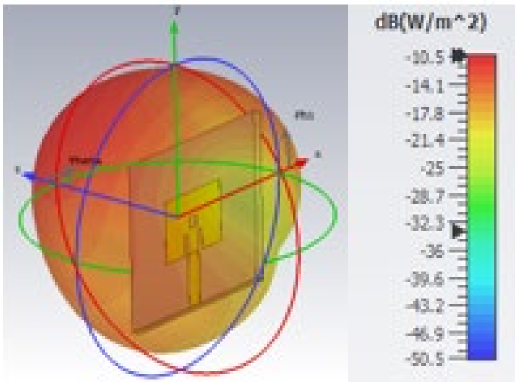


Fig. 6. Radiation pattern

4. Conclusion

In this paper, a rectangular UWB microstrip patch antenna with a compact size of 36mm × 32mm × 1.6mm is presented. This antenna can operate within 5.1256 GHz to 6.4391 GHz with UWB properties. It has a return loss of -17.927 dB, directivity of 5.3 dBi and a gain of 3.51 dBi at 5.8 GHz. The proposed rectangular microstrip patch antenna has been designed and simulated using CST Studio Suite 2019. The goal of this project is to design an UWB antenna which can operate within IoT, 5G and RFID applications. The simulated antenna can operate from 5.1256 GHz to 6.4391 GHz with a bandwidth of 1.3 GHz. This shows that the antenna satisfies the characteristics for the IoT, 5G and RFID applications.

References

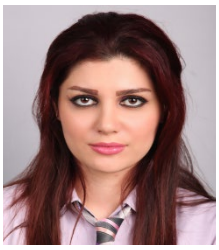
1. M. Mokayef and Y. Zahedi, "Prospects of 5G Communications," Journal of Independent Studies and Research – Computing, vol. 14, no. 1, 2016
2. O. Vermesan et al. "Internet of Things strategic research roadmap," in Internet of Things—Global Technological and Societal Trends From Smart Environments and Spaces to Green ICT, River, pp. 9–52, 2011.
3. Juels A. RFID security and privacy: A research survey. IEEE Journal on Selected Areas in Communications, 24(2): 381-394, 2006
4. Indrasen Singh et al, Int. J. Comp. Tech. Appl., Vol 2 (5), 1595-1599, 2011.
5. Federal Communication Commission, "First order and report: Revision of part 15 of the Commision's rules regarding UWB transmission systems," April 22, 2002.
6. H. S. J. Chen et al., "A Compact, Highly Efficient and Flexible Polymer Ultra-Wideband Antenna," in IEEE Antennas and Wireless Propagation Letters, vol. 14, pp. 1207-1210, 2015
7. R. A. A. Othman, D. a. Sagaran, M. Mokayef and W. I. I. R. b. W. M. Nasir, "Effective LiFi communication for IoT applications", *2018 IEEE 4th International Symposium in Robotics and Manufacturing Automation (ROMA)*, pp. 1-4, 2018.
8. Mastaneh Mokayef, MHD Amen Summakieh, "An Ultra-Wideband Antenna for IoT Connectivity" International Journal of Internet of Things and Web Services, 2017
9. . Verma, L. Mahajan, R. Kumar, H. S. Saini and N. Kumar, "A small microstrip patch antenna for future 5G applications," 2016 5th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, pp. 460-463, 2016

Authors Introduction

Mr. Gurney Nga

He has received his bachelor degree from the Faculty of Engineering and Built Environment (FETBE) UCSI University in 2021.

Asst. Prof. Dr. Mastaneh Mokayef



She has received her PhD from Wireless Communication Centre Faculty of Electrical Engineering in University Technology Malaysia (UTM) in 2014. She is a Chartered Engineer ((The Institution of Engineering and Technology (IET), United Kingdom) 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).

Mr. Shahid Manzoor

He is currently working as lecturer in UCSI University Kuala Lumpur Malaysia and doing his PhD in communication and electronic from UNITEN, Malaysia.He received his MS degree in Electrical & Electronic Engineering from Universiti Teknologi PETRONAS (UTP) in July-2008. He received his BS degree in Computer Engineering (With Honors) from University of Engineering and Technology Taxila (UET), Pakistan in Jan-2005.

Mr. Manickam Ramasamy

He is currently working as lecturer in UCSI University Kuala Lumpur Malaysia and doing his PhD. He is a Chartered Engineer (The Institution of Engineering and Technology (IET), United Kingdom). He is also Member, The Institution of Engineering and Technology (IET), United Kingdom , Graduate Member, Board of Engineers Malaysia (BEM), Malaysia, and Graduate Member Institute of Electrical and Electronic Engineering (IEEE).

Drone Performance Analysis Based on SNR Factor

Gershom Phiri
Department of Electrical and Electronic Engineering
Faculty of Engineering, Technology & Built Environment
UCSI University, Kuala Lumpur, 56000, Malaysia

Mastaneh Mokayef
Department of Electrical and Electronic Engineering
Faculty of Engineering, Technology & Built Environment
UCSI University, Kuala Lumpur, 56000, Malaysia

Sew Sun Tiang
Department of Electrical and Electronic Engineering
Faculty of Engineering, Technology & Built Environment
UCSI University, Kuala Lumpur, 56000, Malaysia

Wong Chin Hong
Department of Electrical and Electronic Engineering
Faculty of Engineering, Technology & Built Environment
UCSI University, Kuala Lumpur, 56000, Malaysia

E-mail: mastaneh@ucsiuniversity.edu.my , Gershomp8@gmail.com , shahid@ucsiuniversity.edu.my , ManickamRamasamy@ucsiuniversity.edu.my
www.ucsiuniversity.edu.my/

Abstract

In this paper, a depth analysis of a drone performance based on the signal to noise ratio (SNR), has been provided. The MATLAB software is used to simulate the effect of frequency, power and distance on the performance of drones. The SNR optimization has been proposed as well. The light has been shed on the drone performance in wildfire scenario.

Keywords: UAV, SNR, flying base station, cellular decomposition

1. Introduction

The human's craving to fly high in the sky developed as early as on time as its common sense. In any case, it took a long effort to make this fantasy real [1]. An enormous number of scientists had worked on this topic and it had requested such a significant number of courageous individuals, until at long last men could rise starting from the earliest stage. By that, people eagerness towards the aviation led success. From the start they conquered the air balloons, later by airships, lastly with planes. Meanwhile, the plan to use a machine that can fly without a person on board has consistently been in the scientists mind. This idea isn't unexpected in any way, because such systems advantages are obvious. We don't need to check with the death of the on-board personnel, if the airplane is destroyed for reasons unknown. Also, we can utilize them for such exhausting errands, similar to aerial surveillance. At last, their financial advantage position is unquestionable, because of the reality, that as a rule a UAV's first price is lower than the cost of a conventional aircraft. Like a great deal of objects that are used even in civilian life, UAVs are the result of the improvements done during military conflicts. Despite the fact that there were some unmanned balloons used as early as the middle eighteenth century to destroy the enemy. These kind of airplane showed up in the First World War. As nowadays every nation has their own UAV research program [2].

The primary unnamed aerial vehicle UAV guideline was proposed in 1944, directly after the World War II. The principal globally perceived flight guideline, the Chicago Convention, brought up that the activity of UAVs ought to be approved to guarantee the wellbeing of kept an eye on common airplane [9]. Since 2000, because of the fast improvement of UAV and its expanding ubiquity, UAV guidelines have developed both broadly and globally. In 2002, the United Kingdom and Australia originally distributed their UAV guidelines. In 2006, the International Civil Aviation Organization (ICAO) reported that it was important to give a globally recognized enactment for common tasks of UAVs. Since 2012, an expanding number of nations have built up their own UAV guidelines [10].

5G or so called the fifth generation of wireless communication is the new revolution and evolution in a wireless field that speed up the advent of IoT. 5G is an essential infrastructure that allows all communications to be linked through the IoT at anytime and anywhere. The main advantages of 5G are providing larger capacity, higher data-rate, lower latency and higher application reliability. It is 1000 times more capacity, 100 times faster and 100 times more connectivity compared to 4G [1].

IoT has accelerated the development of many devices around us and has brought up possibility of many applications. Its vision is to build a fully connected environment to ensure more convenient and better human living by improving the issues like energy management, climate change, automation and transportation, healthcare and treatment, logistic management and other related fields [2].

Radio frequency identification (RFID) is a key technology for making the IoT to automatically detect objects by using wireless communication. In recent years, RFID has been widely used in many places, such as logistic management and production. RFID is one of the main IoT technologies and one of the ten most important technologies in the 21st century, as an automatic detection and data capture technology [3]. RFID uses frequency from 30 Hz to 5.8 GHz which depends on the applications [4].

According to the Federal Communications Commission (FCC) in February 2002, the specifications for UWB systems to use in the band of 3.1 GHz to 10.6 GHz and a large bandwidth of more than 500 MHz [5]. UWB provides higher data rates with very low radiation power. Hence, this technology is getting more significant and popular in wireless communication systems [6].

In this paper, the study and design of an UWB microstrip patch antenna is presented. Microstrip patch antenna is chosen over other types due to its compact size, light weight, high bandwidth, multiband properties, low cost and high gain [7]. The proposed antenna needs to cover three of the applications as well as meeting the requirements of UWB. The rise of the IoT has accelerated the development of many devices around us and has opened the possibility of myriad applications. The IoT vision is to build a fully connected environment. This can be done by using smart objects and devices to produce data and transmit through the internet automatically for decision making purpose [7-9]. This is to ensure more convenient and better human living by improving the issues like energy management, climate change, automation and transportation, healthcare and treatment, logistic management, and other related fields. The IoT devices includes remote monitoring, tracking, collection of data, manufacturing and also for media applications. Nowadays, there are many IoT applications which are already identified globally i.e smart city, smart home, smart logistics, smart transportation, smart healthcare, and smart agriculture [8-10].

2. Background

The surveying of UAV based systems for traffic monitoring and management. Despite having a lot of research on the subject. Unmanned aerial vehicles (UAVs) are proven to be an applicable and less time-consuming alternative to real time traffic monitoring and management, providing the eye in the sky solution to the problem [1-3]. However, synthetic aperture radar (SAR) processing in general is not enough to detect and focus moving targets on the ground, therefore he described other techniques that are described and considered. Ideas for a long-track velocity were proposed but it was shown that clutter suppression and vehicle acceleration are major challenges for a current estimation of the motion status of road vehicles.

Deployment problem as reducing the number of UAVs and maximizing the load balance among them, which is subject to two main constraints, UAVs should form a robust backbone network and they should keep connected with the fixed based stations.to solve this optimizing problem with low complexity, he proposed a hybrid algorithm to solve them stepwise [4]. . First, a centralized greedy search algorithm is used to heuristically obtain the minimum number of UAVs and their suboptimal positions in a dis-continuous space. Then, a distributed motion algorithm is adopted which enables each UAV to autonomously control its motion towards the optimal position in a continuous space. The proposed algorithm is applicable to various scenarios where UAVs are deployed alone or with fixed base stations regardless of the user equipment distribution. Extensive simulations validate the proposed algorithms despite everything discussed in the paper the main purpose was to make sure they deploy minimum UVAs to evenly serve as many user equipment as possible while guaranteeing a robust backbone network [5].
5G and beyond 5G will be more important for the upcoming wireless networks wireless networks that can potentially facilitate wireless broadcast and support high rate transmissions compared to the communication with fixed infrastructure and he provided the exhaustive review of various 5G techniques based on UAV platforms such as networks layer, joint communication and computing. However in this survey they provided a brief understanding of UAV communication in 5G and beyond 5G [6-8].wireless networks and providing an overview on recent research activities an UAV communication combining the 5G techniques from the viewpoint physical layer, joint communication and computing.

UAVs can be connected to cellular networks as new types of user's equipment and how UAV offer unprecedented opportunity dynamically repositioning themselves to boost coverage, spectral efficiency and user quality experience as well as the obstacles and opportunities for helping cellular communication with UAV based flying relay and based station. However with all that discussed the outcome of the studies show some major interferences issues arising from height of UAVs but most of them can be addressed with the current 4G system even though 5G will be better to deal with UAV related challenges as a proven fact.

The research in [3] mainly focused on the point on the present comprehensive Survey of the literature on the location optimization of the unmanned aerial vehicles mounted base station in new gerent on wireless networks and the generic problem on the finding of the location of the base station and some solution to solve the challenges that are covered in the literature review and later gave the form of the mathematical formulation of flying base station location problem. However, the future research awareness will be to ensure the reliability of the services provided by flying base station.

3. Methodology

The about the operation of drone according to its coverage area, and methods to improve the drove lasting time in the air and a practical scenario in which UAV can be implemented and operate has been introduced. Moreover, the effect of SNR on the operation of the UAV and the optimization techniques have been discussed in this section.

3.1. *Coverage Path Planning*

Given an area of interest composed by the robot's free space and its limits, the coverage path planning (CPP) problem comprises of planning a path which covers the whole entire environment considering the vehicle's movement limitations and sensor's characteristics, while avoiding passing over obstacles [5]. In an aerial context, the workspace obstacles can represent no-flight zones (NFZ) that the UAV should not consider during the planning stage, e.g. areas close to air terminals or very tall buildings.
A few techniques can even deal with areas of interest containing NFZ which must be avoided from during coverage. These no-fly zones can represent areas where coverage is basically unnecessary or areas where the UAVs are not permitted to fly. Distinctive decompositions strategies are typically embraced to lessen the concavities of complex zones or to part the area into littler cells to facilitate the coverage task.

3.2. *Cellular Decomposition*

One of the major worries about the CPP problem is to ensure a total coverage of the scenario. This is generally accomplished by applying cellular decomposition in the area of interest, splitting the target free space into cells in order to simplify the coverage In literature there are different cellular decomposition methods and the most well-known used in CPP problem including UAVs is approximate cellular decomposition. The approximate cellular decomposition method approximates the area into a lot of regular cells. These regular cells for the most part expect a square structure, yet they can also be represented either in a triangle or hexagonal structure Grid based methods can be applied over approximate areas to generate coverage paths. The size of the cells for the most part fits the robots measurements while considering coverage using land robots. However, in aerial coverage the UAVs fly at a specific altitude from the ground carrying a camera as a sensor to perform the task. For this situation, the footprint of the camera in the UAV while it's flying is shown in Fig. 1 bellow.

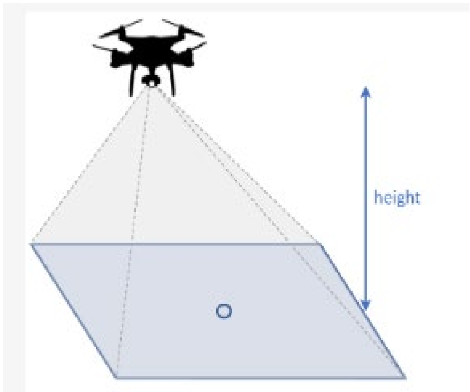


Fig. 1. Square area presentation overview of a flying drone.

3.3. *Drone Performance Matrics*

3.1.1. *Flying Altitude*

Every nation has their own laws of how high one can fly their drone. While it might have the option to fly a thousands of feet in the air [9]. There's a set law in Malaysia as indicated by Malaysia's national aviation authority, the civil aviation authority of Malaysia (CAAM) that permits drone owners to fly up to 400 feet high and you can't fly within 50 meters of any individual. In the European Union, you can fly an automaton up to 500 feet from the beginning. This is to prevent any interference with airplanes.

3.4. *Mission Completion Time*

For a drone to last long on its flight to complete its mission you need a long lasting battery that can at least last up to 25 minutes. Aside from the battery the weight can affect your drone flight time. Drone add-on also plat a major a major role of any extensions attached, one of the biggest interest of drone ad-on is the camera therefor the heavier the camera is the less time your drone flight will be and the weather matters too, you have in a calm weather and avoid strong winds, rains and any other unsuitable weather [8]. Strong wind can majorly affect your drone flight time because the battery works harder to power through it. Drone can be applied and be useful in different scenarios and the following is an example of a practical example of WILDFIRES in which a drone is being used.

3.5. *Wildfire*

A wildfire, Wild land fire or rural fire is an uncontrolled fire in a region of burnable vegetation happening in rural areas. Depending upon the sort of vegetation present, a rapidly spreading wildfire can likewise be grouped all as a wildfire, bushfire (in Australia), desert fire, forest fire, grass fire, hill fire, vegetation fire or veld fire. Numerous organization think about wildfire as unplanned and undesirable fire, while wild fire is a more extensive term that include prescribed fire just as Wild land fire use (WFU) these are additionally called response fires. Fossil charcoal demonstrates that wildfires started not long after the presence of terrestrial plants 420 million years prior. Wildfire's event since the commencement of terrestrial life welcomes conjecture that fire probably had pronounced developmental consequences on most ecosystem widely varied vegetation. Earth is a characteristically flammable planet attributable to its cover of carbon-rich vegetation, seasonally dry climate, atmospheric oxygen, and broad lightning and volcanic ignitions [8-10]. In the case of wildfire to monitor the situation on the ongoing fire a drone is used.in this disastrous situation we need a quad copter drone with savor motors, battery, pay load of 6kg and a thermal camera sensor. The UAV is fitted with an integrated thermal camera sensor for monitoring a controlled burn and use during post fire cleanup for hot spots detection from above. The following parameters must be considered in order for UAV to work to its maximum potential: frame, transmitter, receiver, power, propellers, motors, ESCs (electrical speed controller) [9].

4. Signal to Noise Ratio

Signal-to-noise ratio (abbreviated SNR or S/N) is a measure used in science and engineering that compares the level of a desired signal to the level of background noise. SNR is defined as the ratio of signal power to the noise power, often expressed in decibels. A ratio higher than 1:1 (greater than 0 dB) indicates more signal than noise.

The transmission of information from one point to another is called data link. When this transition is accomplished using wireless means the data link is now called radio link or radio modem

In unmanned aviation, the data link permits the transmission and gathering of information between the autopilot in the aircraft and the Ground Control Station (GCS).in order to transmit information we have two links called uplink and downlink. SNR is considered when receiving information using data links, SNR compares the level of a desired signal to the level of background noise. The higher the SNR the better signal quality.it is given by the formula below [10]. When measuring signal-to-noise-ratio there are two basic elements to the measurement, one is noise level and the other is the signal. The signal often includes noise due to the way measurements are made in which the signal level is assumed to be much larger than the noise hence it is considered signal plus noise.

5. Results and discussion

In this paper, the main objective is to analyse the drone performance based on signal to noise ratio. Therefore a parametric study on the effects of changing distance, frequency and transmit power of the UAV are analysed and these are the only parameters that we are focused on in this paper.

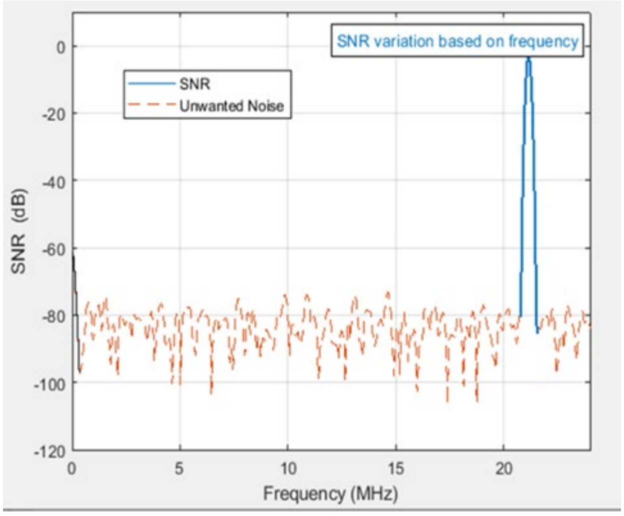


Fig. 2. SNR Variation in relation to frequency

Fig. 2 shows the SNR variation in subject to frequency, the graph also shows the noise within space and signal strength closer to the transmitter. The SNR to a given frequency band is sensitive as different frequency band will vary causing the SNR to be affected. Typically, the input voltage for SNR threshold level of 10dB is stated and the noise level threshold of -90dBm is being used. Based on the SNR the transmitter and the receiver will negotiate a data rate with which to communicate. In a free world propagation where there is nothing blocking SNR the frequency plays a major role in communication system. While the transmit power and the distance is maintained the frequency is affected where we have buildings and trees, the signal wave is affected hence causing the SNR to be affected so if we are going to a lower frequency the SNR is reduced and If we going to a higher frequency the SNR increases hence making the connection strong. Therefore, every move away we lose the frequency and the higher the SNR the better because higher data rate can be used to increase the communication.

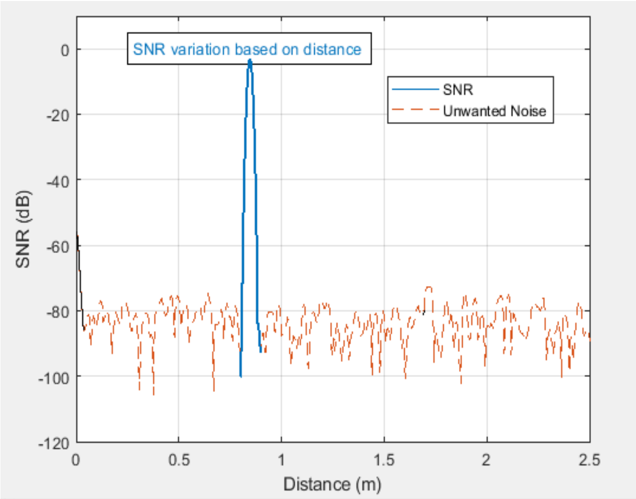


Fig. 3. SNR variation subjected to distance

Fig.3 shows the SNR as function to distance, the graph also shows noise within space and signal strength over a distance as a curve that has the best signal strength closer to the transmitter, the SNR threshold level is 10 dB however for a good performance SNR must be greater than 20 dB, to analyze the performance the frequency threshold of 2.4GHz and power threshold of 1W are maintained and the distance is changed, the SNR is affected as we move away from the signal. When we are at a distance of 1m and frequency of 10MHz with a transmit power of 1W, we have a stronger signal due to less noise that come around the receiver but as we move further away we get lower SNR caused by noise that can come from the remote control, phone, Bluetooth, trees and building distracting the signal more, hence the SNR is affected by distance. Therefore, SNR at a larger distance will require a higher frequency and power.

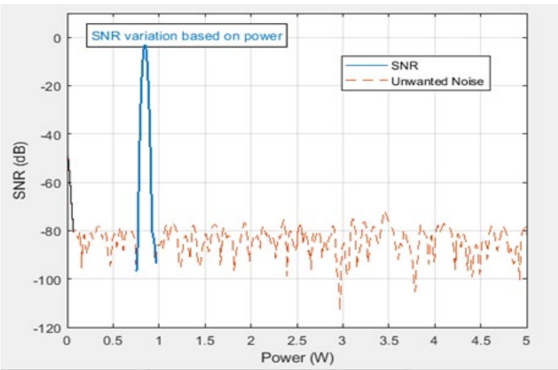


Fig. 4. Effect of power on SNR variation

Figure 4 shows the SNR variation with power, the noise and signal level strength. The noise reduce the distance that the transmitted signal can travel the main concern here is to make sure the transmitted power is high enough to overcome thermal noise that is contributed by the communication system. The transmitted power decreases over distance in watts/m2 and density hence the power received is already lower than that which was transmitted. Transmit power must increase to ensure high SNR, therefore, to overcome covering a larger distance of SNR we increase our power or amplify the noise.

6. Conclusion

The importance of a UAV operation in recent years have been shown in this paper. It can be very useful for temporal communication in disastrous scenario. However issues with UAV application and factors affecting them must be addressed and where they can fly according to the rule and regulation for UAV according to the region and country. On the other hand wireless network communication is one of the key points that make unnamed air vehicles more useful other than other methods of commutations that are currently used. This paper describes the effects of distance, frequency and power on signal to noise ratio. As depicted in results, the larger the distance the more they affect the signal and the more frequency results to the better SNR. The result also proved the same effect from increment in power however, more power is required to overcome thermal noise in transmission.

References

1. N. Gaiman, "signal to noise," bloomsbury Publishing PLC, vol. 4, pp. 45-50, 2019.
2. O. Vermesan et al. “Internet of Things strategic research roadmap,” in Internet of Things—Global Technological and Societal Trends From Smart Environments and Spaces to Green ICT, River, pp. 9–52, 2011.
3. A. Crimmins, J. Balbus, J. Gamble, C. Beard, L. Jantarasami and D. M. Mills, "Impacts of Climate Change on Human Health in the United States: A Scientific Assessment.," Eds. U.S. Global Change Research Program, Washington, DC, p. 312, 2019
4. D. M, . M. G. P. López, P. Wang and Z. Lin, "The area spectralEfficiency Monotonically Grow as small cells go dense," ieee, pp. 1-7, 2018.
5. N. Silver, The signal and the noise, New york: penguin putnam, 2018.
6. Y. Zahedi, R. Ngah, U. Oknonkwo, S. Nunoo and M. Mokayef, "UWB channel measurement and development

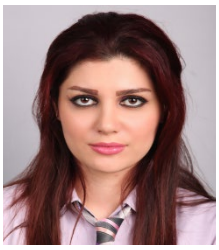
of scatterer identification algorithm," *2014 2nd International Conference on Electronic Design (ICED)*, 2014, pp. 511-514, doi: 10.1109/ICED.2014.7015860.
7. G. S. Ganesan and M. Mokayef, "Multi-Purpose Medical Drone for the Use in Pandemic Situation," 2021 IEEE Microwave Theory and Techniques in Wireless Communications (MTTW), 2021, pp. 188-192, doi: 10.1109/MTTW53539.2021.9607300.
8. A. A. Elijah, M. Mokayef and E. E. Hussin, "The Effect of Different Position and Size of U-Shape Slot on a Microstrip Patch Antenna for IoT Application," 2018 IEEE 4th International Symposium in Robotics and Manufacturing Automation (ROMA), 2018, pp. 1-6
9. Y. Zahedi, R. Ngah and M. Mokayef, "Path Loss Modeling for Outdoor UWB Channels," 2018 IEEE 4th International Symposium in Robotics and Manufacturing Automation (ROMA), 2018, pp. 1-5
10. V. M. l. D. d. Olga Boric-Lubecke, Sources of Noise and Signal to Noise Ratio, toronto: cananda books, 2018.

Authors Introduction

Mr. Garshom Phiri

He has received his bachelor degree from the Faculty of Engineering and Built Environment (FETBE) UCSI University in 2021.

Asst. Prof. Dr. Mastaneh Mokayef



She has received her PhD from Wireless Communication Centre Faculty of Electrical Engineering in University Technology Malaysia (UTM) in 2014. She is a Chartered Engineer ((The Institution of Engineering and Technology (IET), United Kingdom) 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).

Asst. Prof. Dr. Sew Sun Tiang

She received the Bachelor of Engineering (Hons.) Electronics majoring in Telecommunications from Multimedia University (MMU) in 2008 and Ph.D degree in Electrical & Electronic Engineering from Universiti Sains Malaysia (USM) in 2014. She is currently working as an Assistant Professor in the Faculty of Engineering, Technology & Built Environment, UCSI University. Her research interests include antenna design, wireless communication, optimization, and machine 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.