ICAROB 2020

PROCEEDINGS OF THE 2020 INTERNATIONAL CONFERENCE ON ARTIFICIAL LIFE AND ROBOTICS

January 13-16, 2020
B-Con Plaza, Beppu, Oita, JAPAN
25th AROB International Meeting Series

Editor-in-Chief
Masanori Sugisaka
Editors: Yingmin Jia, Takao Ito, Ju-Jang Lee

ISBN 978-4-9908350-5-7

Proceedings of The 2020 International Conference on

ARTIFICIAL LIFE AND ROBOTICS

(ICAROB2020)

January 13-16, 2020
B-Con Plaza, Beppu, Oita, JAPAN
25th AROB International Meeting Series

Editor-in-Chief
Masanori Sugisaka
Editors: Yingmin Jia, Takao Ito, Ju-Jang Lee
ISBN 978-4-9908350-5-7

Contents

SPONSERED

ALife Robotics Corporation Ltd.

ORGANIZED BY

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

SUPPORTED BY

IEEE Fukuoka Section (Japan)

CO-ORGANIZED BY

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

ADVISORY COMMITTEE CHAIRS

Moshe Kam (New Jersy Institute of Technology, Former IEEE President, USA)
Kensuke Harada (Osaka University, Japan)

ADVISORY COMMITTEE

Adam Grzech (Wroclaw University of Technology, Poland)
Bruce Eisenstein (Drexel University, Former IEEE President, USA)
Hidenori Kimura (RIKEN & Wasada University, Japan)
Jeffrey Johnson (The Open University, UK)
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)
Jangmyung Lee (Pusan National University, Korea)
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 (AIST, 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)
Eiji Hayashi (Kyushu Institute of Technology, Japan)
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, Nippon Bunri University (NBU), 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 this journal, International Journal of Robotics, Networking and Artificial Life (JRNAL). From now on, ALife Robotics Corporation Ltd. is in charge of management of both the conference and the journal. 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 2020 International Conference on Artificial Life and Robotics (ICAROB2020) 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 an international journal ROBOTICS, NETWORKING AND ARTIFICIAL LIFE 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)

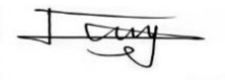
Masanori Sugisaka

Masanori Sugisaka

General Chair of ICAROB

It is my great honor to invite you all to The 2020 International Conference on Artificial Life and Robotics (ICAROB 2020).
This Conference is changed as the old symposium from the first (1996) to the Eighteenth (2013) annually which were organized by Oita University, Nippon Bunri University(NBU), 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 2020 International Conference on Artificial Life and Robotics (ICAROB2020) (25th 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 hope that fruitful discussions and exchange of ideas between researchers during Conference (ICAROB2020) 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

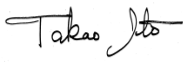
It is my great pleasure to invite you to The 2020 International Conference on Artificial Life and Robotics (ICAROB 2020), in B-Con Plaza, Beppu, Oita, Japan from January 13 to 16, 2020.

ICAROB develops from the AROB that was created in 1996 by Prof. Masanori Sugisaka and will celebrate her 25th birthday in 2020. 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 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 meeting you at ICAROB in Beppu and wishing you enjoy your stay in Japan.

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



Takao Ito

Co General Chair of ICAROB

It is my great honor to invite you all to the 2020 International Conference on Artificial Life and Robotics (ICAROB 2020), to the wonderful city of Beppu, Oita city, Oita Prefecture, Japan.

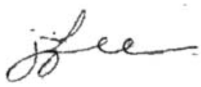
The ICAROB has its long history. First launched in 1996 as ISAROB, this former organization of ICAROB, was developed under the strong leadership and yeoman efforts of the President—the internationally famous Professor Masanori Sugisaka, who is widely acknowledged as the father of AROB. Our symposium has brought together many research scholar, 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. ICAROB has provided a foundation for unifying the exchange of scientific information on the study of man-made systems that exhibit the behavioral characteristics of natural living systems, including software, hardware and wetware. Our conference shapes the development of artificial life, extending our empirical research beyond the territory circumscribed by life-as-we-know-it and into the domain of life-as-it-could–be. It will provide us a good place to present our new research results, innovative ideas, and valuable information about artificial intelligence, complex systems theories, robotics, and management of technology.

The conference site is the B-con Plaza, one of the most famous international convention centers in Kyushu island, Japan. You can find many fantastic scenic spots and splendid historical places in Beppu, Oita city. Do enjoy your stay and take your time to visit Beppu, Oita city.

I eagerly look forward to personally meeting you in Beppu, Oita city, during the ICAROB 2020 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 7th ICAROB 2020 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 7th ICAROB.

I'm looking forward to meeting you at the 7th ICAROB in Beppu, Oita Prefecture and wishing you all the best.

GENERAL SESSION TOPICS

GS1 Neural Networks (6)	GS2 Control Techniques (5)
GS3 Vision & Image Processing (4)	GS4 Robotics 1 (4)
GS5 Robotics 2 (4)	GS6 Application Techniques (8)
GS7 Poster (4)	

ORGANIZED SESSION TOPICS

OS1 Machine vision and Mobile robot (4)	OS2 Advanced Control (4)
OS3 Intelligent Control (4)	OS4 Intelligent Life and Systems (7)
OS5 Artificial Life and Robotics (7)	OS6 Theory and Implementation of Neuromimetic Systems (4)
OS7 Robotic Manipulation (4)	OS8 Advanced Research of Engineering and Management (4)
OS9 Recognition and Control (19)	OS10 AI Applications (8)
OS11 System and Control (16)	OS12 Advances in Theory and Education on Control (4)
OS13 Natural Computing (5)	OS14 Software Development Support Method (5)
OS15 Recognition and System (11)	OS16 Media Information Processing and Artificial Intelligence (4)
OS17 Machine Learning Technologies for Human Understanding (2)	OS18 Service Robotics (7)
OS19 Advanced Information Processing Applications (5)	OS20 Artificial Intelligence for Embedded Systems and Robotics (5)
OS21 Mathematical Informatics (4)	OS22 Robot Competitions for Social Contribution (5)
OS23 Advances in Field Robotics and Their Applications (5)	OS24 Robot Intelligence and Factory Automation (6)

1/13(Mon.) 17:30-19:30	Welcome Party (Beppu Kamenoi Hotel)
1/14(Tue.) – 1/16(Thu.)	ICAROB Secretariat
1/16(Thu.)	Farewell Party (Conference Site: 3F, Meeting Room 32)

TIME TABLE (1/14)

1/14(Tue.)	Conference Room	Meeting Room 31	Meeting Room 32	Meeting Room 33	Meeting Room 1	Meeting Room 4
8:40-	Registration (3F)					
9:00-10:00		OS5 Artificial Life and Robotics (7) Chair: Chung-Wen Hung Will be end at 10:45	OS22 Robot Competitions for Social Contribution(5) Chair: Yasunori Takemura Will be end at 10:15	OS6 Theory and Implementation of Neuromimetic Systems (4) Chair: Takashi Kohno	OS20 Artificial Intelligence for Embedded Systems and Robotics (5) Chair: Hakaru Tamukoh Will be end at 10:15	OS12 Advances in Theory and Education on Control (4) Chair: Takao Sato
10:00-10:20	Coffee break					
10:20-10:50	Opening Ceremony (Conference Room)					
11:00-12:00	Chair: Evgeni Magid Plenary Speech PS-1(Conference room) Takashi Kohno					
12:00-13:00	Lunch					
13:00-14:00	Chair: Ju-Jang Lee Plenary Speech PS-2(Conference Room) Eiji Hayashi					
14:00-14:20	Coffee break					
14:20-15:20	Chair: Jangmyung Lee Invited Speech IS-2, IS-3(Conference Room) Henrik Hautop Lund, Luigi Pagliarini					
15:20-15:40	Coffee break					
15:40-17:55		GS6 Application Techniques (8) Chair: Hazry Desa	OS4 Intelligent Life and Systems (7) Chair: Kuo-Hsien Hsia OS18-1 Service Robotics (2) Chair: Evgeni Magid	OS 23 Advances in Field Robotics and Their Applications (5) Keisuke Watanabe	OS13 Natural Computing (5) Chair: Marion Oswald	OS24 Robot Intelligence and Factory Automation (6) Chair: Eiji Hayashi

Meeting Room 31: Committee waiting room and Rest room

TIME TEBLE (1/15)

1/15(Wed.)	Conference Room	Meeting Room 31	Meeting Room 32	Meeting Room 33	Meeting Room 1	Meeting Room 4
8:40-	Registration					
9:00-10:00			OS7 Robotic Manipulation (4) Chair: Kensuke Harada	GS7 Poster (4) Chair: Jiwu Wang	OS18-2 Service Robotics (5) Chair: Evgeni Magid Will be end at 10:15	GS4 Robotics 1 (4) Chair: Noritaka Sato
10:00-10:15	Coffee break					
10:15-11:15			OS3 Intelligent Control (4) Chair: Yingmin Jia	OS8 Advanced Research of Engineering and Management (4) Chair: Takao Ito		GS5 Robotics 2 (4) Chair: Hidehiko Yamamoto
11:15-11:30	Coffee break					
11:30-12:30	Chair: Mohamed Rizon Plenary Speech PS-4 (Conference Room) Naoki Suganuma					
12:30-13:30	Lunch					
13:30-14:30	Chair: Marion Oswald Plenary Speech PS-3 (Conference Room) Jeffrey Johnson					
14:30-14:45	Coffee break					
14:45-16:15			OS10-1 AI Applications (6) Chair: Mohamed Rizon	OS16 Media Information Processing and Artificial Intelligence (4) Chair: Yasunari Yoshitomi		GS1 Neural Networks (6) Chair: Akira nakamura
16:15-16:30	Coffee break					
16:30-17:45			OS10-2 AI Applications (2) Chair: Mohamed Rizon	OS19 Advanced Information Processing Applications (5) Chair: Toru Hiraoka		GS2 Control Techniques (5) Chair: Chian C. Ho
18:30-20:30	Banquet: Beppu Kamenoi Hotel					

Meeting Room 31: Committee waiting room and Rest room (available from 9:00 a.m.)

TIME TABLE (1/16)

1/16(Thu.)	Meeting Room 31	Meeting Room 32	Meeting Room 33	Meeting Room 1	Meeting Room 4
8:40-	Registration				
9:00-10:15		OS2 Advanced Control (4) Chair: Yingmin Jia	OS14 Software Development Support Method (5) Chair: Tetsuro Katayama		
10:15-10:30	Coffee break				
10:30-11:00	Chair: Yingmin Jia Invited speech IS-1(Meeting Room 31) Jangmyung Lee				
11:00-12:00	Lunch				
12:00-13:15	OS11 System and Control (16) Chair: Yizhun Peng	OS1 Machine vision and Mobile robot (4) Chair: Wei Liu	OS17 Machine Learning Technologies for Human Understanding (2) Chair: Masao Kubo		
13:15-13:30	Coffee Break				
13:30-14:45	OS9 Recognition and Control (19) Chair: Fengzhi Dai	OS21 Mathematical Informatics (4) Chair: Makoto Sakamoto	GS3 Vision & Image Processing (4) Chair: Joo Kooi Tan		
Farewell Party (15:00-15:30) Meeting Room 32					

Meeting Room 31: Committee waiting room and Rest room(available from 9:00 a.m.)

The 2020 International Conference on ARTIFICIAL LIFE AND ROBOTICS (ICAROB2020)

January 13 (Monday)

17:30-19:30 Welcome Party (Beppu Kamenoi Hotel)

January 14 (Tuesday)

10:20-10:50

Opening Ceremony (Conference Room)

Chair: Marion Oswald (Technische Universität Wien, Austria)

Welcome Addresses

1. General Chairman of ICAROB	Masanori Sugisaka (ALife Robotics Co., Ltd., Japan)
2. Co-General Chairman of ICAROB	Yingmin Jia (Beihang University, China)
3. Co-General Chairman of ICAROB	TaKao Ito (Hiroshima University, Japan)
4. Co-General Chairman of ICAROB	Ju-Jang Lee (KAIST, Korea)

January 15 (Wednesday)

Banquet: Beppu Kamenoi Hotel

18:30-20:30

Chair: Takao Ito (Hiroshima University, Japan)

Welcome Addresses

Prof. Yingmin Jia (Beihang University, China)
Prof. Ju-Jang Lee (KAIST, Korea)
Prof. Henrik Hautop Lund (Technical University of Denmark, Denmark)
Prof. Jangmyung Lee (Pusan National University, Korea)
Prof. Mohd Rizon (University of UCSI, Malaysia)

.

TECHNICAL PAPER INDEX

January 14 (Tuesday)

8:40-Registration

Conference Room
10:20-10:50 Opening Ceremony
Chair: Marion Oswald (Technische Universität Wien, Austria)

11:00-12:00
Plenary Speech PS-1
Chair: Evgeni Magid (Kazan Federal University, Russia)

PS-1 *Towards neuromimetic computing*
Takashi Kohno (University of Tokyo, Japan)

13:00-14:00
Plenary Speech PS-2
Chair: Ju-Jang Lee (KAIST, Korea)

PS-2 *Innovative Robot - Robot/AI for Factory Automation-*
Eiji Hayashi (Kyushu Institute of Technology, Japan)

14:20-14:50
Invited session IS-2
Chair: Jangmyung Lee (Pusan National University, Korea)
IS-2-1 *Body and Brain Training with Big Data and AI*
IS-2-2 *Body and Brain Training with Big Data and AI 2 – A Pilot Test of Falls Prevention*
IS-2-3 *Playware Ball – Initial Development Impressions of an Intelligent Ball*
Henrik Hautop Lund (Technical University of Denmark, Denmark)

14:50-15:20
Invited session IS-3
Chair: Jangmyung Lee (Pusan National University, Korea)

IS-3-1 *PixelBeing – An Eco-sustainable Approach to Robotics and AI*
IS-3-2 *Social Play with Modular Playware*
Luigi Pagliarini[1,2,3], Henrik Hautop Lund[1]
([2]Academy of Fine Arts of Macerata, [3] ISIA Design, Italy), ([1]Technical University of Denmark, Denmark)

.

Meeting Room 31
9:00-10:45 OS5 Artificial Life and Robotics (7)
Chair: Chung-Wen Hung (National Yunlin University of Science and Technology, Taiwan)
Co-Chair: Ching-Chun Chuang (National Yunlin University of Science and Technology, Taiwan)

OS5-1 *An FSK based Industrial Analog Signal Transmission*
Po-Yun Shih, Chung-Wen Hung, Chau-Chung Song
(National Yunlin University of Science and Technology, Taiwan)

OS5-2 *X-Y Platform Synchronous Control with CANopen*
Yu-Ming Guan, Chung-Wen Hung, Shih-Ting Yu, Yu-Kai Chen
(National Yunlin University of Science and Technology, Taiwan)

OS5-3 *A PFC Converter with Voltage Double Characteristic for Universal Input Voltage Applications*
Ching-Chun Chuang, Hung-Chi Lee, Chih-Chiang Hua, Chih-Wei Chuang, Chuan-Ming Niu
(National Yunlin University of Science and Technology, Taiwan)

OS5-4 *Design of a Low-pulse High-current LLC Resonant Converter for EDM Applications*
Yu-Kai Chen, Min-Feng Lee, Yung-Chun Wu, and Jui-Yang Chiu
(National Formosa University, Taiwan)

OS5-5 *Networking Integration and Monitoring System with CANopen Controller for Intelligent Production Line of Tool Machine*
Chau-Chung Song, Yu-Wei Ho, Chen-Pang Chen, Yu-Kai Chen
(National Formosa University, Taiwan)

OS5-6 *Illumination Manipulation and Specular Reflection Analysis of Still Image with Single Object*
Hsuan T. Chang, Chi-Jie Chen
(National Yunlin University of Science and Technology, Taiwan)

OS5-7 *Robotics Education for the 2019 MakeX Robotics Competition*
Jia-Ming Hsiao (Far East University, Taiwan)

15:40-17:40 GS6 Application Techniques (9)
Chair: Hazry Desa (Universiti Malaysia Perlis, Malaysia)

GS6-1 *A Promoting Method of Role Differentiation Using the Learning Rate Getting Periodically Negative Value in Multi-agent Reinforcement Learning*
Masato Nagayoshi, Simon Elderton (Niigata College of Nursing, Japan),
Hisashi Tamaki (Kobe Univ. Japan)

GS6-2 *Verification of a Combination of Gestures Accurately Recognized by Myo Using Learning Curves*
Kengo Kitakura, Hideyuki Tanaka (Hiroshima University, Japan)

.

GS6-3 *Augmentative and Alternative Communication Device Based on Head Movement to Aid Paralyzed Victims with Speech Disabilities*
Vihanga Ashinsana Wijayasekara, Torin Wirasingha
(Informatics Institute of Technology, Sri Lanka)

GS6-4 *The research about editing system of performance information for player piano. –Make inferences about whole musical composition by using DP matching system-*
Mami Ezaki, Eiji Hayashi (Kyushu Institute of Technology, Japan)

GS6-5 *Evaluating Public Perception using Fuzzy Logic: A case study of Praeksa Mai dumpsite in Samut Prakan, Thailand (withdraw)*
Sun Olapiriyakul, Khemika Kongpetch
(Sirindhorn International Institute of Technology, Thammasat University, Thailand*)*

GS6-6 *A Development of a Model CubeSat with an Amateur Radio Transceiver for Education on Satellite Communication*
Masahiro Tokumitsu, Kentarou Konishi (National Institute of Technology, Yonago College, Japan),
Taku Takada (National Institute of Technology, Kochi College, Japan)
Fumio Asai (Member of the Radio Amateur Satellite Corporation)
Makoto Wakabayashi (National Institute of Technology, Niihama College, Japan)

GS6-7 *Performance Evaluations on Data Estimation Technique with Statistical Properties of Telemetry Data for Corrupted Data in Amateur Satellite Communication*
Yusuke Teranishi, Masahiro Tokumitsu (National Institute of Technology, Yonago College, Japan)
Taku Takada (National Institute of Technology, Kochi College, Japan)
Fumio Asai (Member of the Radio Amateur Satellite Corporation)
Makoto Wakabayashi (National Institute of Technology, Niihama College, Japan)

GS6-8 *Heritage Building Design Properties: Development of As Built Drawing by UAV Application via 3D Laser Scanner*
Hazry. D, Azizan. M. A, Noriman. N.Z, Romeli. N (Universiti Malaysia Perlis, Malaysia)

GS6-9 *A Virtual System for Measurements and Analysis of the Respiratory Sounds for Diagnosis of Respiratory System*
Ali S. AlMejrad (King Saud University, Kingdom of Saudi Arabia

Meeting Room 32
9:00-10:15 OS22 Robot Competitions for Social Contribution (5)
Chair: Yasunori Takemura (Nishinippon Institute of Technology, Japan)
Co-Chair: Kazuo Ishii (Kyushu Institute of Technology, Japan)

OS22-1 *Real-Time Self Localization for Autonomous Robot of RoboCup MSL*
Kaori Watanabe, Yuehang Ma, Tetsuya Yoshida, Hidekazu Suzuki

.

(Tokyo Polytechnic University, Japan)

OS22-2 *Behavior Selection System for Soccer Robot Using Neural Network*
Moeko Tominaga[1], Yasunori Takemura[2], Kazuo Ishii[1]
([1]Kyushu Institute of Technology, [2]Nishinippon Institute of Technology, Japan)

OS22-3 *Development of the Auto Measurement System for Cedars in a Forest Using a Drone*
Keiji Kamei, Masahiro Kaneoka, Ken Yanai, Masaya Umemoto, Hiroki Yamaguchi, Kazuki Osawa (Nishinippon Institute of Technology, Japan)

OS22-4 *Report on the 5th Tomato-harvesting Robot Competition*
Yasunori Takemura[1], Takayuki Matsuo[2], Takashi Sonoda[1], Kazuo Ishii[3]
([1]Nishinippon Institute of Technology, [2]National Institute of Technology, Kitakyushu College, [3]Kyushu Institute of Technology, Japan)

OS22-5 *Reports on the 7th Underwater Robot Festival in Kitakyushu*
Yuya Nishida[1], Takashi Sonoda[2], Takayuki Matsuo[3], Shinsuke Yasukawa[1], Masanori Sato[4], Yasunori Takemura[2], Kazuo Ishii[1]
([1]Kyushu Institute of Technology, [2]Nishinippon Institute of Technology, [3]National Institute of Technology, Kitakyushu College, [4]Nagasaki Institute of Applied Science, Japan)

15:40-17:25 OS4 Intelligent Life and Systems (7)
Chair: Kuo-Hsien Hsia (National Yunlin University of Science and Technology, Taiwan)
Co-Chair: I-Hsien Liu (National Cheng-Kung University, Taiwan)

OS4-1 *Development of the IoT Module with AI Function Using STM32 Chip*
Jr-Hung Guo, Evgeni Magid*, Kuo-Hsien Hsia, Kuo-Lan Su
(National Yunlin University of Science and Technology, Taiwan)
(*Kazan Federal University, Russia)

OS4-2 *Markerless Indoor/Outdoor Augmented Reality Navigation Device Based on ORB-Visual-Odometry Positioning estimation and ORB-Visual-Mapping Image Registration*
Chian C. Ho, Guan-Jie Wang (National Yunlin University of Science and Technology, Taiwan)

OS4-3 *Apply Adaptive Control Approach for Mobile Robot Path Following*
Chun-Chi Lai, Chia-Jen Lin, Kuo-Hsien Hsia, Kuo-Lan Su
(National Yunlin University of Science and Technology, Taiwan)

OS4-4 *Malware Classification Using Deep Learning*
Cheng-Hsiang Lo, Ta-Che Liu, I-Hsien Liu, Jung-Shian Li, Chuan-Gang Liu, Chu-Fen Li
(National Cheng Kung University, Taiwan)

OS4-5 *Robust Control of Nonholonomic Wheeled Mobile Robot with Hybrid Controller Approach*

.

Ho-Nien Shou (National Yunlin University of Science and Technology, Taiwan)

OS4-6 *Design and Implementation of Microsatellite Detumbling Control with PWPF: Verification Using PIL*
Ho-Nien Shou (Air Force Institute of Technology, Taiwan)

OS4-7 *A Study of Applying Computer-assisted Language Learning to English Course for Junior College Students in Taiwan*
Shu-Hua Huang, I-Hsien Liu (Air Force Institute of Technology, Taiwan)

17:25-17:55 OS18-1 Service Robotics (2)
Chair: Evgeni Magid (Kazan Federal University, Russia)
Co-Chair: Kuo-Lan Su (National Yunlin University of Science and Technology, Taiwan)
Co-Chair: Kuo-Hsien Hsia (National Yunlin University of Science and Technology, Taiwan)

OS18-6 *Pick and Place of Large Object Based on 3D vision*
Hsien-Huang P. Wu, Jia-Kun Xie (National Yunlin University of Science & Technology, Taiwan)

OS18-7 *Application of the MyRIO Based Mobile Robot Using Vision System*
Bo-Jun Yang[1], Kuo-Hsien Hsia[1], Kuo-Lan Su[1], Evgeni Magid[2]
([1]National Yunlin University of Science & Technology, Taiwan)
([2]Kazan Federal University, Russia)

Meeting Room 33
9:00-10:00 OS6 Theory and Implementation of Neuromimetic Systems (4)
Chair: Takashi Kohno (University of Tokyo, Japan)
Co-Chair: Timothée Levi (University of Tokyo, Japan)

OS6-1 *Real –time pattern recognition implementation on FPGA in multi-SNNs*
Xia Yang, Kazuyuki Aihara, Timothée Levi, Takashi Kohno (University of Tokyo, Japan)

OS6-2 *Experimental results of a biomimetic silicon synaptic circuit*
Ashish Gautam, Takashi Kohno (University of Tokyo, Japan)

OS6-3 *Towards Modeling Cholinergic Modulation for Neuromorphic Computing*
Naruaki Takano, Takashi Kohno (University of Tokyo, Japan)

OS6-4 *Bioelectrical Signal Analysis of Mouse Cardiomyocyte Culture recorded on Thin-Film-Transistor Sensor Arrays*
Anne-Claire Eiler, Junichi Sugita, Satoshi Ihida, Hiroshi Toshiyoshi, Katsuhito Fujiu, Thimothée Lévi, Agnes Tixier-Mita (The University of Tokyo, Japan)

.

15:40-17:10 OS23 Advances in Field Robotics and Their Applications (6)
Chair: Keisuke Watanabe (Tokai University)
Co-Chai: Kazuo Ishii (Kyushu Institute of Technology)

OS23-1 *Sea Trials for Benthos Sampling Using Autonomous Underwater Vehicle*
Yuya Nishida[1], Shinsuke Yasukawa[1], Takashi Sonoda[2], Keisuke Watanabe[3], Kazuo Ishii[1]
([1]Kyushu Institute of Technology, [2]Nishinippon Institute of Technology, [3]Tokai University, Japan)

OS23-2 *Field Experiments of Underwater Image Transmission for AUV*
Shinsuke Yasuakwa[1], Yuya Nishida[1], Jonghyun Ahn[2], Takashi Sonoda[3], Kentaro Yanagise[1], Keisuke Watanabe[4], Kazuo Ishii[1]
([1]Kyushu Institute of Technology, [2]Hiroshima Institute of Technology, [3]Nishinippon Institute of Technology, [4]Tokai University, Japan)

OS23-3 *Development of Subsea Creature Monitoring Station for AUV Exploration Assistance*
Keisuke Watanabe[1], Koshi Utsunomiya[1], Amir Sadiq[1], Daichi Hiramaki[1], Kyoko Takashima[1], Kazuo Ishii[2] ([1]Tokai University, [2]Kyushu Institute of Technology, Japan)

OS23-4 Consideration on Installation Method of In-situ Drilling Platform through Simulations(withdraw)
Keisuke Watanabe (Tokai University, Japan)

OS23-5 *Environment Map Generation in Forest Using Field Robot*
Noboru Takegami, Eiji Hayashi and Ryusuke Fujisawa (Kyushu Institute of Technology, Japan)

OS23-6 *Graph-Based Path Generation for Area Coverage*
Ayumu Tominaga, Eiji Hayashi, Ryusuke Fujisawa (Kyushu Institute of Technology, Japan)

Meeting Room 1
9:00-10:15 OS20 Artificial Intelligence for Embedded Systems and Robotics (5)
Chair: Hakaru Tamukoh (Kyushu Institute of Technology, Japan)

OS20-1 *A Hardware-Oriented Echo State Network for FPGA Implementation*
Kentaro Honda, Hakaru Tamukoh (Kyushu Institute of Technology, Japan)

OS20-2 *Network with Sub-Networks*
Ninnart Fuengfusin, Hakaru Tamukoh (Kyushu Institute of Technology, Japan)

OS20-3 *A Study on Fast Pick-and-Place Method for Home Service Robots using 3D point clouds*
Tomohiro Ono, Hakaru Tamukoh (Kyushu Institute of Technology, Japan)

OS20-4 *Acceleration of training dataset generation by 3D scanning of objects and Evaluation of recognition accuracy*
Yushi Abe, Yutaro Ishida, Tomohiro Ono, Hakaru Tamukoh
(Kyushu Institute of Technology, Japan)

.

OS20-5 *Anomaly Detection Using Autoencoder Trained with Reversed Color Models*
Obada Al aama, Hakaru Tamukoh (Kyushu Institute of Technology, Japan)

15:40-16:55 OS13 Natural Computing (5)
Chair: Marion Oswald (TU Wien, Austria)
Co-Chair: Yasuhiro Suzuki (Nagoya University, Japan)

OS13-1 *Emergence of Adaptive Behavior in Simulations by Using Abstract Rewriting System on Multisets*
Yasuhiro Suzuki (Nagoya University, Japan)

OS13-2 *Implementing the Euler and Runge-Kutta Method by Using Abstract Rewriting System on Multisets*
Yasuhiro Suzuki (Nagoya University, Japan)

OS13-3 *Extracting Tactile Sensation from Body Movement and Converting it into Vibrotactile Using the Tactile Score Bit*
Yasuhiro Suzuki (Nagoya University, Japan)

OS13-4 *A Tactile Sense Centered Virtual Reality Game by Using Biometric Feedback*
Yoshihito Ushida, Yasuhiro Suzuki (Nagoya University, Japan)

OS13-5 *A Method of Extracting Sensibility from Time Series Data and Converting it to Vibrotactile*
Yasuhiro Suzuki (Nagoya University, Japan)

Meeting Room 4
9:00-10:00 OS12 Advances in Theory and Education on Control (4)
Chair: Takao Sato (University of Hyogo, Japan)
Co-Chair: Masanori Takahashi (Tokai University, Japan)

OS12-1 *Proposal teaching materials the concepts and principles of machine learning for use in education*
Shinichi Imai, Yusuke Shiba (Tokyo Gakugei University, Japan)

OS12-2 *Objective Evaluation of the Educational Effects on the Feedforward, Feedback and PID Control*
Yugo Tokura, Takao Sato, Ryota Yasui, Natsuki Kawaguchi, Nozomu Araki and Yasuo Konishi (University of Hyogo, Japan)

OS12-3 *Programming Learning of Temperature Control for Science Class of Elementary School*
Yoshihiro Ohnishi[1], Takeshi Nakano[2], Teruyuki Tamai[1], Shinnosuke Mori[1], Kazuo Kawada[3] ([1]Ehime University, [2]Ehime University Elementary School, [3]Hiroshima University, Japan)

.

OS12-4 *Actuator Fault-Tolerant Control using a Spiking Neuron Model*
Masanori Takahashi (Tokai University, Japan)

15:40-17:10 OS24 Robot Intelligence and Factory Automation (6)
Chair: Eiji Hayashi (Kyushu Institute of Technology)
Co-Chair: Kazuo Ishii (Kyushu Institute of Technology)

OS24-1 *Deep Object 6-Dof Pose Estimation Using Semantic Pixel-wise Segmentation*
Victor Pujolle, Eiji Hayashi (Kyushu Institute of Technology, Japan)

OS24-2 *Autoencoder with Spiking in Frequency Domain for Anomaly Detection of Uncertainty Event*
Umaporn Yokkampon, Sakmongkon Chumkamon, Eiji Hayashi
(Kyushu Institute of Technology, Japan)

OS24-3 *Evaluation of the Relationships Between Saliency Maps and Keypoints*
Ryuugo Mochizuki, Kazuo Ishii
(Kyushu Institute of Technology, Japan)

OS24-4 *Robot Motion and Grasping for Blindfold Handover*
Jiraphan Inthiam, Sackmongkon Chumkamon, Umaporn Yokkampon, Eiji Hayashi
(Kyushu Institute of Technology, Japan)

OS24-5 Gait Learning Method for Quadrupedal Robot Using Chaos Time-series Analysis
Yuehang Ma, Kaori Watanabe, Hidekazu Suzuki
(Tokyo Polytechnic University, Japan)

OS24-6 *Development of Antagonistic High Power Joint Mechanism with Cams*
Katsuaki Suzuki[1], Yuya Nishida[1], Takashi Sonoda[2], Kazuo Ishii[1]
([1]Kyushu Institute of Technology, [2]Nishinippon Institute of Technology, Japan)

<u>January 15 (Wednesday)</u>

8:40-Registration

Conference Room

11:30-12:30

Plenary Speech PS-4

Chair: Mohamed Rizon (UCSI University, Malaysia)

.

PS-4 *Towards fully automated driving in urban areas*
Naoki Suganuma (Kanazawa University, Japan)

13:30-14:30
Plenary Speech PS-3
Chair: Marion Oswald (Technische Universität Wien, Austria)

PS-3 *Multidimensional Configurations and Strategy in Robot Soccer*
Ruggero Rossi, **Jeffrey Johnson** (Open University, UK)

Meeting Room 32
9:00-10:00 OS7 Robotic Manipulation (4)
Chair: Kensuke Harada (Osaka University, Japan)
Co-Chair: Tokuo Tsuji (Kanazawa University, Japan)
Co-Chair: Akira Nakamura (AIST, Japan)

OS7-1 *User Interface and Motion Planner for Task Database*
Tokuo Tsuji (Kanazawa University, Japan), Natsuki Yamanobe (AIST, Japan)
Kensuke Harada (Osaka University/AIST, Japan)

OS7-2 *POMDP Action Planning for 6 DoF object recognition on Humanoid*
Masato Tsuru (Osaka University, Japan), Tomohiro Motoda (Osaka University, Japan)
Adrien Escande (CNRS-AIST, Japan), Kensuke Harada (Osaka University/AIST, Japan)

OS7-3 *Cost-oriented Planning for Error Recovery in an Automation Plant*
Akira Nakamura[1], Natsuki Yamanobe[1], Ixchel Ramirez Alpizar[1],
Kensuke Harada[1,2], Yukiyasu Domae[1] ([1] AIST, Japan, [2] Osaka University, Japan)

OS7-4 *Real-time Planning Robotic Palletizing Tasks using Reusable Roadmaps*
Takumi Sakamoto (Osaka University, Japan), Weiwei Wan (Osaka University/AIST)
and Kensuke Harada (Osaka University/AIST)

10:15-11:15 OS3 Intelligent Control (4)
Chair: Yingmin Jia (Beihang University (BUAA), China)
Co-Chair: Weicun Zhang (University of Science and Technology Beijing, China)

OS3-1 *No Free Lunch Principle in Agent Swarm Systems: One Case Study*
Yunzhong Song[1], Fengzhi Dai[2], Huimin Xiao[3], Shumin Fei[4]
([1]Henan Polytechnic University, [2]Tianjin University of Science and Technology, [3]Henan University

.

of Economics and Law, [4]South East University, China)

OS3-2 *Analyzing the Controllability and Observability of Discrete-Time Delayed LTI Systems with Data-Based Methods*
Zhuo Wang (Beihang University, China)

OS3-3 *An Optimal Collective Control Strategy Based on Vicsek Model*
Yongnan Jia, Weicun Zhang, Yue Liu, Qing Li
(University of Science and Technology Beijing, China)

OS3-4 *Distributed Rotating Encirclement Control of Strict-Feedback Multi-Agent Systems Using Bearing measurements*
Tengfei Zhang, Yingmin Jia (Beihang University (BUAA), China)

14:45-16:15 OS10-1 AI Applications (6)
Chair: Mohamed Rizon (UCSI University, Malaysia)
Co-Chair: Ang Chun Kit (UCSI University, Malaysia)

OS10-1 *Effects of Variable Arm Length on UAV Control Systems*
M. Rizon, CK. Ang, MI. Solihin (UCSI University, Malaysia)
Zuradzman M. R, H. Desa, Shahriman A. B., Wan Khairunizam(UniMAP, Malaysia)
I. Zunaidi (University of Sunderland, UK)

OS10-2 *EEG based drowsiness detection using relative band power and short time fourier transform*
Pranesh Krishnan, Sazali Yaacob, Annapoorni Pranesh Krishnan (UniKL, Malaysia)
Mohamed Rizon, Ang Chun Kit (UCSI University, Malaysia)

OS10-3 *Mathematical Model Implementation of SPWM fed Three-phase Induction Motor Drive Using MATLAB Simulink*
Amir Rasyadan, Sazali bin Yaacob, Pranesh Krishnan (UniKL, Malaysia)
Mohamed Rizon, Ang Chun Kit (UCSI University, Malaysia)

OS10-4 *Implementation of X-mean Clustering Algorithm for Wireless Sensor Networks*
Abdelrahman Radwan, Nazhatul Hafizah Kamarudin, Mahmud Iwan Solihin, Hungyang Leong, Chun Kit Ang (UCSI University, Malaysia)

OS10-5 *Robust H_∞ controller design for flexible link manipulator based on constrained meta-heuristics optimization algorithms*
Mahmud Iwan Solihin, Lim Wei Hong, Chun Kit Ang, Mohamed Rizon, Abdelrahman Radwan
(UCSI university, Malaysia)

OS10-6 *Classification of Facial Nerve Paralysis Based on Regional Evaluation*
Wan Syahirah W Samsudin, Rosdiyana Samad (Universiti Malaysia Pahang, Malaysia)
Kenneth Sundaraj (Universiti Teknikal Malaysia Melaka, Malaysia)

.

Mohamed Rizon (UCSI University, Malaysia)
Mohd Zaki Ahmad (Hospital Tuanku Ampuan Afzan, Malaysia)

16:30-17:00 OS10-2 AI Applications (2)
Chair: Mohamed Rizon (UCSI University, Malaysia)
Co-Chair: Ang Chun Kit (UCSI University, Malaysia)

OS10-7 *Introduction of Forehead Lesion Assessment with House-Brackmann Score for Facial Nerve Paralysis Evaluation*
Wan Syahirah W Samsudin, Rosdiyana Samad (Universiti Malaysia Pahang, Malaysia)
Kenneth Sundaraj (Universiti Teknikal Malaysia Melaka, Malaysia)
Mohamed Rizon (UCSI University, Malaysia)
Mohd Zaki Ahmad (Hospital Tuanku Ampuan Afzan, Malaysia)

OS10-8 *Intelligent Wearable Biofeedback Fuzzy Logic Based Device for Monitoring and Treatment of Voice Loudness*
Ali S. AlMejrad (King Saud University, Kingdom of Saudi Arabia)

Meeting Room 33
9:00-10:00 GS7 Poster (4)
Chair: Jiwu Wang (Beijing Jiaotong University, China)

POS7-1 *The Development and Evaluation of Fig's Leaf Syrup*
Shang-Hui Li, Pei Hi Zheng, I Chih Chiang, Yu Ting Su, Syue Sheng Lin (Far East University, Taiwan)

POS7-2 *The Research and Development of Fruit Puffed Rice*
Shang-Hui Li, Yi-Ru Wang, Yi Ting Liu, En-Yi Lu, Cheng Han Li, Fang Quan Zhang (Far East University, Taiwan)

POS7-3 *The Research of Heath Western Cusine – A Study of Aloe in Cooking*
HsiYing Hsieh (Far East University, ROC Taiwan)

POS7-4 *The Influence of Attitude, Subjective Norm, Perceived Behavior Control on Purchase Intention – A Study of the Green Restaurants in Taiwan*
HsiYing Hsieh (Far East University, ROC Taiwan)

10:15-11:15 OS8 Advanced Research of Engineering and Management (4)
Chair: Chair: Takao Ito (Hiroshima University, Japan)
Co-Chair: Ammar A. Al-Talib (UCSI University, Malaysia)

OS8-1 *A Rack and Pinion Driven Mechanical Footstep Power Generator*
Ammar A. Al-Talib, K S See (UCSI University, Malaysia)

.

OS8-2 *Table Tennis Using Arduino For Seniors' Healthcare*
Reem Ali Abdullah, Mastaneh Mokayef, Miad Mokayef, Sew Sun Tiang, Wei Hong Lim
(UCSI University, Malaysia)

OS8-3 *A Study on Prevention of Predictive Failures using IoT Technology*
Tsutomu Ito[1], Hiroshi Sakai[1], Takao Ito[2], Seigo Matsuno[3], Makoto Sakamoto[4]
([1]Hino Motors Ltd, [2]Hiroshima University, [3]Ube National College of technology, [4]University of Miyazaki, Japan)

OS8-4 *Did a mismatch between the ASBJ's standard-setting and its organizational structure occur?*
Kensuke Ogata (Osaka City University, Japan)

14:45-15:45 OS16 Media Information Processing and Artificial Intelligence (4)
Chair: Yasunari Yoshitomi (Kyoto Prefectural University, Japan)
Co-Chair: Masayoshi Tabuse (Kyoto Prefectural University, Japan)

OS16-1 *Effectiveness of Data Augmentation in Pointer-Generator Model*
Tomohito Ouchi, Masayoshi Tabuse (Kyoto Prefectural University, Japan)

OS16-2 *Mouse Cursor Control System Using Facial Movements*
Masayoshi Tabuse[1], Manase Mizobe[2], Yasunari Yoshitomi[1], Taro Asada[1]
([1]Kyoto Prefectural University, [2]TORAY ENGINEERING Co.,Ltd, Japan)

OS16-3 *Facial Expression Synthesis Using Vowel Recognition for Synthesized Speech*
Taro Asada[1], Ruka Adachi[2], Syuhei Takada[3], Yasunari Yoshitomi[1], Masayoshi Tabuse[1]
([1]Kyoto Prefectural University, [2]Software Service, Inc., [3]Seika Town Hall, Japan)

OS16-4 *Speech Synthesis of Emotions in a Sentence Using Vowel Features*
Rintaro Makino[1], Yasunari Yoshitomi[2], Taro Asada[2], Masayoshi Tabuse[2]
([1]SoftBank Corp., [2]Kyoto Prefectural University, Japan)

16:30-17:45 OS19 Advanced Information Processing Applications (5)
Chair: Toru Hiraoka (University of Nagasaki, Japan)
Co-Chair: Masaharu Hirota (Okayama University of Science, Japan)

OS19-1 *Extraction of Irrelevant Sentences from Online Hotel Reviews*
Shogo Watanabe, Masaharu Hirota, Tetsuya Oda (Okayama University of Science, Japan)

OS19-2 *Inferring Home Location of Foreign Tourists Based on Travel Routes Extracted from Social Media Sites*
Lugasi Chen, Masaharu Hirota (Okayama University of Science, Japan)

OS19-3 *The IoT Solution to Archive and Play the Digital Library of Kamishibai*
Motohide Yoshimura, Ayumi Eikawa (The University of Nagasaki, Japan)

.

OS19-4 *A Research on Prediction of Inter-firm Relationships*
Takao Ito[1], Tsutomu Ito[2], Matsuno Seigo[3], Rajiv Mehta[4], Makoto Sakamoto[5]
([1]Hiroshima University, Japan, [2]Hino Motors Ltd, Japan, [3]Ube National College of Technology, Japan, [4]New Jersey Institute of Technology, USA, [5]University of Miyazaki, Japan)

OS19-5 *Generation of Arbitrarily-Oriented Ripple Images Using Circular-Sector-Type Smoothing Filter and Inverse Filter*
Toru Hiraoka (University of Nagasaki, Japan)

Meeting room 1

9:00-10:15 OS18-2 Service Robotics (5)

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

OS18-1 *Modelling Autonomous Parallel Parking Procedure for Car-like Robot Avrora Unior in Gazebo Simulator*
Dinir Imameev[1], Ksenia Shabalina[1], Artur Sagitov[1], Kuo-Lan Su[2], Evgeni Magid[1]
([1]Kazan Federal University, Russia)
([2]National Yunlin University of Science and Technology, Taiwan)

OS18-2 *Traffic Sign Recognition Algorithm for Car-like Robot Avrora Unior*
Nikita Nikiforov[1], Ksenia Shabalina[1], Artur Sagitov[1], Kuo-Hsien Hsia[2], Evgeni Magid[1]
([1]Kazan Federal University, Russia)
([2]National Yunlin University of Science and Technology, Taiwan)

OS18-3 *Robotic palpation modeling for KUKA LBR IIWA using Gazebo Simulator*
Artur Shafikov[1], Artur Sagitov[1], Hongbing Li[2], Natalia Schiefermeier-Mach[3], Evgeni Magid[1]
([1]Kazan Federal University, Russia), ([2]Shanghai Jiao Tong University, China)
([3]Health University of Applied Sciences, Austria)

OS18-4 *Remote Control Application for "Servosila Engineer" Robot on Android Mobile Devices*
Daniel Kiryanov, Roman Lavrenov (Kazan Federal University, Russia)

OS18-5 *Network Failure Detection and Autonomous Return for PMB-2 mobile robot*
Dmitry Bereznikov, Aufar Zakiev (Kazan Federal University, Russia)

Meeting room 4

9:00-10:00 GS4 Robotics 1 (4)

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

GS4-1 *Design and Modeling of an Automatic Cartesian Farming Robot*
Wisanu Jitviriya, Amornphun Phunopas
(King Mongkut's University of Technology North Bangkok, Thailand)
Eiji Hayashi (Kyushu Institute of Technology, Japan)

GS4-2 *K-APF Algorithm to Avoid Obstacles in Path Planning*

.

Dong-Kyo Jeong, Jang-Myung Lee (Pusan National University, Korea)

GS4-3 *Selection System of Robot type for cell assembly production (Production efficiency comparison of single arm robot and double arm robot)*
Takahiro Watanabe, Hidehiko Yamamoto, Takayoshi Yamada (Gifu University, Japan)

GS4-4 *Curvature Surface Magnetic Wheel Climbing Robot with Adaptive Electromagnetic Adhesive Force*
Richit Palangwatanakul, Apisit Thungsang, Kaned Thungod (Mahasarakham University, Thailand)
Arsit Boonyaprapasorn, Wanachart Borisut (Chulachomklao Royal Military Academy, Thailand)
Thavida Maneewarn, Suriya Natsupakpong
(King Mongkut's University of Technology Thonburi, Thailand)
Thunyaseth Sethaput
(Sirindhorn International Institute of Technology, Thammasat University, Thailand)

10:15-11:15 GS5 Robotics 2 (4)
Chair: Hidehiko Yamamoto (Gifu University, Japan)

GS5-1 *Intention Classification of a User of a Walking Assist Cart by Using Support Vector Machine*
Noritaka Sato, Tomoki Yokotani, Yoshifumi Morita (Nagoya Institute of Technology, Japan)

GS5-2 *Development of an Environmentally Adaptable Autonomous Mobile Robot*
Naokazu Iwata, Joo Kooi Tan (Kyushu Institute of Technology, Japan)

GS5-3 *Automated Task and Path Management for Industrial AGVs in Foam Manufacturing Plant*
Amornphun Phunopas[1], Wisanu Jitviriya[2], Noppadol Pudchuen[3], Songklod Tunsiri[4], Eiji Hayashi[5]
([1,2,3]King Mongkut's University of Technology North Bangkok, Thailand,
[4]Urban Community Development College, Thailand, [5]Kyushu Institute of Technology, Japan)

GS5-4 *Object Tracking Method Considering Time Series Information Using Re3 with Stochastic Depth*
Taichi Kitayama, Hyoungseop Kim (Kyushu Institute of Technology, Japan)

14:45-16:15 GS1 Neural Networks (6)
Chair: Akira Nakamura (AIST, Japan)

GS1-1 *Neural Network and Internal Resistance based SOH classification for lithium battery*
Jong-Hyun Lee[1], Hyun-Sil Kim[2], In-Soo Lee[1] ([1]Kyungpook National University, Korea)
([2]Naval Combat Systems PMO Agency For Defense Development, Korea)

GS1-2 *Estimation of Self-Posture of a Pedestrian Using MY VISION and Deep Learning*
Tomoyuki Kurosaki, Joo Kooi Tan (Kyusyu Institute of Technology, Japan)

.

GS1-3 *Simultaneous Space Object Recognition and Pose Estimation by Convolutional Neural Network*
Roya Afshar, Zhongyi Chu (Beihang University, China)
Shuai Lu (Beijing University of Chemical Technology, China)

GS1-4 *An error correction mechanism for reliable chemical communication systems*
Masashi K. Kajita (The University of Tokyo, Japan)

GS1-5 *A Reinforcement Learning-Based Path Planning Considering Degree of Observability*
Yong Hyeon Cho, Chan Gook Park (Seoul National University, Korea)

GS1-6 *A Performance Analysis of Pose Estimation Based on Two-View Tracking and Multi-State Constraint Kalman Filter Fusion*
Tae Ihn Kim, Jae Hyung Jung, Chan Gook Park (Seoul National University, Korea)

16:30-17:45 GS2 Control Techniques (5)
Chair: Chian C. Ho (National Yunlin University of Science and Technology, Taiwan)

GS2-1 *ORB-SLAM based Sensor Fusion Algorithm for Real-Time Precision Driving*
Yong·Jin Ock, Zhan·Ming Gu, Jang·Myung Lee (Pusan National University, Korea)

GS2-2 *Gait Control of A Four-Legged Robot with Fuzzy-PID Controller*
Arphakorn Kunha, Amornphun Phunopas, Wisanu Jitviriya
(King Mongkut's University of Technology North Bangkok, Thailand)

GS2-3 *Hybrid Force/Position Teaching and Control Method for 6 DoF Manipulator utilizing f-PAWTED*
Quang-Trung Chu, Hiroki Tanaka, Hideki Inuzuka, Yoshifumi Morita (NITech, Japan)
Masao Sakai (Aichi Pref., Japan)

GS2-4 *A Study on Generalized Predictive Control in Consideration of Noise*
Akira Yanou (Kawasaki University of Medical Welfare, Japan)

GS2-5 *Simulation Study on Emergency-Stopping Avoidance Control due to Singularity During Teaching Operation with Parallel Wire-Type Teaching Device*
Hideki Inuzuka (NITech, Japan), Masao Sakai (Aichi Pref., Japan), Yoshifumi Morita (NITech, Japan)

January 16 (Thursday)

8:40-Registration

.

Meeting Room 31
10:30-11:00
Invited session IS-1
Chair: Yingmin Jia (Beihang University (BUAA), China)

IS-1 *Concurrent Localization of Multiple Unmanned Surface Vehicles Using Neural Networks*
Jangmyung Lee (Pusan National University, Korea)

12:00-13:15 OS11 System and Control (16)
Chair: Yizhun Peng (Tianjin University of Science and Technology, China)
Co-Chair: Huailin Zhao (Shanghai Institute of Technology, China)

OS11-1 *Design of Humanoid Soccer Robot Based on STM32*
Yuheng Zhang[1], Yulong Peng[2], Yizhun Peng[1], Lianchen Zhao[1], Zhou Zhang[1], Wanlong Peng[1]
([1]Tianjin University of Science and Technology, [2] Tianjin Tianke Intelligent and Manufacture Technology Co., LTD, China)

OS11-2 *A Design and Implementation of Intelligent Cradle*
Ting Zhao[1], Qing He[1], Yulong Peng[2], Zhou Yang[1], Zhenjiang Chen[1], Shuo Jiang[1]
([1]Tianjin University of Science and Technology, [2] Fenyang College of Shanxi Medical University, China)

OS11-3 *A Design and Implementation of Intelligent Networking Bookcase*
Zhou Zhang, Yajun Li, Yizhun Peng, Hucheng Wang, Yuqi Zhao
(Tianjin University of Science and Technology, China)

OS11-4 *A Design and Implementation of Quad-rotor UAV*
Junjie Lin, Chunxia Zhang, Yizhun Peng, Ting Zhao, Zhengke Xu
(Tianjin University of Science and Technology, China)

OS11-5 *Characteristic Analysis and Synchronization Control of a Non-equilibrium System*
Lianchen Zhao, Xinyu Zhang, Yizhun Peng, Yuheng Zhang, Hongyan Jia,
(Tianjin University of Science and Technology, China)

OS11-6 *Classification and Recognition of Baby Cry Signal Feature Extraction Based on Improved MFCC*
Zhenjiang Chen, Yizhun Peng, Di Li, Zhou Yang, Nana Wang
(Tianjin University of Science and Technology, China)

OS11-7 *Design of Space Remote Sensing Data Storage Platform Based on Distributed File System*
Di Li, Yizhun Peng, Ruixiang Bai, Zhenjiang Chen, Lianchen Zhao
(Tianjin University of Science and Technology, China)

OS11-8 *Research on Semantic Map Establishment of Parking Lot Based on Deep Learning and Multi-sensor*

.

Shiqian Zhang, Yizhun Peng, Ruixiang Bai, Yuheng Zhang, Tianye Jian, Wanlong Peng
(Tianjin University of Science and Technology, China)

OS11-9 *Survey on Kinematics Calibration Technology of Manipulator*
Zhou Yang, Yizhun Peng, Nana Wang, Yuheng Zhang, Tianye Jian
(Tianjin University of Science and Technology, China)

OS11-10 *Visualization Analysis of Web Crawler Evolution Retrieval Research Based on KG*
Zhenjiang Chen, Jiamian Wang, Yizhun Peng, Di Li, Lianchen Zhao
(Tianjin University of Science and Technology, China)

OS11-11 *Circuit Simulation of Synchronized Novel 4D Chaotic Systems*
Yuhan Zhang, Hong Niu (Tianjin University of Science and Technology, China)

OS11-12 *Crowd Counting Network with Self-attention Distillation*
Li Wang, Huailin Zhao, Zhen Nie, Yaoyao Li (Shanghai Institute of Technology, China)

OS11-13 *Path Planning Based on Improved Artificial Potential Field Method*
Feifan Xu, Huailin Zhao, Zhen Nie, Xin Zhou, Zheheng Tao
(Shanghai Institute of Technology, China)

OS11-14 *Self-balancing Car based on Adaptive Fuzzy PID Control*
Zhen Nie, Huailin Zhao, Lu Sun, Xiongfeng Zhong
(Shanghai Institute of Technology, China)

OS11-15 *Crowd Counting Method Based on Improved CSRnet*
Huailin Zhao, Shengyang Lu, Li Wang, Yaoyao Li (Shanghai Institute of Technology, China)

OS11-16 *Graph-based Global Reasoning Network for Crowd Counting*
Li Wang, Huailin Zhao, Zhen Nie, Yaoyao Li
(Shanghai Institute of Technology, China)

13:30-14:45 OS9 Recognition and Control (19)
Chair: Fengzhi Dai (Tianjin University of Science and Technology, China)
Co-Chair: Qiang Wei (Jianghan University, China)

OS9-1 *A Design of Ocean Current Velocity Measuring Device*
Jichao Zhao[1], Fengzhi Dai[1,3], Xin Ma[2], Fengkun Wang[1], Haokang Wen[1], Hongbo Hao[1], Qianqian Zhang[1]
([1]Tianjin University of Science and Technology, [2]National Oceanic Administration Beihai Marine Environmental Monitoring Center Station, [3]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

OS9-2 *Device Design Based on TDS Water Quality Detection*
Jichao Zhao[1], Fengzhi Dai[1,2], Di Yin[1], Yuhui Cheng[1], Fengkun Wang[1], Leixin Han[1], Qianqian Zhang[1]
([1]Tianjin University of Science and Technology, [2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

OS9-3 *Design of Film Forming Rate Measuring Instrument based on Polyurethane Material*
Jichao Zhao[1], Fengzhi Dai[1,2], Fengkun Wang[1], Haokang Wen[1], Hongbo Hao[1], Qianqian Zhang[1]
([1]Tianjin University of Science and Technology, [2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

OS9-4 *Research on the Intelligent Aircraft Design based on STM32*
Hongbo Hao[1], Fengzhi Dai[1,2], Haokang Wen[1], Jichao Zhao[1], Qianqian Zhang[1]
([1]Tianjin University of Science and Technology, [2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

OS9-5 *Research and Design of Gain Controllable System in RF Receiver*
Haokang Wen[1], Fengzhi Dai[1,2], Jichao Zhao[1], Hongbo Hao[1], Hong Niu[1], Qianqian Zhang[1]
([1]Tianjin University of Science and Technology, [2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

OS9-6 *Design of Packaging Bottle Recycling System based on Internet of Things*
Yujie Yan[1], Fengzhi Dai[1,2], Shiwei Wu[1], Yuanyuan Xi[1], Huanhuan Li[1]
([1]Tianjin University of Science and Technology, [2] Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China, China)

OS9-7 *Research on Image Super-Resolution Reconstruction Based on Deep Learning*
Lingran An[1],Fengzhi Dai [1,2], Yasheng Yuan[1]
([1]Tianjin University of Science and Technology, [2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

OS9-8 *Research on Fatigue Detection Method Based on Deep Learning*
Yasheng Yuan[1], Fengzhi Dai[1,2], Lingran An[1], Di Yin[1], Yuxuan Zhu[1], Yujie Yan[1]
([1]Tianjin University of Science and Technology, [2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

OS9-9 *Fatigue Driving Monitoring System based on the EEG*
Yuxuan Zhu[1], Fengzhi Dai[1,2], Di Yin[1], Yasheng Yuan[1]
([1]Tianjin University of Science and Technology, [2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

OS9-10 *Research of the Control Strategy of Vienna Rectifier Circuit based on the Vector Control*
Yuxuan Zhu[1], Fengzhi Dai[1,2], Di Yin[1], Yasheng Yuan[1]
([1]Tianjin University of Science and Technology, [2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

OS9-11 *Research on Emotion Classification based on EEG*
Di Yin[1], Fengzhi Dai[1,3], Mengqi Yin[2], Jichao Zhao[1]
([1]Tianjin University of Science and Technology, [2]Hebei University of Chinese Medicine, [3]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

OS9-12 *Identification of Synthetic Pigment based on Fluorescence Spectroscopy Combined with RBF Neural Network*

Di Yin[1], Fengzhi Dai[1,2], Yuxuan Zhu[1], Yasheng Yuan[1]
([1]Tianjin University of Science and Technology, [2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

OS9-13 *Simulation Study of 3D Reconstruction in Electromagnetic Tomography with Multi-layer Sensors Array*
Yuanli Yue, Ze Liu, Yu Miao (Beijing Jiaotong University, China)

OS9-14 *Visualized the Knowledge Map in Children's Minds: A Study on Cognitive Structure Measurement*
Qiang Wei, Hua Dong, Yi-tong Zhang, Ao-nan Zhang (Jianghan University, China)

OS9-15 *Escape Route of Subway under Fire Conditions: An Experimental Study in Virtual Reality Environment*
Hua Dong, Qiang Wei, Qing-qing Zhang, Lan-lan Fang (Jianghan University, China)

OS9-16 *Research on the Smart Home Design based on Single-chip Microcomputer*
Hongbo Hao[1], Fengzhi Dai[1, 2], Haokang Wen[1], Jichao Zhao[1]
([1]Tianjin University of Science and Technology, [2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

OS9-17 *Design of a Portable Instrument for Measuring Heart Rate and Blood Oxygen*
Haokang Wen[1], Fengzhi Dai[1,2], Jichao Zhao[1], Hongbo Hao[1], Qianqian Zhang[1]
([1]Tianjin University of Science and Technology, [2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

OS9-18 *Principal Component Analysis of Wine Based on Three-dim Fluorescence Spectra*
Di Yin, Fengzhi Dai, Yuxuan Zhu, Yasheng Yuan
(Tianjin University of Science and Technology, China)

OS9-19 *Control Design of Intelligent Device for Living Environment of Senile Apartment*
YashengYuan[1], Fengzhi Dai[1,2], Shengbiao Chang[3], Lingran An[1], Di Yin[1],
Yuxuan Zhu[1], Yujie Yan[1]
([1]Tianjin University of Science and Technology, [2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, [3]Dawning Information Industry Co., LTD, China)

Meeting Room 32
9:00-10:00 OS2 Advanced Control (4)
Chair: Yingmin Jia (Beihang University (BUAA), China)
Co-Chair: Weicun Zhang (University of Science and Technology Beijing, China)

OS2-1 *Multiple Model Adaptive Control Based on Switching/Weighting Intelligent Fusion Algorithm*
Weicun Zhang, Yongnan Jia, Qing Li
(University of Science and Technology Beijing, China)

OS2-2 *Detection Algorithm of Porosity and Crack Defects on Surface of Micro-precision Glass Insulated Terminals*
Qunpo Liu[1], Gaowei Wang[1], Fuzhong Wang[1], Manli Wang[1], Hanajima Naohiko[2]

([1]Henan Polytechnic University, P.R.China) ([2]Muroran Institute of Technology, Japan)

OS2-3 *Concept Drift Adaptation by Multi-stream Data Knowledge Transfer*
Hongpeng Yin, Chenglin Liao, Yi Chai (Chongqing University, China)

OS2-4 *Neuro-adapative Control of High-speed Trains under Uncertain Wheel-rail Relationship*
Zhechen Wang, Yingmin Jia (Beihang University (BUAA), China)

12:00-13:00 OS1 Machine vision and Mobile robot (4)
Chair: Wei Liu (Beijing Jiaotong University, China)
Co-Chair: Jiwu Wang (Beijing Jiaotong University, China)

OS1-1 *Route Planning of Teleoperation Mobile Robot Based on Virtual Reality Technique*
Jiwu Wang, Xuechun Yuan, Chenyang Li, Zhikun Song (Beijing Jiaotong University, China)

OS1-2 *Geometric Measurement Based on The Single Image with a Rectangle Structure*
Jiwu Wang, Bo Dai, Jiangyue Wang (Beijing Jiaotong University, China)

OS1-3 *Heavy-duty Spherical Mobile Robot Driven by Five Omnidirectional Wheels*
Wei Liu, Junyang Sun, Runjiao Wang, Guanwang Geng, Lian Luo
(Beijing Jiaotong University, China)

OS1-4 *Robustness Analysis of Visual SLAM Based on Gazebo Simulation Environment*
Jiwu Wang, Yafan Liu (Beijing Jiaotong University, China)

13:30-14:30 OS21 Mathematical Informatics (4)
Chair: Makoto Sakamoto (University of Miyazaki, Japan)
Co-Chair: Amane Takei (University of Miyazaki, Japan)

OS21-1 *Hidden Surface Removal for Interactions between User's Bare Hands and Virtual Objects in Augmented Reality*
Takahiro Ishizu, Makoto Sakamoto, Kenji Sakoma, Takahiro Shinoda, Amane Takei
(University of Miyazaki, Japan), Takao Ito (Hiroshima University, Japan)

OS21-2 *Proposal of Interactive Projection Mapping using Human Detection by Machine Learning*
Takahiro Shinoda, Makoto Sakamoto, Takahiro Ishizu, Kenji Sakoma, Amane Takei
(University of Miyazaki, Japan), Takao Ito (Hiroshima University, Japan)

OS21-3 *Fundamental Study on Control of CG Characters by Electroencephalography (EEG) Analysis*
Kenji Sakoma, Makoto Sakamoto, Takahiro Ishizu, Takahiro Shinoda, Amane Takei
(University of Miyazaki, Japan), Takao Ito (Hiroshima University, Japan)

.

OS21-4 *Development of parallel microwave analysis code: ADVENTURE_Fullwave*
Amane Takei (University of Miyazaki, Japan)

Meeting Room 33
9:00-10:15 OS14 Software Development Support Method (5)
Chair: Tetsuro Katayama (University of Miyazaki, Japan)
Co-Chair: Tomohiko Takagi (Kagawa University, Japan)

OS14-1 *Behavioral Modeling Technique for Multiple Objects of Software Using Extended Place /Transition Nets with Attributed Tokens*
Tomohiko Takagi, Ryo Kurozumi (Kagawa University, Japan)

OS14-2 *Learning Support Technique of Software Visual Modeling Using Place/Transition Nets*
Yuki Ue, Tomohiko Takagi (Kagawa University, Japan)

OS14-3 *Redundant Test Cases Elimination on Code Coverage with Distance and Correlation Measurement Method*
Mochamad Chandra Saputra[1], Tetsuro Katayama[1], Yoshihiro Kita[2], Hisaaki Yamaba[1], Kentaro Aburada[1], and Naonobu Okazaki[1]
([1]University of Miyazaki, [2]Tokyo University of Technology, Japan)

OS14-4 *The Measurement of Class Cohesion using Semantic Approach*
Bayu Priyambadha[1], Tetsuro Katayama[1], Yoshihiro Kita[2], Hisaaki Yamaba[1], Kentaro Aburada[1], and Naonobu Okazaki[1]
([1]University of Miyazaki, [2]Tokyo University of Technology, Japan)

OS14-5 *Proposal of an Algorithm to Generate VDM++ by Using Words Extracted from the Natural Language Specification*
Yasuhiro Shigyo[1], Tetsuro Katayama[1], Yoshihiro Kita[2], Hisaaki Yamaba[1], Kentaro Aburada[1], Naonobu Okazaki[1] ([1]University of Miyazaki, [2]Tokyo University of Technology, Japan)

12:00-12:30 OS17 Machine Learning Technologies for Human Understanding (2)
Chair: Masao Kubo (National Defense Academy of Japan, Japan)

OS17-1 *Customization of Contents for Acquisition of Skills of FPS without Trainer*
Masao Kubo, Takeshi Ueno, Hiroshi Sato (National Defense Academy, Japan)

OS17-2 *Visual Classification of Malware by Few-shot Learning*
Tran Kien, Masao Kubo, Hiroshi Sato (National Defense Academy, Japan)

13:30-14:30 GS3 Vision & Image Processing (4)

.

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

GS3-1 *Detecting Pedestrians and Moving Directions by a MY VISION System*
Kenta Hori, Seiji Ishikawa, Joo Kooi Tan (Kyushu Institute of Technology, Japan)

GS3-2 *Human Motion Recognition Using TMRIs*
Cao Jing, Youtaro Yamashita, Joo Kooi Tan (Kyusyu Institute of Technology, Japan)

GS3-3 *Automatic Extraction of Abnormalities on Temporal CT Subtraction Images Using Sparse Coding and 3D-CNN*
Yuichiro Koizumi, Noriaki Miyake, Huimin Lu, Hyoungseop Kim
(Kyushu Institute of Technology, Japan)
Takatoshi Aoki (University of Occupational and Environmental Health, Japan)
Shoji Kido (Osaka University, Japan)

GS3-4 *Design of a Data-Driven Multi Controllers Using VRFT and Ensemble Learning*
Takuya Kinoshita, Yuma Morota, Toru Yamamoto (Hiroshima University, Japan)

OS15 Recognition and System (11)

OS15-1 *Efficient Detection Device for Wafer Physical Defects*
Jianyong Chen, Xiaoyan Chen, Chundong Zhao
(Tianjin University of Science and Technology, China)

OS15-2 *Wafer Defect Detection Method based on Machine Vision*
Chundong Zhao, Xiaoyan Chen, Jianyong Chen
(Tianjin University of Science and Technology, China)

OS15-3 *Analysis and Circuit Design of a Novel 4D Chaotic System*
Yan Sun, Yongchao Zhang, Jiaqi Chen (Tianjin University of Science and Technology, China)

OS15-4 *Research on Synchronous Control of a Novel 4D Dissipative Chaotic System*
Jiaqi Chen, Yongchao Zhang, Yan Sun (University of Science and Technology, China)

OS15-5 *EEG classification based on common spatial pattern and LDA*
Lei Wang[1], Zixuan Li[2] ([1]Tianjin University of Science and Technology, China)
([2]Dongbei University of Finance & Economics, China)

OS15-6 *Dynamic Characteristics Analysis of the Shimizu–Morioka Chaotic System*
Wenxin Shi, Hongyan Jia (Tianjin University of Science and Technology, China)

OS15-7 *Research on the Motion Track of High-speed Objects*

.

Qianqian Zhang[1,2], Fengzhi Dai[1], Jichao Zhao[1], Haokang Wen[1], Hongbo Hao[1]
([1]Tianjin University of Science and Technology, [2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

OS15-8 *Research on Surface Defect Detection of Aluminum based on Image Processing*
Xuemin Liu[1], Ce Bian[3], Di Yin[2], Yuxuan Zhu[2], Yasheng Yuan[2]
([1]China Petroleum Engineering and Construction Corporation, [2] Tianjin University of Science and Technology, [3]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

OS15-9 *Research on the Control of Multi Position Production Line based on PLC*
Yong Hou[1], Hao Wang[1], Runhua Mao[2], Xuemin Liu[3]
([1]Tianjin University of Science and Technology, [2]Tianjin Tianke Intelligent and Manufacture Technology CO, LTD, [3]China Petroleum Engineering and Construction Corporation, China)

OS15-10 *Research on acoustic source localization system based on acoustic holography*
Xiuqing Wang, Xiaoyun Jia (Tianjin University of Science and Technology, China)

OS15-11 *Anti-interference Method of Electrical Fast Transient for Fire Alarm of Substation*
Guo Wangyong, Ju Zhenfu, Chen Guang, Zhu Bo
(Beijing NARI Yihe Environmental Protection Technology Co., Ltd, China)

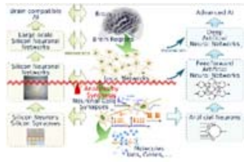
Farewell Party

PS abstracts

PS-1 Towards neuromimetic computing

Takashi Kohno (University of Tokyo, Japan)

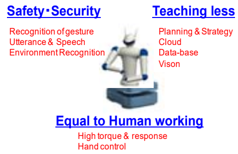
Deep neural network brought a new era of neuro-inspired machine learning. Its splendid performance in pattern recognition is attracting both scientific and industrial resources and accelerating the field of "AI". In these days, neuromimetic computing is developing its presence in this field as a potential approach to the post-deep neural network, the next-generation AI. Silicon neuronal network (SNN) is an approach to the neuromimetic computing, which tries to construct an electronic circuit version of the nervous system. Its strong potential to innovate the computing paradigm from digital computing to "brain-compatible" computing facilitated IBM and Intel to developed SNN chips, but it still stays in the embryonic stage because of a great hurdle. In this talk, we overview and discuss about trials in this field.



PS-2 Innovative Robot - Robot/AI for Factory Automation-

Eiji Hayashi (Kyushu Institute of Technology, Japan)

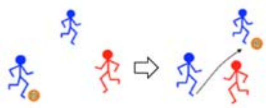
We have engaged in a regional revitalization project during the period from 2018 to 2022 as a subsidy grant which the Japan's government are promoting, by Yaskawa Electric Corporation, City of Kitakyushu and Kyushu Institute of Technology. In this project has been developed innovative robots combined Robot / Artificial Intelligence technologies. We think that AI and robots are already everywhere, especially not engineers due to the influence of the media. Certainly, their technologies have made enormous advances and progress in recent years like Siri, self-driving car system and so on as well as deep-learning neural network on image processing. The innovative robot for FA robot which means a revolution in productivity into next generation is described about what such a robot needs.



PS-3 Multidimensional Configurations and Strategy in Robot Soccer

Ruggero Rossi, Jeffrey Johnson (Open University, UK)

Recognizable structures appear repeatedly in robot soccer. Typically they involve configurations of one to five players from each team. Some structures are highly desirable as part of temporal sequences with the ball moving forwards towards the goal with the player structures, where some sequences ending in goals. A particular structure is the *defenders dilemma* when a red player attempts to win the ball from a blue player. If red attempts to tackle the blue player it passes the ball to another blue player. If red moves to prevent this the blue player with the ball dribbles past him. We illustrate these configurations and show that they give a tactical advantage. A winning strategy is to move to form these structures while denying them to the opposition.



PS-4 Towards fully automated driving in urban areas.

Naoki Suganuma (Kanazawa University, Japan)

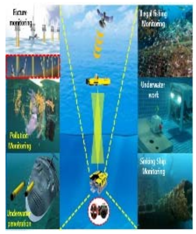
In recent years, research on autonomous driving technology has been actively conducted. In our laboratory, research and development on autonomous vehicle started from 1998 aiming for autonomous driving in urban areas. Additionally, our laboratory started the first public road test among Japanese University in Japan from 2015, and it has already been conducted public road experiment over 14,000 km. In this presentation, we will outline the technologies required for autonomous driving based on the knowledge obtained from these achievements and describe the current state and issues of autonomous driving technologies.

IS abstracts

IS-1 Concurrent Localization of Multiple Unmanned Surface Vehicles Using Neural Networks

Jangmyung Lee (Pusan National University, Korea)

A concurrent localization of multiple USVs (Unmanned Surface Vehicle) has been proposed using neural networks in order to resolve the problems of the probability-based filters used in the existing USV localization. The multiple USVs are effective for port surveillance, ocean reconnaissance, and so on, when the concurrent localization is available. Several probability-based filters, such as, EKF, KF, and UKF have been utilized so far. Since these algorithms are externally affected by wind and waves on the sea surface, it becomes difficult to accurately control the navigation along the desired trajectory. Using the proposed neural network, the multiple USVs can be navigated to survey the under-water condition effectively



IS-2 Body and Brain Training with Big Data and AI

Henrik Hautop Lund (Technical University of Denmark, Denmark)

Utilizing a Big Data and AI approach, we developed a novel playful method for screening people for potential physical and cognitive shortages. The method creates a body and brain performance map for each individual, and the Big Data analysis provides a basis for automatically identifying the particular abilities, which may be underperforming in an individual. Further, several studies including randomized controlled trials with the Moto Tiles system have shown that particular Moto Tiles game play will increase performance of particular abilities, even after short-term play. Thereby, the proposed system can automatically generate personalized training protocols for the individual by selecting and providing the right Moto Tiles games for the individual to play to improve the underperforming abilities. The suitability of the method was tested in a small effect test with seniors with mild dementia at a care institution in Denmark. The results show that the seniors with dementia who were screened to be at high risk of falling, within the short period of training with the automatically generated personalized protocol increased their skills to no longer be at risk of falling.

IS-3 PixelBeing

Luigi Pagliarini[1,2,3], Henrik Hautop Lund[1]
([1]Academy of Fine Arts of Macerata, Italy), ([2]ISIA Design, Italy)
([1]Technical University of Denmark, Denmark)

In this conceptual paper, we describe and define the range of possible applications and the technical contours of a robotic biotechnological system to be worn on the body for playful interactions. Moving from earlier works on Wearable and Modular Robotics we described how, by using modular robotics for creating wearable, it is possible to explore a self-sustainable and flexible system, consisting of freely inter- changeable input/output modules that through the use of solar, mechanical, and other sources of renewable energy are able to suit some specific tasks. Here, we drive the attention on early prototypes to show the potentialities of such an approach, and focus on depicting possible application in the future of sustainable electronics domain. Indeed, our artistic experiment is a clear example of how to scale down electronics to an eco-friendly level, which can still create playful and useful interactions for many applications.

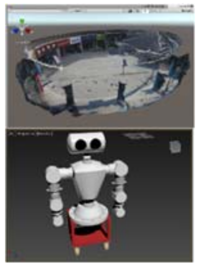
OS abstracts

OS1 Machine vision and Mobile robot (4)

OS1-1 Route Planning of Teleoperation Mobile Robot Based on Virtual Reality Technique

Jiwu Wang, Xuechun Yuan, Chenyang Li, Zhikun Song (Beijing Jiaotong University, China)

Mobile teleoperation robot is one of the effective methods to help operators to work in complex environments. However, the time delay by distance is a key factor that restricts its application. To solve this problem, the motion trajectory of the robot is simulated using virtual reality technology and the obtained optimization data are applied to control teleoperation robot. In order to improve the efficiency of simulation, a 3D simulation environment of the real world is built in the unity 3D platform, and the physical model and kinematic model of the robot are established. At the same time, the environmental conditions of various adjustable parameters such as illumination and material are simulated in the platform. The physics engine for the robot is added in the simulation platform and the physical properties of the robot are configured. Finally, the error between the simulated and real motion trajectories is analyzed, and the results show the developed method is effective to solve the time delay for the remote control.



OS1-2 Geometric Measurement Based on The Single Image with a Rectangle Structure

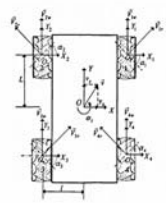
Jiwu Wang, Bo Dai, Jiangyue Wang (Beijing Jiaotong University, China)

Image-based geometric measurement is getting more attention in machine vision field due to its contact-less and low-cost characteristics. Here the measurement of single image with the rectangle structure is studied, in which only one side length of the rectangle is known, and the coordinates of points in the measured plane are obtained. The intrinsic parameters of the camera are calibrated firstly by using the three vanishing points of mutually orthogonal directions. Then the homography between the image plane and the model plane which the rectangle structure belongs to is derived, and the plane measurement method of the latter is described in details. Furthermore, the measurement experiment is done with the single image taken by the industrial camera, and the validity of the proposed method is verified by comparing with real data.

OS1-3 Heavy-duty Spherical Mobile Robot Driven by Five Omnidirectional Wheels

Wei Liu, Junyang Sun, Runjiao Wang, Guanwang Geng, Lian Luo
(Beijing Jiaotong University, China)

This paper presents a heavy-duty spherical mobile robot with 5 omnidirectional wheels. The chassis is supported by four drive omni-directional wheels, and they touch the internal spherical shell. The omni-directional wheel arranged above can prevent the chassis from overturning. Four omnidirectional wheels are driven by two motors so that both forward and steering motions can be achieved simultaneously. The four omni-directional wheels can support heavy load. The mechanism of the robot is described in detail, and its motion and dynamics are built to analysis. Preliminary multi-body dynamics simulation and analysis show that it has good function and dynamic characteristics etc.



OS1-4 Robustness Analysis of Visual SLAM Based on Gazebo Simulation Environment

Jiwu Wang, Yafan Liu (Beijing Jiaotong University, China)

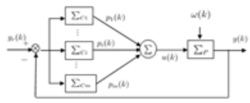
The visual SLAM technology applied to mobile robots is mostly for complex and variable unknown environments. Therefore, in order to reduce the experimental cost and improve the robustness of the technology, it is necessary to test the SLAM technology in the virtual simulation environment before the robot enters the actual working conditions. The Gazebo simulation software used in this paper is a three-dimensional multi-robot dynamics simulation system, which can provide high-fidelity physical simulation environment and a complete set of robot sensor models for SLAM testing. In this paper, we respectively perform simulation tests on advanced technology such as ORB-SLAM2, DSO, and Vins-mono, and obtain experimental data and compare them. Then, based on the test results, we propose corresponding improvement suggestions for these SLAM technology.

OS2 Advanced Control (4)

OS2-1 Multiple Model Adaptive Control Based on Switching/Weighting Intelligent Fusion Algorithm

Weicun Zhang, Yongnan Jia, Qing Li (University of Science and Technology Beijing, China)

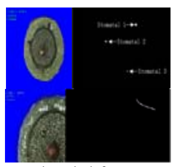
A multiple model adaptive control scheme based on switch/weighting intelligent fusion algorithm is presented with the following contributions: 1) switching/weighting intelligent fusion algorithm against disturbances and noises of the system; 2) stability and convergence analysis of the corresponding closed-loop control system.



OS2-2 Detection Algorithm of Porosity and Crack Defects on Surface of Micro-precision Glass Insulated Terminals

Qunpo Liu[1], Gaowei Wang[1], Fuzhong Wang[1], Manli Wang[1], Hanajima Naohiko[2]
([1]Henan Polytechnic University, P.R.China), ([2]Muroran Institute of Technology, Japan)

A micro-precision glass insulated terminal made of glass powder and metal wire sintered by a special process. Aiming at the defects of the pores and cracks on the surface of the glass-insulated terminal, this paper proposes a method for surface defects inspection based on image processing. Pre-processing operations such as background removal, noise filtering, and G-channel feature extraction are performed on the acquired image with defects. The global threshold segmentation algorithm is used to divide it into some regions. And the feature parameters are established based on the factors such as circularity, aspect ratio, tightness, the length of the contour extract area etc.. The identification of defects is realized based on feature vectors and defect determination algorithms of each region. The test samples are selected including simples with pores, cracks and no defects on the surface. The experimental results show that the algorithm can identify the samples with pores and cracks correctly.

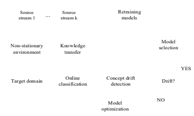


Pore and crack defect test results

OS2-3 Concept Drift Adaptation by Multi-stream Data Knowledge Transfer

Hongpeng Yin, Chenglin Liao, Yi Chai (Chongqing University, China)

The classifier of new concept cannot be retrained after concept drift unless massive new concept instances gathered. To handle this issue, this paper proposed a multi-stream data knowledge transfer approach when just few new concept data are collected. Multi-stream data are represented by labeled source data streams and one unlabeled target data stream. Several sub-classifiers learnt from source data streams individually are composed to an ensemble to predict the target. Empirical studies indicate the effectiveness comparing with other state-of-the-art methods.

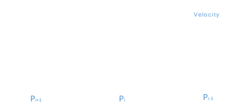


Multi-stream data knowledge transfer approach

OS2-4 Neuro-adapative Control of High-speed Trains under Uncertain Wheel-rail Relationship

Zhechen Wang, Yingmin Jia (Beihang University (BUAA), China)

Traditional automatic controller designing in train systems is almost based on urban rail transit where the influence of changing wheel-rail relationship caused by the variation of speed and environment is ignored. However, high-speed railway operates in more open environment and higher speed, which leading to a more complex variation of wheel-rail relationship occurring. In this paper, we design an automatic train controller in high-speed railway which can realize the automatic velocity tracking even if the uncertain and nonlinear variation of complex wheel-rail relationship happens. First of all, the train dynamic model is established where the wheel-rail relationship is expressed as an uncertain unknown function and the train operation system is expressed as a third-order nonlinear system. Then, a neural network adaptive controller is designed by using the backstepping method and barrier Lyapunov function. Bade on this controller, position and velocity tracking errors are semi-globally uniformly ultimate boundedness. Finally, the effectiveness of the algorithm is verified by simulation experiments.



Multi-point model schematic

OS3 Intelligent Control (4)

OS3-1 No Free Lunch Principle in Agent Swarm Systems: One Case Study

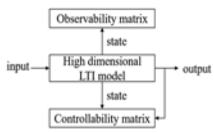
Yunzhong Song[1], Fengzhi Dai[2], Huimin Xiao[3], Shumin Fei[4]
([1]Henan Polytechnic University, China), ([2]Tianjin University of Science and Technology, China), ([3]Henan University of Economics and Law, China), ([4]South East University, China)

This note comes with information flooding of the control action of the leader in leader follower framework. In this situation, even observer for the leader velocity is built in a distributed style, the agent swarm system in a whole could not be classified into distributed one. One case study example will be borrowed to demonstrate the truth of the no free lunch principle in agent swarm systems.

OS3-2 Analyzing the Controllability and Observability of Discrete-Time Delayed LTI Systems with Data-Based Methods

Zhuo Wang (Beihang University, China)

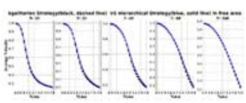
This paper presents a series of novel data-based methods to analyze the state/output controllability and state observability of discrete-time delayed linear time-invariant (LTI) systems. The parameter matrices are assumed to be unknown. In order to analyze the above system characteristics, we first augment the system into a high dimensional LTI model, then employ the measured state/output data to directly construct the state/output controllability and state observability matrices of this high dimensional model, whose ranks play the role of the criteria of corresponding characteristics of the original discrete-time delayed LTI system. These data-based methods have two merits: first, it is not necessary to identify the unknown parameter matrices for characteristics determination, which thus reduces the computational effort greatly; second, the calculation complexity of them is lower because of the relatively simple data-based matrix construction methods.



OS3-3 An Optimal Collective Control Strategy Based on Vicsek Model

Yongnan Jia, Weicun Zhang, Yue Liu, Qing Li (University of Science and Technology, China)

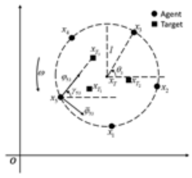
Vicsek model is one of the most famous models taking advantage describing the collective behaviors of self-propelled particles by using simple interaction rules. These particles update their directions according to the average value of their neighbors, who located in its communication range. However, as we all know, individual difference exists widely among social animals. Different individuals may have different influence on the same particle. Therefore, we propose an optimal method considering the individual difference when updating the directions of each particle. According to the simulation results, compared with the standard Vicsek model, the optimal model is more efficient under specific conditions.



OS3-4 Distributed Rotating Encirclement Control of Strict-Feedback Multi-Agent Systems Using Bearing measurements

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

This paper investigates the distributed multi-target rotating encirclement formation problem of strict-feedback multi-agent systems using the targets' bearing angles and the agents' known positions, where all agents are forced to achieve even circular formation around the targets' geometric center. Firstly, an estimator is proposed for each agent to localize the neighbor targets. Secondly, based on the trajectory planning method, a reference trajectory is constructed by three estimators, which are used to obtain the targets' geometric center, the reference rotating radius and angular. Then, the proposed adaptive neural dynamic surface control law forces each agent to move along the reference trajectory. Finally, the performance of our proposed control scheme is verified by a numerical simulation example.



OS4 Intelligent Life and Systems (7)

OS4-1 Development of the IoT Module with AI Function Using STM32 Chip

Jr-Hung Guo, Evgeni Magid*, Kuo-Hsien Hsia, Kuo-Lan Su
(National Yunlin University of Science and Technology, Taiwan)
(*Kazan Federal University, Russia)

The applications of the Internet of Things (IoT) have been widely used in our daily lives with the advancement of related software and hardware technologies. In order to make these IoT modules more intelligent, many IoT modules are incorporated with artificial intelligent algorithms. In this paper, an IoT module with STM32 chip as the main controller is developed. This module uses fuzzy Analytic Hierarchy Process (fuzzy-AHP) and Adaptive Fusion Method to improve the correctness and self-learning ability of the sensor. In terms of communication, the IoT module has Ethernet, Wi-Fi, and LoRa communication interfaces. And we also built a web server on this module, so that the IoT module can operate directly in the browser. Finally, we also developed a monitoring system. Through this monitoring system, multiple IoT modules can be constructed into a sensor network. This monitoring system can also use the same algorithm to correct and isolate data from modules or sensors in the network. These make this IoT module more intelligent and applicable in different areas.

OS4-2 Markerless Indoor/Outdoor Augmented Reality Navigation Device Based on ORB-Visual-Odometry Positioning estimation and ORB-Visual-Mapping Image Registration

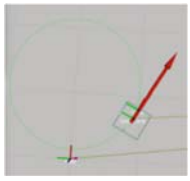
Chian C. Ho, Guan-Jie Wang (National Yunlin University of Science and Technology, Taiwan)

For markerless indoor/outdoor Augmented Reality Navigation (ARN) technology, camera pose is inevitably the fundamental argument of positioning estimation and pose estimation, and floor plane is indispensably the fiducial target of image registration. Based on Oriented FAST and Rotated BRIEF (ORB) feature with descriptors, this paper proposes ORB-visual-odometry positioning estimation and ORB-visual-mapping image registration to improve camera pose estimation and floor plane detection for making ARN more precise and reliable with real-time performance. Experimental results show both ORB-visual-odometry positioning estimation and ORB-visual-mapping image registration have higher accuracy and reliability than conventional well-known camera-pose-based positioning estimation and floor-plane-based image registration methods, respectively, for ARN. On the other hand, markerless indoor/outdoor ARN technology with proposed two methods have seamlessly been implemented on the portable Android platform and have smoothly been verified to co-work well on the portable Android platform.

OS4-3 Apply Adaptive Control Approach for Mobile Robot Path Following

Chun-Chi Lai, Chia-Jen Lin, Kuo-Hsien Hsia, Kuo-Lan Su
(National Yunlin University of Science and Technology, Taiwan)

The state of the art of mobile robot autonomous navigation is composed with global path planning and local path follower. For example, the global path planning majorly applies the shortest path A* algorithm in an occupancy map. And the local path follower such as Dynamic Window Approach (DWA) which computes the velocity commands from the objective function that includes obstacle clearance and progress towards the goal. However, with the obstacle interference, DWA will often lead robot with a great deviation from the global path and non-smooth trajectory. In this work we apply adaptive control method to feed the goals for DWA's computing. The experimental result shows that DWA will follow the global path smoothly.



OS4-4 Malware Classification Using Deep Learning

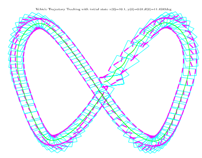
Cheng-Hsiang Lo, Ta-Che Liu, I-Hsien Liu, Jung-Shian Li, Chuan-Gang Liu, Chu-Fen Li
(National Cheng Kung University, Taiwan)

We'll display two different kinds of experiments, which are NIDS-based and Dynamic-based analysis shows how artificial intelligence (AI) helps us detecting and classify malware. On the NIDS-based intrusion detection, we use CICIDS2017 as a research dataset, embedding high dimensional features and find out redundant features in the raw dataset by Random Forest algorithm, reach 99.93% accuracy and 0.3% of the false alert rate. We extract the function calls in malware data by the method proposed in this paper to generate text data. The algorithm n-gram and TF-IDF are used to process text data, converts them into numeric features, and by another feature selection methods, we reduce the training time, achieve 87.08% accuracy, and save 87.97% training time in Dynamic-based analysis.

OS4-5 Robust Control of Nonholonomic Wheeled Mobile Robot with Hybrid Controller Approach

Ho-Nien Shou (National Yunlin University of Science and Technology, Taiwan)

This study proposes a control strategy to solve the nonholonomic mobile robot trajectory tracking problem on the basis of Cerebellar Model Articulation Controller (CMAC). Mobile robot needs two controllers to provide the control demands. One controller is mathematically described in terms of robot's kinematics; while the other is given by dynamics equations. To implement the speed control to track the reference trajectory, we apply the Lyapunov theory to obtain the virtual speed control command. Our simulation is performed in Matlab/Simulink environment, and the results verify the effectiveness of the controller algorithm.



OS4-6 Design and Implementation of Microsatellite Detumbling Control with PWPF: Verification Using PIL

Ho-Nien Shou (Air Force Institute of Technology, Taiwan)

The micro-satellites due to the limits of weight and power, this kind of micro-satellite with low design cost and high precision requires reducing unnecessary attitude sensors and controllers to ensure the precision of attitude control. In this article, an estimating technology about measuring angle speed with a gyro-less is mentioned. This technology is based on the period change in which the earth's magnetic field gets along the track. Only by using a three-axis magnetometer, it can produce the data from the micro-satellite measuring the earth magnetic field. Besides, the measures of the three-axis angle speed and attitude angle can be gotten through Kalman filtering. The purpose of this article is mainly to explore the problem of attitude stumbling control of micro-satellites departing away a carrier to enter a track. It is realized by a thruster to proceed with the confinement of the satellite moving, the attitude stable control and processor-in-the-loop (PIL).

OS4-7 A Study of Applying Computer-assisted Language Learning to English Course for Junior College Students in Taiwan

Shu-Hua Huang, I-Hsien Liu (Air Force Institute of Technology, Taiwan)

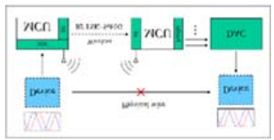
This study aims to explore the impact of the integration of Computer-assisted language learning (Call) into traditional English course on the learning effectiveness at Junior college in Taiwan. To achieve the above purpose, this study adopts the pre-experimental design with one-group pretest-posttest design. This study adopts quantitative and qualitative methods for data analysis and obtains the following conclusion: The application of Computer-assisted English learning has positive and significant progress on students' English learning effectiveness. And students have positive comments on English blended learning, which can enhance students' English learning motivation.

OS5 Artificial Life and Robotics (7)

OS5-1 An FSK based Industrial Analog Signal Transmission

Po-Yun Shih, Chung-Wen Hung, Chau-Chung Song
(National Yunlin University of Science and Technology, Taiwan)

An FSK based industrial analog signal transmission is proposed in this paper. Due to the advantage of digital transmission, such as noise immunity and error check, FSK based digital transmission are popular. Although the analog signal modulation already is well-known and mature, but it is easily interfered by the noise, especially in industrial environment. When the carry frequency is occupied, changing frequency is necessary but tough work for sure. The FSK based wireless transmission is used to perform the wireless transmission of analog signal in the industrial. It also had good effects of saving the wiring between device to device and breaks the restriction of the device movement. Experiment results show that the proposed method could translate wirelessly the analog signal under 16kHz bandwidth.



OS5-2 X-Y Platform Synchronous Control with CANopen

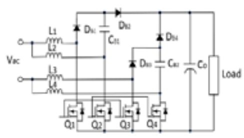
Yu-Ming Guan, Chung-Wen Hung, Shih-Ting Yu, Yu-Kai Chen
(National Yunlin University of Science and Technology, Taiwan)

An X-Y platform consisting of three-phase hybrid stepper motor driver and a controller that supports the CANopen protocol is proposed in this paper. And the communication network of the X-Y platform is based on CAN-bus. In addition, the system is provided with the synchronization signal which meets the CiA301 communication protocol and CiA402 motion control protocol. The interpolation mode which meets CiA402 standard is implemented to improve the precision of position control. Compared with interpolation mode, the other control method, cross-coupling control is also implemented to decrease the synchronization error of biaxial motion control.

OS5-3 A PFC Converter with Voltage Double Characteristic for Universal Input Voltage Applications

Ching-Chun Chuang, Hung-Chi Lee, Chih-Chiang Hua, Chih-Wei Chuang, Chuan-Ming Niu
(National Yunlin University of Science and Technology, Taiwan)

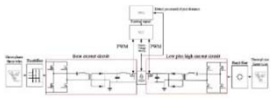
In this paper, a current-fed bridgeless power factor correction rectifier with voltage-double is proposed for a hybrid electric vehicle charging system. The proposed PFC rectifier was simulated on a 3.4-kW prototype. The differences of the simulated THD and PF between the proposed converter and the conventional interleaved PFC converter are insignificant. The proposed rectifier shows an improved low-line efficiency compared to its conventional counterpart under 1.5-kW output power.



OS5-4 Design of a Low-pulse High-current LLC Resonant Converter for EDM Applications

Yu-Kai Chen, Min-Feng Lee, Yung-Chun Wu, Jui-Yang Chiu
(National Formosa University, Taiwan)

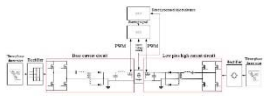
In this paper, we have developed and improve the technology of performance equalization control of Electric discharge machining (EDM) applications. We change the output current and control the current waveform of the EDM. To control the current and the loss of the electrodes and increased the processing speed and the processed object will not be too Rough we have designed a smoothly rising current to let the electrode lossless; to processing the super-hard alloy we have designed a low pulse high output current. These two converters are designed to operate with three-phase three-line ac input voltage and output will be controlled two different dc voltages during the EDM process, the method employed magnitude and frequency control to enable the converter to operate at zero voltage and zero current transitions. During the EDM process will have two different output current first is the base current, the base current is to reduce the electrode the current will be smoothly rising to the base current, when the base current has risen to the stable state, the second LLC will be started output the low pulse high current waveform of the new circuit for the EDM applications.



OS5-5 Networking Integration and Monitoring System with CANopen Controller for Intelligent Production Line of Tool Machine

Chau-Chung Song, Yu-Wei Ho, Yu-Kai Chen (National Formosa University, Taiwan)

In this paper, the development of CANopen node controller and the integration of intelligent production line is focused and studied by applying the CANopen protocol. The CAN bus is used to serve as the communication backbone by using CANopen as the high application-layer protocol. With regard to intelligent production line, the SCADA system is established with real-time monitoring and data acquisition functions to achieve the real-time communication between the operator and the tool machines. The node controller with CANopen is mainly applied to control the movements of each axis motor of measurement modules for tool machine. While the basic network management and data transmission is followed by the dictionary objects in the CANopen CiA301 protocol, the controller needs to implement with the motor motion control specification subjected to the CiA402 protocol. This proposed controller is also additionally equipped with a motor-specific incremental encoder to receive the motor feedback signal, so as to realize the motor speed control and high-precision positioning control. The SCADA system communicates with each node through CAN Bus to complete the task of constructing and managing the entire network. Further, the control commands are released and the effective data transmission and reception processing is performed on all the nodes, so that the operator can immediately monitor the condition of the production line. In the side of data collection, the real-time data is received for data analysis and processes, and displayed on the computer interface for monitoring and facilitating the safety decision-making of the field operator. Finally, the functional testing and system performance of the entire system network monitoring and management will be evaluated and implemented in this paper.



OS5-6 Illumination Manipulation and Specular Reflection Analysis of Still Image with Single Object

Hsuan T. Chang, Chi-Jie Chen (National Yunlin University of Science and Technology, Taiwan)

A lighting manipulation scheme for still images is proposed in this paper. By using the dichromatic-based model, the reflection component can be resolved through a single-image specular reflection removal method based on the characteristics of color constancy. Finally, we can generate a virtual illumination through the recombination of the reflection component. The scenes illumination can be estimated using the proposed automatic method without knowing the illumination spectra, three-dimensional object modeling, or texture databases. Experimental results show that the method is useful for handling the single or multicolor objects in scenes. However, the proposed method suffers some limitations in producing realistic simulation on the dichromatic reflection model. The future work will focus on relaxing the limits so that more lighting effects can be achieved.

OS5-7 Robotics Education for the 2019 MakeX Robotics Competition

Jia-Ming Hsiao (Far East University, Taiwan)

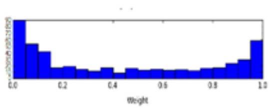
This paper is proposed to illustrate the robotics education experience for 2019 MakeX Robotics Competition in Taiwan. It is promoted in the class to make us of multidisciplinary learning within the fields of science and technology. Students are passionate about innovation by engaging them in exciting Robotics Competition. The spirit of creativity, teamwork, fun and sharing is appreciated in the competition. Alliance competition and cooperation strategy in competition is another topic in teaching and learning.

OS6 Theory and Implementation of Neuromimetic Systems (4)

OS6-1 Real –time pattern recognition implementation on FPGA in multi-SNNs

Xia Yang, Kazuyuki Aihara, Timothée Levi, Takashi Kohno (University of Tokyo, Japan)

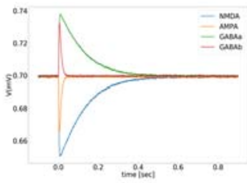
By mimicking or being inspired by the nervous system, Neuromorphic systems are designed to realize robust and power-efficient information processing by highly parallel architecture. Spike timing dependent plasticity(STDP) is a common method for training Spiking Neural Networks (SNNs) for pattern recognition. Here, we present a real-time STDP implementation on FPGA in SNN. Equipped with Ethernet Interface, FPGA allows online configuration as well as data input and output all in real-time. We show that this STDP implementation can achieve pattern recognition task.



OS6-2 Experimental results of a biomimetic silicon synaptic circuit

Ashish Gautam, Takashi Kohno (University of Tokyo, Japan)

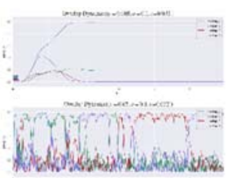
We present the experimental results of a biomimetic silicon synaptic circuit capable of exhibiting both excitatory and inhibitory dynamics, all the synaptic circuits proposed till date have either excitatory or inhibitory dynamics alone. This single circuit is capable of emulating the dynamics of the major synaptic populations received by a neuron. The first order dependence of the synaptic current on the instantaneous value of the membrane potential is also taken into account, with the synaptic current being proportional to the difference between a tunable synaptic reversal potential and membrane potential of the postsynaptic neuron.



OS6-3 Towards Modeling Cholinergic Modulation for Neuromorphic Computing

Naruaki Takano, Takashi Kohno (University of Tokyo, Japan)

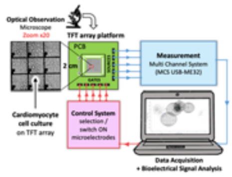
Digital Spiking Silicon Neuron (DSSN) model is a qualitative neuron model specifically designed for digital circuit implementation which exhibits high biological plausibility. In this study we introduced an additional slow negative feedback variable which corresponds to a slow potassium current (3-variable DSSN model) and constructed an autoassociative network. We observed the network dynamics by altering the magnitude of the slow negative feedback current which is known to be controlled by cholinergic modulation, and the strength of neuronal interaction. By altering these parameters, we obtained various pattern retrieval dynamics, such as chaotic transitions within stored patterns or stable and high retrieval performance. We will also briefly discuss potential applications of these results for neuromorphic computing.



OS6-4 Bioelectrical Signal Analysis of Mouse Cardiomyocyte Culture recorded on Thin-Film-Transistor Sensor Arrays

Anne-Claire Eiler, Junichi Sugita, Satoshi Ihida, Hiroshi Toshiyoshi, Katsuhito Fujiu,
Thimothée Lévi, Agnes Tixier-Mita (The University of Tokyo, Japan)

Bio-electricity is at work in our body and the living world around us all the time. A problem occurring in the cardiac cell network can range from minor to fatal inconvenience. However, the mechanism of cardiac signaling network is still poorly understood. In this work, the electrical potentials generated by cardiomyocytes cultured on an array of microelectrodes have been measured using Thin-Film-Transistor (TFT) sensor arrays. The electrical signals have then been analyzed using a Matlab program developed for the bioelectrical processing of electrogenic cells. The recorded signals were filtered for the detection of bioelectric spikes, and grouped into clusters according to their similar features. Our analysis revealed the change of spike amplitudes and durations with modulation of cell culture conditions such as the temperature. Through this analysis, we thus demonstrated the possibility of obtaining accurate spike sorting and analysis from extracellular recordings on TFT array.



OS7 Robotic Manipulation (4)

OS7-1 User Interface and Motion Planner for Task Database

Tokuo Tsuji (Kanazawa University, Japan), Natsuki Yamanobe (AIST, Japan)
Kensuke Harada (Osaka University/AIST, Japan)

In this paper, we present a GUI tools for handing task motion data on robotic manipulation. The data covers robot motion, human motion and object information, We especially present its data structure, data registration, data search, GUI, and API.

OS7-2 POMDP Action Planning for 6 DoF object recognition on Humanoid

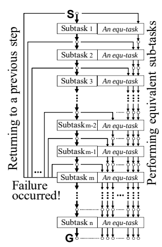
Masato Tsuru (Osaka University, Japan), Tomohiro Motoda (Osaka University, Japan)
Adrien Escande (CNRS-AIST, Japan), Kensuke Harada (Osaka University/AIST, Japan)

In this research, we plan the action of a humanoid robot to find a model given object. For humanoid robot automation, deciding next actions by itself is desired. Because of various noises or occlusions in observation data, robot cannot recognize objects easily. To overcome this problem, our robot walks around and observes target area from various directions. In deciding a next observation position, we use POMDP (Partially Observable Markov Decision Process) and we can obtain a merged point cloud effectively. From that point cloud, we estimate a target object's position and orientation by using 3D model file. Totally, we implemented POMDP, functions of dealing with point cloud data, and controller of whole body of humanoid robot HRP-2, and we tried it on real world, real robot.

OS7-3 Cost-oriented Planning for Error Recovery in an Automation Plant

Akira Nakamura[1], Natsuki Yamanobe[1], Ixchel Ramirez Alpizar[1],
Kensuke Harada[1,2], Yukiyasu Domae[1] ([1]AIST, Japan, [2] Osaka University, Japan)

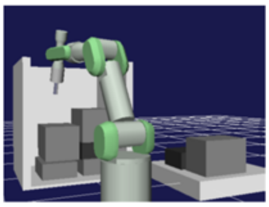
In an automation plant where robots play an active part, not only simple repetition tasks but also complicated tasks are carried out. An error is more likely to occur in such difficult work. That is why the improvement of the technique to perform recovery of an error is necessary. The task often is re-executed after returning to previous step, in the case of a big error. Therefore, it becomes the important problem to decide both the past step that it should return to and the recovery planning after returning. In this paper, cost-oriented planning of error recovery taking account of these two subjects is proposed.



OS7-4 Real-time Planning Robotic Palletizing Tasks using Reusable Roadmaps

Takumi Sakamoto (Osaka University, Japan), Weiwei Wan (Osaka University/AIST)
Kensuke Harada (Osaka University/AIST)

This paper focuses on robotic motion planning for performing the palletizing or de-palletizing tasks. When a robot performs such tasks, a robot usually iterates similar pick-and-place again and again. Taking into account such feature of the palletizing or de-palletizing task, we propose a robotic motion planning approach re-using the previously constructed roadmaps. We propose two methods, R-PRM and R-RRT*, utilizing the previously constructed roadmaps in the PRM and RRT*, respectively. We experimentally confirm that both methods significantly reduce the calculation time compared with the conventional methods on motion planning.



OS8 Advanced Research of Engineering and Management (4)

OS8-1 A Rack and Pinion Driven Mechanical Footstep Power Generator

Ammar A. Al-Talib, K S See (UCSI University, Malaysia)

Kinetic energy from human footsteps during locomotion is generally wasted. It is possible to convert this energy into electrical energy in a non-conventional way. Due to the increase in population, the energy consumption had reached its crisis level as the fossil fuel is depleting drastically over the years. Hence, the waste energy from human footsteps can provide a promising solution to this issue. In this project, some modifications have been made to refine the performance and efficiency of power generation from mechanical footstep generators of previous works. A prototype consists of only rack and pinion mechanism had been fabricated to justify and evaluate the feasibility of the proposed concept. The power is to be generated in this system when force is applied due to the weight of a person stepping on the top plate and will cause a rack and pinion combination to rotate two DC generators. The individuals with weight ranging from 35kg to 75kg have been invited to participate in the experiments by stepping onto the top plate in the first set of experiments and then by jumping in the second set of experiments. The results obtained have been compared with the theoretical results.

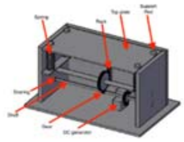
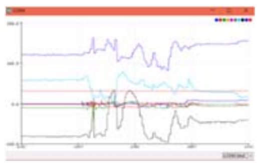


Fig. 1: The Proposed assembly for the mechanical footstep power generator used in the current study.

OS8-2 Table Tennis Using Arduino For Seniors' Healthcare

Reem Ali Abdullah, Mastaneh Mokayef, Miad Mokayef, Sew Sun Tiang, Wei Hong Lim
(UCSI University, Malaysia)

Virtual reality (VR) has been shown to function well as an assistive technology to physical therapy for senior users. Seniors, from retirement home residents, form a unique user group in this field, due to their characteristics and demands. A VR controller was implemented using Arduino, MPU9565 and OpenCV to control a Table Tennis game made on Unity3D. The satisfactory results have been obtained in terms of tracking accuracy.



OS8-3 A Study on Prevention of Predictive Failures using IoT Technology

Tsutomu Ito[1], Hiroshi Sakai[1], Takao Ito[2], Seigo Matsuno[3], Makoto Sakamoto[4]([1]Hino Motors Ltd, Japan, [2]Hiroshima University, Japan, [3]Ube National College of Technology, [4]University of Miyazaki, Japan)

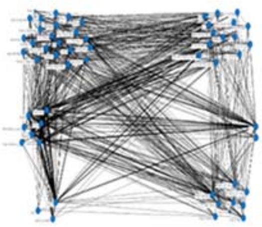
As main power of conveyor and drop lift in manufacturing systems, especially in car manufacturing factory, electric motors play important role in industries. Many electric motors equipped brake system to determine where they should stop. This brake system need to be maintained to keep proper brake gap. Measuring the brake gap of electric motors is one of the main job of maintenance department, but most of the motors located in unsafe spaces, like upper end of the lift or under pit. In this study, we introduced new approach to measure the brake gap of motor to reduce maintenance risk and cost. The method is using vibration acceleration to measure the motor gap in running production system based on mechanical model and IoT technology. Experience results suggested the effectiveness of our research.

Fig.1 Real time measuring of vibration acceleration

OS8-4 Did a mismatch between the ASBJ's standard-setting and its organizational structure occur?

Keisuke Ogata (Osaka City University)

This paper examines the relationship between standard development and organizational structure of the ASBJ, Japanese accounting standard setter. According to the previous literature, the standard setter (re)structured its organization in order to execute strategies and goals that it sets by itself or is expected to achieve. In fact, the ASBJ slowed down the development of revolutionary and improved accounting standards during the period of 2009-2012. However, the ASBJ formed the organization to move ahead on developing such standards, that is, the accounting-profession-centric organization, based on organization analysis in this paper. This paper indicates that an exogeneous shock changed interests in some domestic actors, and consequently caused this mismatch.

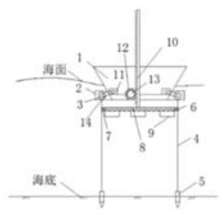


OS9 Recognition and Control (19)

OS9-1 A Design of Ocean Current Velocity Measuring Device

Jichao Zhao[1], Fengzhi Dai[1,3], Xin Ma[2], Fengkun Wang[1],
Haokang Wen[1], Hongbo Hao[1], Qianqian Zhang[1]
([1]Tianjin University of Science and Technology, China, [2]National Oceanic Administration Beihai Marine Environmental Monitoring Center Station, China,
[3]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

Ocean current refers to the relatively stable flow of the surface water of the ocean in a certain direction in a certain direction, and it is the main regulator of the thermal environment on the surface of the earth. At present, the ocean current velocity detection at sea is basically a field measurement by a staff member in a boat, and the adjustment of the detector position during the measurement process is very inconvenient. This paper designs a new submarine current velocity measuring device, including measuring ship, cable, fixed anchor and ocean current velocity detector. The cable on the reel is retracted through the cable retracting room, and the end of the cable is fixedly provided with a fixed anchor. The fixed anchor is fixedly provided with a plurality of ocean current flow rate detectors, and the ocean current flow rate detector has a built-in control chip, a wireless communication module, and Pressure Sensor. This device can flexibly adjust the depth of the detector according to actual needs, so as to measure the flow velocity of different depths.



OS9-2 Device Design Based on TDS Water Quality Detection

Jichao Zhao[1], Fengzhi Dai[1,2], Di Yin[1], Yuhui Cheng[1],
Fengkun Wang[1], Leixin Han[1], Qianqian Zhang[1]
([1]Tianjin University of Science and Technology, China, [2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

The traditional water quality detection method is sampling and detection by artificial nodes, which is easily interfered by weather and environment. The design is a new water quality detection device for marine waters and other waters. The device is spherical in shape and uses a built-in sensor and control system to detect the water quality in the form of drifting. This device has built-in micro control system, TDS detection module, power voltage regulator module, GPS module and wireless signal transmission module. The TDS module detects the concentration of total dissolved matter in the water, and the GPS module measures the current location of the device, which sends the data to the micro-control system. The micro control system sends the water quality information and the geographical location information to the network through the wireless transmitting module, and the user can view the water quality status by the device in real time through the Internet. The power supply voltage regulator module provides the different voltages required by the entire device to achieve the function of detecting the current water quality.

OS9-3 Design of Film Forming Rate Measuring Instrument based on Polyurethane Material

Jichao Zhao[1], Fengzhi Dai [1,2], Fengkun Wang[1], Haokang Wen[1], Hongbo Hao[1], Qianqian Zhang[1]
([1]Tianjin University of Science and Technology, China, [2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

With the development and research of polyurethane materials, it has become widely used in the field of antibacterial, waterproof and moisture permeable, and medical. Therefore, this paper mainly studies on the determination of the film formation rate of polyurethane, and designs a film formation rate measuring instrument based on Atmega328P single chip microcomputer and BH1750 light intensity measurement chip. This design compares the advantages and disadvantages of the currently widely used film formation rate measuring instrument and improves it on the basis of it. By improving its multiple sensors into a high-precision main sensor, measurement requirements are reduced, and the size of the instrument is also reduced, making it easier and more portable. Moreover, it can be used not only to measure the film formation rate of polyurethane, but also to determine the film formation rate of other materials.

OS9-4 Research on the Intelligent Aircraft Design based on STM32

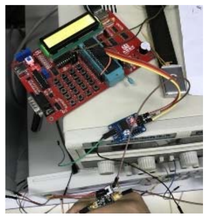
Hongbo Hao [1], Fengzhi Dai [1,2], Haokang Wen[1], Jichao Zhao[1], Qianqian Zhang[1]
([1]Tianjin University of Science and Technology, China, [2] Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

With STM32f767 igt6 (M7 series) as the core, the designed aircraft will finally realize the intelligence of the aircraft by using the optimized attitude algorithm (cascade PID and single-stage PID share the optimized control), the omni-directional ultrasonic radar detection barrier collision prevention technology, the long-distance wireless transmission technology (to realize the timely transmission of images), navigation technology (Beidou and GPS double positioning to realize more accurate positioning), voice recognition technology, man-machine interaction technology, and wireless local area network technology.

OS9-5 Research and Design of Gain Controllable System in RF Receiver

Haokang Wen [1], Fengzhi Dai [1,2], Jichao Zhao[1], Hongbo Hao[1], Hong Niu[1], Qianqian Zhang[1]
([1]Tianjin University of Science and Technology, China, [2] Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

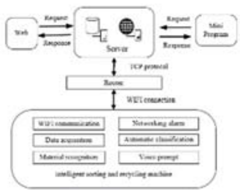
This paper studies and designs a gain controlled radio frequency amplifier system. As the hardware core of the whole system, the variable gain amplifier realizes the data function of the amplifier combined with the single chip microcomputer. After adjusting the gain of RF amplifier, the input signal can be amplified or attenuated, and the ideal output signal can be obtained. Using ADS simulation, the gain controlled RF amplifier system designed in this paper has better control effect, meets higher index, and has stable performance. The experiment shows that the gain controlled RF amplifier system has certain practicability.



OS9-6 Design of Packaging Bottle Recycling System based on Internet of Things

Yujie Yan[1], Fengzhi Dai[1,2], Shiwei Wu[1], Yuanyuan Xi[1], Huanhuan Li[1]
([1]Tianjin University of Science and Technology, China, [2] Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

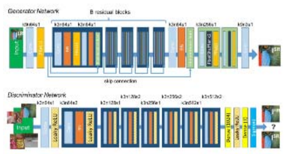
The problems of environmental pollution and resource waste need to be solved. The existing packaging bottle recycling system in the market of China has the problems of high cost and maintenance cost, imperfect function and inflexible rebate mode. Based on the technology of intelligent detection, wireless communication and software development, this paper presents a comprehensive system of packaging bottle classification and recycling, which includes intelligent classification and recycling machine, Mini Program and Web management platform. The system can automatically identify and classify the metal bottles and plastic bottles, and it can also give integrals to the user, which has certain environmental protection significance.



OS9-7 Research on Image Super-Resolution Reconstruction Based on Deep Learning

Lingran An[1], Fengzhi Dai[1,2], Yasheng Yuan[1]
([1]Tianjin University of Science and Technology, China, [2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

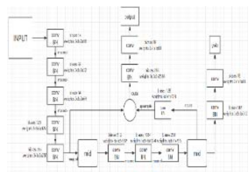
This paper mainly applies the relevant theories of deep learning to image super-resolution reconstruction technology. By comparing four classical network models used for image super-resolution (SR), finally a generative adversarial network (GAN) is selected to implement image super-resolution, which is called SRGAN. SRGAN consists of a generator and a discriminator that uses both perceived loss and counter loss to enhance the realism of the output image in detail. The data sets used by the training network are partly from the network and partly from the artificial. Compared with other network models, the final trained SRGAN network is above average in PSNR and SSIM values. Although it is not optimal, the output high-resolution images are the best in the subjective feelings of human eyes, and the reconstruction effect in the image details is far higher than that of other networks.



OS9-8 Research on Fatigue Detection Method Based on Deep Learning

Yasheng Yuan[1], Fengzhi Dai[1,2], Lingran An[1], Di Yin[1], Yuxuan Zhu[1], Yujie Yan[1]
([1]Tianjin University of Science and Technology, China, [2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

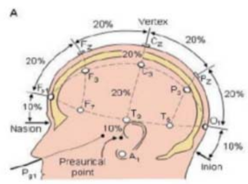
Aiming at the shortage of the existing state of deep learning to test people fatigue, this paper presents a new idea to detect fatigue. First, a video image format is saved to frames, and then the YOLOV3 - Tiny algorithm is used to detect faces in images. Compared with traditional OpenCV image process and other deep learning face recognition methods, YOLOV3-Tiny's advantages are mainly that the network is simple, the computation is small, and it can run on the mobile side or the device side. After face recognition, the recognized face is separate out, and then use OpenCV to process the face, roughly divide the face into three areas: the left eye area, right eye area and mouth area. Finally, YOLOV3 is used to identify the state of each region, and the improved PERCLOS algorithm is used to judge fatigue.



OS9-9 Fatigue Driving Monitoring System based on the EEG

Yuxuan Zhu[1], Fengzhi Dai[1,2], Di Yin[1], Yasheng Yuan[1]
([1]Tianjin University of Science and Technology, China, [2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

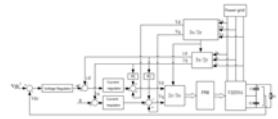
This paper analyzes how to acquire the EEG signal, how to extract and analyze the EEG signal characteristic rhythm and how to estimate the fatigue degree and provide the voice reminder. Firstly, four characteristic rhythms of EEG signals are extracted by wavelet packet decomposition, and the characteristics are analyzed by relative energy frequency spectrum of rhythms. Then, according to the classification of fatigue degree, the energy value of δ wave in F3, F4 and C3 channels is selected as the basis of judging driver fatigue to classify and estimate the fatigue degree of EEG signals. Finally, when the threshold reaches 0.4, the sound card of the computer is called to prompt the tired driver.



OS9-10 Research of the Control Strategy of Vienna Rectifier Circuit based on the Vector Control

Yuxuan Zhu[1], Fengzhi Dai[1,2], Di Yin[1], Yasheng Yuan[1]
([1]Tianjin University of Science and Technology, China, [2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

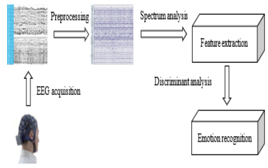
This paper analyzes the working principle of VIENNA rectifier and establishes the transformation model of rectifier in three different mathematical coordinate systems. First is to obtain the control structure of the converter, which is a double closed loop control structure (the voltage is the outer loop control and the current is the inner loop control). By using a feed forward control strategy to solve the problem of phase-to-phase coupling. By adding the voltage equalization loop and using the midpoint balance algorithm to solve the problem of the voltage imbalance caused by the rectifier under load. Finally, uses MATLAB to build the model, simulates and verifies the established model system.



OS9-11 Research on Emotion Classification based on EEG

Di Yin[1], Fengzhi Dai[1,3], Mengqi Yin[2], Jichao Zhao[1]
([1]Tianjin University of Science and Technology, China, [2]Hebei University of Chinese Medicine, China, [3]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

Research shows that human emotion production is closely related to the activity correlation of cerebral cortex, so the research of emotion classification by EEG provides a reliable basis. The feature extraction and classification application of EEG have made rapid development, so we combine EEG with emotion to study emotion classification. However, there are differences between EEG signals of different subjects, which have a certain impact on emotion classification. How to ensure the high accuracy and robustness of recognition is a problem. In view of this problem, the spectrum analysis method is used to extract features to study different subjects in different states. The extracted features are classified into emotion by discriminant analysis algorithm, and the classification effect is satisfactory. There are many methods involved in feature extraction and the space is long, different feature extraction methods will be compared later, so as to improve the robustness and efficiency of emotional classification of EEG signals.



OS9-12 Identification of Synthetic Pigment based on Fluorescence Spectroscopy Combined with RBF Neural Network

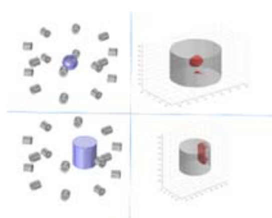
Di Yin[1], Fengzhi Dai[1,2], Yuxuan Zhu[1], Yasheng Yuan[1]
([1]Tianjin University of Science and Technology, China, [2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

Compared with natural pigments, the synthetic pigment is cheap as an important component of food additives, which has good stability, strong coloring ability, etc. But the consumption of excessive synthetic pigment will cause harm to the human body, so an effective method of detecting the pigment is needed. In this paper, 12 samples were selected from 1ug/ml to 100ug/ml different concentrations of carmine and amaranth, the fluorescence spectra of all the samples were measured by the LS-55 fluorescence spectrophotometer. The results showed that the fluorescence peaks were respectively around 450nm and 420nm, redshifts and the fluorescence intensity decreased with the concentration. Two kinds of pigment samples were predicted by RBF neural network model, and the results showed the accuracy of the samples classification was 100% and the error of the sample concentrations was very small, the RBF neural network provides a method for detecting the pigment of foods.

OS9-13 Simulation Study of 3D Reconstruction in Electromagnetic Tomography with Multi-layer Sensors Array

Yuanli Yue, Ze Liu, Yu Miao
(Beijing Jiaotong University, China)

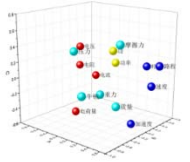
Electromagnetic tomography is an emerging technology of non-destructive testing, which is used in industrial process monitoring and biomedical detection due to its invasive and non-contacting nature. The paper provides the simulation study of 3D reconstruction in electromagnetic tomography with multi-layer sensors array. Various models of the sensor array with different layers and different numbers of coils are established in FEM software——COMSOL Multi-Physics and the changes of sensitivity map correspond to different layers of the sensor array are analyzed, the effects of layer and number of sensor coils to image reconstruction are discussed. The results confirm the feasibility of 3D reconstruction in electromagnetic tomography with Multi-layer Sensors Array and offer a reference to realize it.



OS9-14 Visualized the Knowledge Map in Children's Minds: A Study on Cognitive Structure Measurement

Qiang Wei, Hua Dong, Yi-tong Zhang, Ao-nan Zhang
(Jianghan University, China)

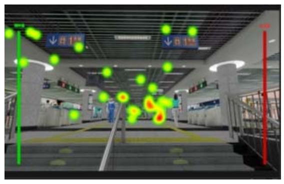
How was knowledge or concepts organized in children's minds? The aim of this study was to explore a measurement based on multidimensional scaling analysis to visualize children's cognitive structure. 120 junior high school students (60 from first grade; 60 from third grade) and 2 teachers participated in this study. These 60 third grade students were divided into three groups based on ranking of physics achievements (20 in A group; 20 in B group; 20 in C group). Fifteen physics concepts selected from high school physics textbook were used in the study. Participants were required to rate relations between two concepts (1 showed close relation; 5 showed far relation), Euclidean distances were computed to operationalize relation-distance between items in target pairs. Compared to first-grade students, third-grade students categorized concepts as organization. Besides, they made the organization tight on the basis of relation-connection among concepts. Compared to A group and B group, C group with lower ranking had more difficulties in categorizing concepts, they spent less time on exercises which would have influences on finding relations among concepts further on making cognitive structure. The results indicate that the more effects students got in learning, the higher level cognitive structure they had.



OS9-15 Escape Route of Subway under Fire Conditions: An Experimental Study in Virtual Reality Environment

Hua Dong, Qiang Wei, Qing-qing Zhang, Lan-lan Fang (Jianghan University, China)

How to finding way for escaping from subway when people were in fire emergency? A virtual Reality (VR) was taken in our study in order to investigate people's stress-level and information searching behavior. In this VR experimental study, participants were required to finding way for escaping from a virtual subway station which was in fire emergency as soon as possible. Compared to male group, the female group had significantly higher stress level report, took longer time on searching information from the environment in order to finding way for escaping. Compared to female group, male group had much more focus on Signage. Our results indicated that female were always ready for searching information for environment and for reception, they even were influenced in finding way for escaping. Fire evacuation system for subway (or other seal off environment) should be design considering the response characteristics on this study's conclusion.



OS9-16 Research on the Smart Home Design based on Single-chip Microcomputer

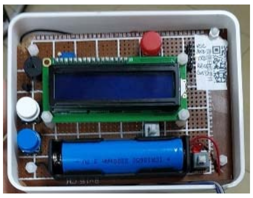
Hongbo Hao[1], Fengzhi Dai[1, 2], Haokang Wen[1], Jichao Zhao[1]
([1]Tianjin University of Science and Technology, China, [2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

The research of the smart home is mainly reflected in the indoor environment temperature, humidity, smoke concentration and the human body to monitor. This indoor environment monitoring system takes the STC89C52 monolithic computer as the control core, using MQ-2 smoke sensor, temperature and humidity sensor, infrared sensor, buzzer and other modules, the main function is to detect the indoor environment temperature and humidity, smoke concentration exceeded the predetermined value, immediately for sound and light alarm; and when the body is detected, immediately to the sound and light alarm. This system can be used in shopping malls anti-theft, Warehouse anti-theft, bank anti-theft and other occasions. The device for the current people's daily life has a strong guiding significance.

OS9-17 Design of a Portable Instrument for Measuring Heart Rate and Blood Oxygen

Haokang Wen[1], Fengzhi Dai[1,2], Jichao Zhao[1], Hongbo Hao[1], Qianqian Zhang[1]
([1]Tianjin University of Science and Technology, China, [2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

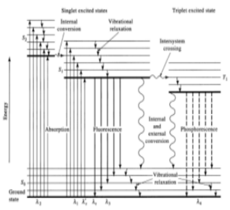
Aiming at the disadvantage of the inconvenience of the current heart rate blood oxygen detection device, a portable heart rate blood oxygen detector is proposed in this paper. The device can detect the heart rate and blood oxygen value of the human body at any time and place according to the needs of the measured person, and can display the heart rate, blood oxygen and temperature on the liquid crystal display. The instrument gets rid of the cumbersomeness of traditional medical appliances, greatly reducing the volume of the instrument and making it easier to carry.



OS9-18 Principal Component Analysis of Wine Based on Three-dim Fluorescence Spectra

Di Yin, Fengzhi Dai, Yuxuan Zhu, Yasheng Yuan
(Tianjin University of Science and Technology, China)

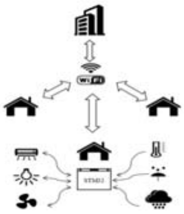
The traditional sensory analysis method can not discriminate the wine objectively and accurately. Compared with the three-dimensional fluorescence technology, it has the advantages of strong selectivity, high resolution and direct access to the fluorescence characteristics of the sample. In this paper, three-dimensional fluorescence spectroscopy Five brands of red wine samples. The five characteristic parameters (mean value, standard deviation, center of gravity coordinates, first order center moment and correlation coefficient) of the three-dimensional fluorescence spectra of five dry red wine samples were extracted respectively. The results showed that the three different dimensions of Cabernet Sauvignon the characteristic parameters of the fluorescence spectrum are similar, based on the analysis of five dry red wine samples by principal component analysis (PCA), it was found that even the same raw material brewing wine due to different origin of raw materials, soil environment, climate and environment will cause the composition of the wine content, which shows a large difference.



OS9-19 Control Design of Intelligent Device for Living Environment of Senile Apartment

YashengYuan[1], Fengzhi Dai[1,2], Shengbiao Chang[3], Lingran An[1], DiYin[1], Yuxuan Zhu[1], Yujie Yan[1]
([1]Tianjin University of Science and Technology, China, [2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China, [3]Dawning Information Industry Co., LTD, China)

At present, with the advancement of science and technology, some smart devices in the society are gradually integrated into the daily lives of the elderly. Now wireless network WIFI technology, real-time network communication function, some new Internet of things technology in the society are more and more accepted by the elderly. This paper researches the existing elderly apartments in the society, and uses the WIFI wireless communication method for information transmission in response to the special physiological needs of the elderly. The bottom layer uses various sensor nodes to transmit indoor environmental data to the upper layer in real time, and can reversely control each indoor electrical equipment through the current environmental index.

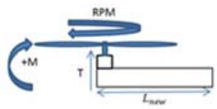


OS10 AI Applications (8)

OS10-1 Effects of Variable Arm Length on UAV Control Systems

M. Rizon, CK. Ang, MI. Solihin (UCSI University, Malaysia),
Zuradzman M. R, H. Desa, Shahriman A. B., Wan Khairunizam(UniMAP, Malaysia)
I. Zunaidi (University of Sunderland, UK)

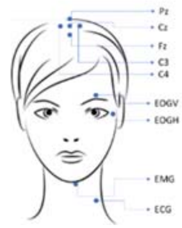
Quadrotor is a type of unmanned aerial vehicle that has been widely used in many applications, such as, policing, surveillance, aerial photography and agriculture. Conventionally, the control of quadrotor flight direction is accomplished by varying speeds of motors or manipulating torques. In this paper, a novel mechanism is proposed. The mechanism uses stepper motors to control the arm length for changing flight directions, while maintaining motors' speed at constant. A mathematical model has been created. The analysis results have shown that varying arm length can effectively control the moment of bending of quadrotors. Increasing the length of arms can result in the increase of the moment of bending without changing speed of motors, thus saving energies. Experimental results have shown that the new mechanism is able to carry more payloads which the motor speed can be utilized fully at 100% while the flight direction is been controlled by changing of the arm length compared to conventional flight control mechanisms.



OS10-2 EEG based drowsiness detection using relative band power and short time fourier transform

Pranesh Krishnan, Sazali Yaacob, Annapoorni Pranesh Krishnan (UniKL, Malaysia)
Mohamed Rizon, Ang Chun Kit (UCSI University, Malaysia)

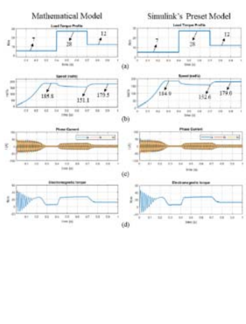
Sleeping on the wheels due to drowsiness is one of the major causes of death tolls all over the world. The objective of this research article is to classify drowsiness with alertness based on the EEG signals using spectral and band power features. A publicly available ULg DROZY database used in this research. The five EEG channels from the raw multimodal signal are extracted and by using a higher order Butterworth low pass filter the high frequency components above 50 Hz are removed. Another bandpass filter bank is designed to separate the raw signals into eight sub bands namely delta, theta, low alpha, high alpha, low beta, mid beta, high beta and gamma. The preprocessed signals are segmented into equal number of frames with a frame duration of 2 seconds using a rectangular time windowing approach with an overlap of 50%. The relative band power based on the short time fourier transform (STFT) was computed for each frame and the features are extracted. The extracted feature sets are further normalized and labelled as drowsy and alert and then combined to form the final dataset. K-fold cross validation method is used. The dataset is trained using KNN and SVM classifiers and the results are compared. The KNN classifier produces 96.1% (dataset1) and 95.5%(dataset2) classification accuracy.



OS10-3 Mathematical Model Implementation of SPWM fed Three-phase Induction Motor Drive Using MATLAB Simulink

Amir Rasyadan, Sazali bin Yaacob, Pranesh Krishnan (UniKL, Malaysia)
Mohamed Rizon, Ang Chun Kit (UCSI University, Malaysia)

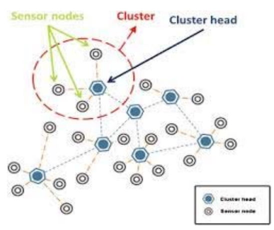
Three-phase induction motors are used in a vast area of applications mainly due to their simplicity, ruggedness and high reliability. With recent advancement in semiconductor technologies, the use of fixed speed induction motor drive is becoming obsolete, majority of the applications now requires inverter-based drives for variable speed operation. In the study of induction motor drive operation, mathematical models are often used to simulate the steady state and transient behavior of induction motor. However, to develop such model is not a straightforward task. Knowing only the equations by themselves are not always enough without some knowledge on solving mathematical equations with the use of computer simulation software. This work presents an approach to implement the mathematical model of a Sinusoidal Pulse Width Modulation (SPWM) fed three-phase induction motor drive in MATLAB Simulink. The sub models include induction motor DQ-model and a voltage source inverter (VSI) fed by SPWM signal generator. The presented model implementation is able to simulate the dynamic behavior of an induction motor operation, this would be useful for further studies on the development of induction motor drive system.



OS10-4 Implementation of X-mean Clustering Algorithm for Wireless Sensor Networks

Abdelrahman Radwan, Nazhatul Hafizah Kamarudin, Mahmud Iwan Solihin, Hungyang Leong, Chun Kit Ang
(UCSI University, Malaysia)

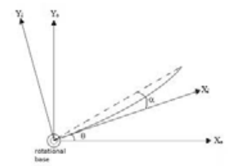
Wireless sensor network (WSN) is a promising technology that has the capabilities to support futuristic applications such as IOT and M2M communication. However, it must overcome crucial constrains such as the limited energy supply and packet routes selection. Clustering algorithm is a potential solution that prolongs the network lifetime when cluster size is balanced, and an optimal number of cluster heads are selected. K-mean algorithm as the one of the popular clustering algorithms preferred over traditional clustering techniques. However, there are some limitation hampering k-means clustering for WSN application such as number of optimal k which is assigned to deal with Cluster Head (CH) selection. In this paper, we propose to apply X-mean algorithm as a new clustering technique for WSN to prolong the network lifetime, i.e. energy consumption optimization. At first, clusters are constructed using tentative CHs and tentative area of centroids in an initial phase. Furthermore, if a cluster meets splitting criteria, new centroids are selected, and new clusters are constructed.



OS10-5 Robust H_∞ controller design for flexible link manipulator based on constrained meta-heuristics optimization algorithms

Mahmud Iwan Solihin, Lim Wei Hong, Chun Kit Ang, Mohamed Rizon,
Abdelrahman Radwan (UCSI university, Malaysia)

The control of flexible manipulators to achieve and maintain accurate positioning is challenging due to the flexible nature of the system. The dynamics is even more complex with parameter variation of the system. Problems arise due to precise positioning requirements and system flexibility which leads to vibration. In this problem, the tracking controller of flexible link manipulators should be able to follow the command of desired angular position and eliminate tip's vibrations while maintaining a fast-tracking response. Robust controller should be designed to deal with the parameter variation of the system. However, parameters tuning of robust controller is a mathematically rigorous by conventional approach. In this paper, H-infinity robust controller tuning using meta-heuristics optimization is proposed. The performance of the design controller will be evaluated and comparison for different meta-heuristics optimization algorithms such as Cuckoo Search (CS), Teaching Learning-based Optimization (TLBO) and Black Hole Optimization Algorithm (BHBO) will be discussed.

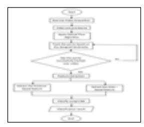


Typical schematic diagram for Flexible Link Manipulator

OS10-6 Classification of Facial Nerve Paralysis Based on Regional Evaluation

Wan Syahirah W Samsudin, Rosdiyana Samad (Universiti Malaysia Pahang, Malaysia)
Kenneth Sundaraj (Universiti Teknikal Malaysia Melaka)
Mohamed Rizon (UCSI University, Malaysia)
Mohd Zaki Ahmad (Hospital Tuanku Ampuan Afzan, Malaysia)

This paper presents an approach of regional evaluation using Kanade - Lucas – Tomasi (KLT) method and extract feature points to classify the normal and patients subjects and also to determine the severity level of paralysis for each region of face. There are Individual Score Table, Total Score Table and Paralysis Score Table were presented. A 100% of accuracy has been obtained in identifying the paralysis with k =3 using k-NN classifier for the Individual Score.



The development of regional evaluation method

OS10-7 Introduction of Forehead Lesion Assessment with House-Brackmann Score for Facial Nerve Paralysis Evaluation

Wan Syahirah W Samsudin, Rosdiyana Samad (Universiti Malaysia Pahang, Malaysia)
Kenneth Sundaraj (Universiti Teknikal Malaysia Melaka)
Mohamed Rizon (UCSI University, Malaysia)
Mohd Zaki Ahmad (Hospital Tuanku Ampuan Afzan, Malaysia)

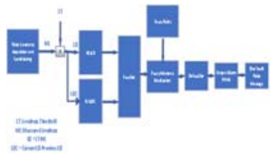
This paper presents a novel approach of forehead lesion assessment using Gabor filter method to identify the type of facial nerve paralysis and proposed a forehead-HB score table to assist the clinicians in assessing the facial nerve function quantitatively, quicker and in a convenient way. A 100% of accuracy has been obtained in identifying the two type of facial lesion, Upper Motor Neuron (UMN) and Lower Motor Neuron (LMN) lesion. Increasing more data may enhance the performance of the accuracy.

The development of forehead lesion assessment

OS10-8 Intelligent Wearable Biofeedback Fuzzy Logic Based Device for Monitoring and Treatment of Voice Loudness

Ali S. AlMejrad (King Saud University, Kingdom of Saudi Arabia)

Development of medical devices are expanding rapidly. This is occurring as results of development of high tech electronics, information technology and wireless communication technology. Integrating these technologies helps to provide good, less cost and more effective healthcare services. In this paper a novel intelligent medical biofeedback device is developed for accurate monitoring of the voice loudness using fuzzy logic compared to proportional control. This accurate monitoring device then enhanced treatment of people unable to control excessive or low loudness due intellectual disability or aggressive behavior causing loudness levels rejected by others. These people with this behavior may often be unaware of their problems and hence an intelligent wearable biofeedback device is necessary to be used by the patients outside the clinic for self-controlling excessive and low loudness without the need to the continuous therapeutic follow-up in speech pathology clinic. The use of the intelligent biofeedback device proved to be an effective due to its enhanced features. The technical design considerations, enhanced features of the device and evaluation will be presented.



Typical illustration for Developed IVLMT Device

OS11 System and Control (16)

OS11-1 Design of Humanoid Soccer Robot Based on STM32

Yuheng Zhang[1], Yulong Peng[2], Yizhun Peng [1], Lianchen Zhao[1], Zhou Zhang[1], Wanlong Peng[1]
([1]Tianjin University of Science and Technology, China)
([2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

Humanoid robot is an important branch in the field of robotics, because its shape and movements are similar to human beings, it provides a good carrier for the related research in the field of artificial intelligence. Humanoid soccer robot of this paper is designed based on STM32 processor. Through theoretical analysis and experimental verification, the stability, working parameters and the division and cooperation of STM32 and 51 series single-chip computers of humanoid soccer robot system are analyzed, and determine the final control scheme.STM32 drives the camera to collect information, process images and make decisions. STC12C5A60S2 controls the robot steering gear to complete the corresponding action. Serial communication is used between the two controllers. The peripheral circuit mainly includes OV7725 camera module and TFT-LCD LCD display module.

OS11-2 A Design and Implementation of Intelligent Cradle

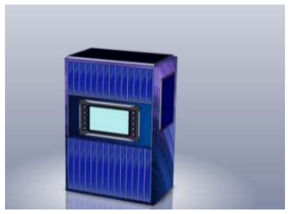
Ting Zhao[1], Qing He[1], Yulong Peng[2], Zhou Yang[1], Zhenjiang Chen[1], Shuo Jiang[1]
([1]Tianjin University of Science and Technology, China)
([2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

The equipment is designed for families to raise children, the device is a cradle of artificial intelligence technology. The data of baby shaking are collected and analyzed by the sensor, so that the cradle bed can bionic cradle shaking. The SVM database training of infant crying can realize the recognition of infant crying and determine the specific meaning of crying, such as hunger, excretion, pain and other factors, and timely inform parents; Based on ROS robot operating system and iflytek platform, it can carry out natural language interaction and autonomous navigation. Parents can call the baby cradle and let it reach the designated position automatically to achieve autonomous obstacle avoidance and path planning without manual interference. At the same time, the camera can view the baby's status and transmit the real-time picture to the mobile phone APP.

OS11-3 A Design and Implementation of Intelligent Networking Bookcase

Zhou Zhang, Yajun Li,Yizhun Peng, Hucheng Wang, Yuqi Zhao
(Tianjin University of Science and Technology, China)

The device is designed for the effective use of book resources. Users can share books in their hands through the device. Multiple devices connected together through wireless network can form a book storage network to realize the flow of paper books. The camera is used for the ISBN code identification of books, the WiFi module is used for the wireless network communication between the device and the background server, the bluetooth module is used for the near distance communication between the device and the mobile terminal, and the infrared module is used for detecting the storage state of the device. Equipment ISBN read, storage of equipment state inspection, such as simple data processing done by the device with a built-in MCU, user registration, user reading habits, such as book search more complex data calculated by the background server processing. By borrowing the concept of distributed computing, implements the equipment, and efficient use of server resources, reduce the workload of the entire system at the same time.



OS11-4 A Design and Implementation of Quad-rotor UAV

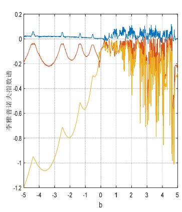
Junjie Lin, Chunxia Zhang, Yizhun Peng, Ting Zhao, Zhengke Xu
(Tianjin University of Science and Technology, China)

In the past 5 years, considerable attention has been paid to unmanned intelligent devices. The Quad-rotor UAV is an unmanned aircraft controlled by radio remote control equipment and self-contained program control device. Our UAV is based on TI MSP432, which can be connected to PC via USB port. This printed circuit board is used to transmit the program for the flight control of UAV. The OpenMV module serves as the data source of the line patrol controller, which is the top-level controller, the same level as the remote controller. The IMU unit calculates information to attitude controller, in order to keep the flight of UAV stable. After the IMU solution information is fused with the optical flow sensor, the information is sent to the horizontal controller to control the flight of the UAV in the horizontal direction. Similarly, After the IMU information is fused with the laser height information, the height controller is applied to control the flight height of UAV.

OS11-5 Characteristic Analysis and Synchronization Control of a Non-equilibrium System

Lianchen Zhao, Xinyu Zhang, Hongyan Jia, Yizhun Peng, Yuheng Zhang
(Tianjin University of Science and Technology, China)

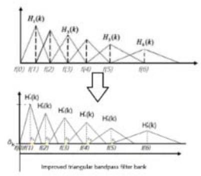
In this paper, the dynamic characteristics of a chaotic system without equilibrium point are studied. Through numerical simulation and theoretical analysis, the chaos characteristics of the system without equilibrium point are studied. In this paper, the nonlinear feedback synchronization control method is used to synchronize the system, and through software simulation, the results are analyzed to determine whether the drive system and the response system are synchronized. At the end of this paper, we give the advantages and disadvantages of this synchronization method.



OS11-6 Classification and Recognition of Baby Cry Signal Feature Extraction Based on Improved MFCC

Zhenjiang Chen, Yizhun Peng, Di Li, Zhou Yang, Nana Wang
(Tianjin University of Science and Technology, China)

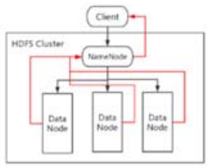
Since MFCC was proposed, it has been widely used in feature extraction of speech signals. However, for some specific sound signals, such as baby crying signal, the direct MFCC feature extraction has a low classification and recognition rate. Through the study of MFCC feature extraction process, it is found that if each filter in the triangle filter bank is shifted up by an $\partial_i(\partial_i \geq 0)$, the recognition rate of the improved MFCC feature extraction is greatly improved.



OS11-7 Design of Space Remote Sensing Data Storage Platform Based on Distributed File System

Di Li, Yizhun Peng, Ruixiang Bai, Zhenjiang Chen, Lianchen Zhao
(Tianjin University of Science and Technology, China)

Due to the large space remote sensing data, a space remote sensing data has seven or eight hundred megabytes or more, and a large amount of space remote sensing data is generated every day for hundreds of GB, TB or even more, so a large amount of space is needed for storage. Space remote sensing data. In order to solve such problems, this paper prepares for the analysis of the subsequent space remote sensing data, and combines the distributed file system to store the space remote sensing data. The installation of the CentOS 6.5 virtual machine through VMware to build an HDFS cluster, through a Namenode node, three Datanode nodes to achieve access to space remote sensing data. Through the upload of the server, the space remote sensing data can be uploaded to the client, and the space remote sensing data can be downloaded through the client.



OS11-8 Research on Semantic Map Establishment of Parking Lot Based on Deep Learning and Multi-sensor

Shiqian Zhang, Yizhun Peng, Ruixiang Bai, Yuheng Zhang, Tianye Jian, Wanlong Peng
(Tianjin University of Science and Technology, China)

With the continuous development of robot technology, SLAM, which is one of the key technologies to realize fully autonomous mobile robots, has become a hot topic in the scientific community. Traditional SLAM technology relies on a single type of sensor to obtain information, and the semantic information of the map is rarely obtained. This paper mainly improves the SLAM mapping technology of the parking lot from the information fusion of sensors and the deep semantics of obtaining map semantic information. Through the lidar and RGB-D camera to obtain multi-modal information, combined with deep learning, extract the semantic information such as the parking lot number and parking position in the image information, help the robot to map and locate faster and better. Finally, simulations were performed using ROS and gazebo to verify the feasibility of the system.

OS11-9 Survey on Kinematics Calibration Technology of Manipulator

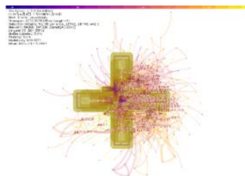
Zhou Yang, Yizhun Peng, Nana Wang, Yuheng Zhang, Tianye Jian
(Tianjin University of Science and Technology, China)

With more and more strict requirements on the quality and accuracy of industrial products, the requirements on the accuracy of robots are increased. Robot calibration is a common method to improve the accuracy. This paper mainly introduces the calibration based on model parameters. Firstly, the calibration of kinematic model parameters is reviewed in four aspects, including kinematic modeling, pose measurement and measurement instruments, parameter identification and error compensation respectively. Secondly, the various modeling methods of kinematics models, the applicable scope of measurement instruments and the algorithms of parameter identification and compensation are compared. Finally, the development trend of calibration is summarized and analyzed.

OS11-10 Visualization Analysis of Web Crawler Evolution Retrieval Research Based on KG

Zhenjiang Chen, Jiamian Wang, Yizhun Peng, Di Li, Lianchen Zhao
(Tianjin University of Science and Technology, China)

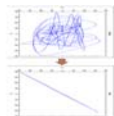
In order to understand the basic situation and future development trend of domestic research on web crawler technology. By using Citespace information visualization analysis software, 2892 web crawler technical literatures in CNKI information technology database from 2000 to 2018 were data mining. From the aspects of literature time distribution, inter-agency cooperation network analysis, co-citation of authors, co-occurrence of keywords and analysis of research frontiers, this paper draws a map of scientific knowledge and sorts out the research background. This paper intuitively reveals the research status, development path, core research groups and research fields of web crawler technology.



OS11-11 Circuit Simulation of Synchronized Novel 4D Chaotic Systems

Yuhan Zhang, Hong Niu (Tianjin University of Science and Technology, China)

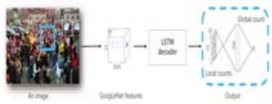
In this paper, synchronization of novel four-dimensional (4D) autonomous chaotic systems, based on the center translation method, is presented. The analog circuit model of the synchronization system is constructed. The numerical and circuitry simulation results are given to illustrate the validity of the synchronization circuitry.



OS11-12 Crowd Counting Network with Self-attention Distillation

Li Wang, Huailin Zhao, Zhen Nie, Yaoyao Li
(Shanghai Institute of Technology, China)

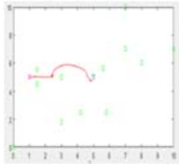
Context information is essential for crowd counting network to estimate crowd numbers, especially in the congested scene accurately. However, shallow layers of common crowd counting networks (i.e., Congested Scene Recognition Network (CSRNet)) don't own large receptive filed so that they can't efficiently utilize context information from the crowd scene. To solve this problem, in this paper, we propose a crowd counting network with self-attention distillation (SADNet). Each input image is firstly sent to the VGG-16 network for feature extracting. Then, the extracted features are processed by the dilated convolutional part for the final crowd density estimation. Specially, we apply self-attention distillation strategy at different locations of the dilated convolutional part to use the global context information from the deeper layers to guide the shallower layers to learn. We compare our method with the other state-of-the-art works on the Shanghai Tech dataset, and the experiment results demonstrate the superiority of our method.



OS11-13 Path Planning Based on Improved Artificial Potential Field Method

Feifan Xu, Huailin Zhao, Zhen Nie, Xin Zhou, Zheheng Tao
(Shanghai Institute of Technology, China)

In this paper, the traditional artificial potential field method is improved. Aiming at the problem that the traditional algorithm cannot pass through the obstacles close to each other and is prone to oscillation near the obstacles, the angle function is added to match the original force field function base on the traditional algorithm, and the stability is enhanced by combining the idea of fuzzy control. Finally, a reasonable and smooth optimal path is obtained by MATLAB simulation. It is proved that the multi-function parallel and multi-algorithm hybrid algorithm is feasible in the field of mobile robot path planning.



OS11-14 Self-balancing Car based on Adaptive Fuzzy PID Control

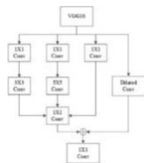
Zhen Nie, Huailin Zhao, Lu Sun, Xiongfeng Zhong (Shanghai Institute of Technology, China)

The self-balancing car is widely studied for its advantages of convenient operation, flexible movement, energy saving and reliability. In this paper, the STM32F105 is used as the main control chip to design a self-balancing car which realizes remote control and video transmission through WIFI. Because the self-balancing car is an unstable nonlinear system, the Kalman filter optimization algorithm is used to fuse the data of gyroscope and accelerometer, and the adaptive fuzzy PID control algorithm is used to control the balance of whole system. The experimental result shows that the system is easy to control, has strong stability, low power consumption and high transmission efficiency.

OS11-15 Crowd Counting Method Based on Improved CSRnet

Huailin Zhao, Shengyang Lu, Li Wang, Yaoyao Li (Shanghai Institute of Technology, China)

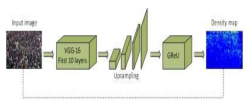
Aiming at the problem of population counting, the research is getting deeper and deeper. CSRnet proposed a method of dilated convolutions instead of convolutional layers and pooling layers. This paper mainly proposes an improved CSRnet crowd counting method, which uses a method similar to the inception-ResNet module to calculate the population density of sparse and dense crowds, and applies this method to ShanghaiTech dataset. The experimental results show that the accuracy of this method has been improved, and the speed of feature extraction has also been improved, compared to CSRnet.



OS11-16 Graph-based Global Reasoning Network for Crowd Counting

Li Wang, Huailin Zhao, Zhen Nie, Yaoyao Li (Shanghai Institute of Technology, China)

Convolutional neural network (CNN) has prompted the crowd counting task to massive progress in recent years. However, filters in the shallow convolutional layer of the CNN only model the local region rather than the global region, which can't capture context information from the crowd scene efficiently. To solve this problem, in this paper, we propose a Graph-based Global Reasoning Network (GGRNet). Each input image is processed by the VGG-16 network for feature extracting, and then the Graph-based Global Reasoning Unit (GGRU) reasons the context information from the extracted feature. Especially, the extracted feature firstly is transformed from the feature space to the interaction space for global context reasoning with the graph convolutional network (GCN). Then, the output of the GCN projects the context information from the interaction space to the feature space. The context information combines with the originally extracted features for final crowd density estimation. We conduct experiments on the UCF-QNRF dataset, and the results demonstrate the effectiveness of our method.



OS12 Advances in Theory and Education on Control (4)

OS12-1 Proposal teaching materials the concepts and principles of machine learning for use in education

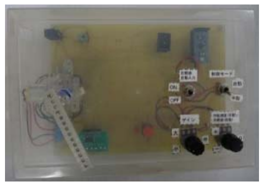
Shinichi Imai, Yusuke Shiba (Tokyo Gakugei University)

In this paper, Technologies such as AI, IoT, and big data are very important contents. However, IoT related content is still in the developing stage and is now evolving at a tremendous speed. It is very important to learn about AI, IoT, big data, etc. Therefore, we propose a teaching tool based on soccer penalty shoots using machine learning. This teaching tool can visually confirm that the goalkeeper robot does not shoot the ball by machine learning.

OS12-2 Objective Evaluation of the Educational Effects on the Feedforward, Feedback and PID Control

Yugo Tokura, Takao Sato, Ryota Yasui, Natsuki Kawaguchi, Nozomu Araki and Yasuo Konishi (University of Hyogo, Japan)

We have newly developed an equipment for the control education. Using the equipment, grade-3 students in Mechanical Engineering course, University of Hyogo, received education on the feedforward, feedback and PID control from the spring to summer in 2019. To objectively evaluate the education effects, the students took tests before and after the control experiment class, and the examination results are compared. The comparison shows that most students precisely understand that the reference tracking performance can be improved by the feedback control and be degraded by the feedforward control.



OS12-3 Programming Learning of Temperature Control for Science Class of Elementary School

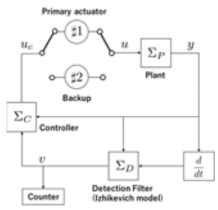
Yoshihiro Ohnishi[1], Takeshi Nakano[2], Teruyuki Tamai[1], Shinnosuke Mori[1], Kazuo Kawada[3]
([1]Ehime University, [2]Ehime University Elementary School, [3]Hiroshima University, Japan)

The programming learning will be made compulsory at elementary schools in 2020. This paper introduces the programming learning at elementary school. In this research, an experimental device on the theme of temperature control was developed in the field of learning electricity in elementary school science. Furthermore, learning activities in which elementary school students perform this programming are also considered. The goal is to learn programming as a technique for efficiently using electrical energy.

OS12-4 Actuator Fault-Tolerant Control using a Spiking Neuron Model

Masanori Takahashi (Tokai University, Japan)

This paper presents a new design method for an actuator fault-tolerant control system (FTCS) using a spiking neuron model. In the proposed FTCS, the Izhikevich neuron model is utilized as a fault detector. When the actuator fails, the neuron model is excited and spikes occurs. Thus, counting up spikes makes it possible to find failures. Compared with the existing AFTCSs, the proposed method has the following advantages: (1) it is possible to set a maximum detection time in advance, and (2) the structure of the control system does not depend on the mathematical model of the plants. It is quite simple even if plants have high orders. In this paper, several numerical simulation results are shown to confirm the effectiveness of the proposed AFTCS.

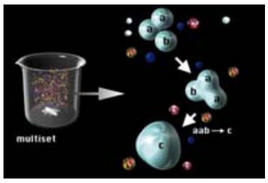


OS13 Natural Computing (5)

OS13-1 Emergence of Adaptive Behavior in Simulations by Using Abstract Rewriting System on Multisets

Yasuhiro Suzuki (Nagoya University, Japan)

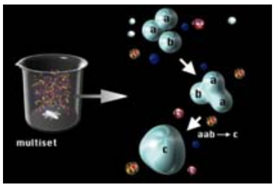
We have developed an Artificial Intelligence system by using a model of chemical reaction, Abstract Rewriting System on Multisets, ARMS, where "intelligence" means that the reaction system can "select" specific molecules to sustain their reactions. We have implemented the reaction system by using an ARMS and have obtained several molecules modified mutated DNA sequences that can sustain the reactions. We confirmed that reaction behaviors in the time series of concentration of non-mutated input molecule and mutated input molecule show oscillations; it would show that the system selects higher concentration one in between non-mutated and mutated one according to its concentration. Since the system exhibits adaptive autonomous behaviors, this DNA reaction networks system realize ARMS.



OS13-2 Implementing the Euler and Runge-Kutta Method by Using Abstract Rewriting System on Multisets

Yasuhiro Suzuki (Nagoya University, Japan)

In this paper, we show that by using a model of chemical reaction, Abstract Rewriting System on Multisets, ARMS, the Euler method, and Runge Kutta method are implemented smoothly. ARMS is a flexible computational model, and it enables us to implement Multi-Agent Systems or P Systems quickly. Hence, by using the proposed method, we can apply the Eular or Runge Kutta method for them. In this paper, we take the Lotka-Volterra model, for example, of the Multi-Agent system and show how to implement it by using the Euler and Runge-Kutta method; then we compare these results.



OS13-3 Extracting Tactile Sensation from Body Movement and Converting it into Vibrotactile Using the Tactile Score Bit

Yasuhiro Suzuki (Nagoya University, Japan)

Body motion, speed and acceleration are extracted using motion capture and converted to tactile score bits and TS bits by discretizing the magnitude. The TS bit is Kansei information, which is information that can be obtained by discretizing the extent and the duration of its size for a specific feature value over time. By using the TS bit, the sensitivity can be extracted from the temporal change in function. In this paper, we will introduce how to convert a dancer's body movement to the TS bit and TS bit to vibrotactile.

OS13-4 A Tactile Sense Centered Virtual Reality Game by Using Biometric Feedback

Yoshihito Ushida, Yasuhiro Suzuki (Nagoya University, Japan)

In this research, we try to construct a Virtual Reality (VR) system centered on the tactile sensation that has used as a secondary. Tactile sensation has a more significant effect on sensibility than audiovisual. Therefore, if tactile sensation and audiovisual sense are combined, presence can be given to a virtual object in the VR space. This system makes a user's biological information accessible in VR space. Then, a virtual object linked to the audiovisual data is presented. A player has to use the virtual object to play this game with physical body movements. And the system feeds back biometric information and deliver the change of biometric data through the sensory presentation to the user.

OS13-5 A Method of Extracting Sensibility from Time Series Data and Converting it to Vibrotactile

Yasuhiro Suzuki (Nagoya University, Japan)

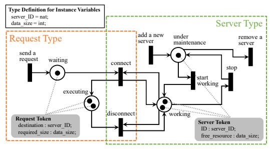
This paper proposes a method to convert sensory information into a vibrotactile sensation. In this method, sensibility is extracted from temporal changes of quantitative data by discretizing the magnitude based on the amount of time change of data and the time change of the data size. Sensibility information obtained by this discretization is called tactile score bits and TS bits. By changing the amplitude and length of the vibrotactile according to the TS bit, we can obtain the vibrotactile transformed from time-series data.

OS14 Software Development Support Method (5)

OS14-1 Behavioral Modeling Technique for Multiple Objects of Software Using Extended Place/Transition Nets with Attributed Tokens

Tomohiko Takagi, Ryo Kurozumi (Kagawa University, Japan)

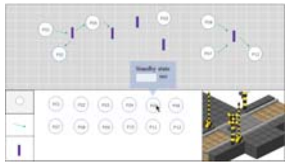
EPN (Extended Place/transition Net) is a formal modeling language to represent the behavior of software that consists of multiple objects. In software modeling using EPN, objects need to be defined individually even if they have the same variables and actions, which causes an increase in model size. This paper shows a novel language called EPNAT (EPN with Attributed Tokens) and modeling technique using it in order to address this problem. In EPNAT, objects are expressed as attributed tokens that are classified into types, and also states and events of objects of the same type are expressed as places and transitions, respectively. Attributed tokens can pass through places and transitions that have the same types as theirs. An EPNAT model can be converted to a VDM++ specification, and allows engineers to check its behavior.



OS14-2 Learning Support Technique of Software Visual Modeling Using Place/Transition Nets

Yuki Ue, Tomohiko Takagi (Kagawa University, Japan)

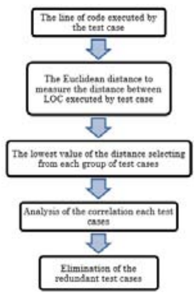
Software modeling is important especially for the development of large and complex software. However, the quality of software models depends on the skill of engineers, and learning it costs time and effort. We propose a learning support technique of software visual modeling to address this problem. The idea of block, which is well-known in the field of visual programming, is introduced into this technique in order that trainees can understand the notation of modeling languages intuitively and can be given guidance about the way to construct proper models. Also, animated graphics are introduced in order that trainees can understand the behavior of their models intuitively. PN (Place/transition Net) is selected as a modeling language in this paper, but other formal modeling languages also can be introduced into this technique.



OS14-3 Redundant Test Cases Elimination on Code Coverage with Distance and Correlation Measurement Method

Mochamad Chandra Saputra[1], Tetsuro Katayama[1], Yoshihiro Kita[2],
Hisaaki Yamaba[1], Kentaro Aburada[1], Naonobu Okazaki[1]
([1]University of Miyazaki, Japan. [2]Tokyo University of Technology, Japan)

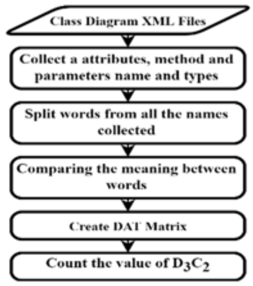
The test cases based on the white box testing will test different control flow paths in a program by executing the input on the test case to the source code and the result is Line of Code (LOC) executed then mainly interested in achieving the possible coverage of the source code. There are several test cases that have similar LOC executed. The distance and correlation measured by comparing the LOC executed by each test case using Euclidean distance. The test case that has the lowest value of distance means high redundancy and possible to execute similar LOC or path. The research tries to eliminate redundant test cases based on that similarity. Several redundant test cases eliminated to get the best test cases. By Euclidean distance, the research finds the redundant test cases on the test suite.



OS14-4 The Measurement of Class Cohesion using Semantic Approach

Bayu Priyambadha[1], Tetsuro Katayama[1], Yoshihiro Kita[2],
Hisaaki Yamaba[1], Kentaro Aburada[1], Naonobu Okazaki[1]
([1]University of Miyazaki, Japan, [2]Tokyo University of Technology, Japan)

The cohesion is one of the design quality indicators in software engineering. The measurement of the value of cohesion is done by looking at the correlation between attributes and methods that are in a class. In Direct Distance Design Class Cohesion (D_3C_2) metrics, attributes and methods are assumed has a good correlation if they have a similar type. But, the similarity of type parameters and attributes do not always indicate that these attributes are managed (correlated) in the method. This study attempts to gain information that can enhance the degree of certainty of a correlation between the methods and attributes. Relatedness between them has seen from closeness the meaning of the attributes, methods, and parameters name. The experimental results have declared an increase in the value of cohesion produced in line with the similarity of meaning.



OS14-5 Proposal of an Algorithm to Generate VDM++ by Using Words Extracted from the Natural Language Specification

Yasuhiro Shigyo[1], Tetsuro Katayama[1], Yoshihiro Kita[2],
Hisaaki Yamaba[1], Kentaro Aburada[1], Naonobu Okazaki[1]
([1]University of Miyazaki, Japan), ([2]Tokyo University of Technology, Japan)

The natural language includes ambiguous expressions. VDM is one of methodology on the formal methods to write the specification without ambiguity. Because VDM++ is written by strict grammar which contains data types and invariants, it is difficult to write a VDM++ specification. This research attempts to generate automatically a VDM++ specification from natural language specification by machine learning. To generate a VDM++ specification, it is necessary to extract words such as function and variable from natural language specification. This paper proposes an algorithm to generate VDM++ specification from the extracted words. An identifier is generated from the extracted words, and then the VDM++ specification is generated by converting the identifiers into VDM++ grammar.

OS15 Recognition and System (11)

OS15-1 Efficient Detection Device for Wafer Physical Defects

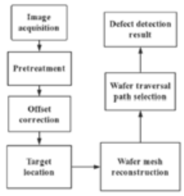
Jianyong Chen, Xiaoyan Chen, Chundong Zhao
(Tianjin University of Science and Technology, China)

Wafer defect detection is an important part of semiconductor manufacturing. In order to improve the efficiency of semiconductor wafer defect detection, this paper designs an efficient visual inspection device. The device uses programmable logic controller (PLC) as controller for the transmission mechanism and uses servo motor as drive device. A CMOS camera is used to capture wafer images, a computer is used for wafer image processing, and results are displayed on a graphical user interface. Camera calibration is implemented by integrating the mapping relationship between the pixel coordinate system and the world coordinate system, the internal and external parameters of the camera and the distortion coefficient. The device proposed in this paper is low in cost and the detection process is stable and reliable. It provides a new solution for wafer defect detection.

OS15-2 Wafer Defect Detection Method based on Machine Vision

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

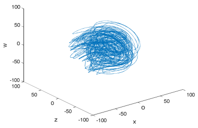
With the development of integrated electronic circuit manufacturing technology, enterprises have put forward higher requirements for the quality of silicon chips. Aiming at the low efficiency of silicon wafer defect detection, this paper proposes an automatic defect detection method based on machine vision. The voiding algorithm based on flood fill can effectively extract the inner contour information of the wafer profile. A rotation correction algorithm is proposed to correct the wafer yaw angle. The actual wafer was used to verify the performance of the proposed method. The results show that the proposed method is effective in detection accuracy.



OS15-3 Analysis and Hardware Implementation of a Novel 4D Chaotic System

Yan Sun, Yongchao Zhang, Jiaqi Chen
(Tianjin University of Science and Technology, China)

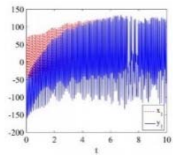
In this paper, a novel 4D chaotic system is proposed. First, the basic dynamic characteristics of this system are analyzed theoretically. Second, dynamical properties of it are investigated by phase trajectory, Poincaré section map, Lyapunov exponent spectrum and bifurcation diagram. Finally, a analog circuit of the system is designed and simulated by Multisim. The experimental results show that the circuit simulation results are consistent with the numerical simulation results.



OS15-4 Research on Synchronous Control of a Novel 4D Dissipative Chaotic System

Jiaqi Chen, Yongchao Zhang, Yan Sun
(University of Science and Technology, China)

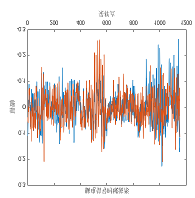
A novel four-dimensional dissipative chaotic system is presented in this paper, which with six nonlinear terms, three variable system parameters and one external excitation input. Through theoretical analysis and numerical simulation analysis of the chaotic characteristics of the system, a self-synchronization of the chaotic system is realized by the method of nonlinear synchronous control strategy, numerical results show that nonlinear synchronous controller is correct and effective.



OS15-5 EEG classification based on common spatial pattern and LDA

Lei Wang[1], Zixuan Li[2]
([1]Tianjin University of Science and Technology, China),
([2]Dongbei University of Finance ＆ Economics, China)

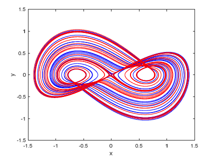
In the BCI system of motor imaging (MI), the effective method of MI feature extraction and classification is the key to the follow-up work. According to the physiological phenomena of event related desynchronization (ERD) and event related synchronization (ERS) of MI EEG signal, the optimal eigenvector representing the state of MI EEG signal corresponding to the maximum eigenvalue was selected, and then the feature extraction method of common spatial pattern (CSP) was proposed, and the classification of MI EEG signals was realized by combining with linear discriminant analysis (LDA). The experiment verified the method of data Ⅲ provided by Graz University, and the recognition accuracy of left and right hand motion EEG signals is 80%. The results of experiment and data analysis show that LDA classifier can be applied to classify the feature vectors extracted from CSP, which has a good recognition accuracy and can lay a theoretical and experimental foundation for the practical application of BCI system based on MI.



OS15-6 Dynamic Characteristics Analysis of the Shimizu–Morioka Chaotic System

Wenxin Shi, Hongyan Jia
(Tianjin University of Science and Technology, China)

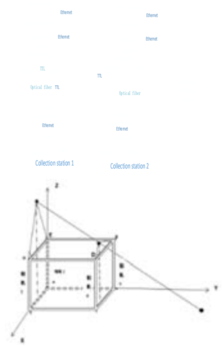
In the paper, we investigate the Shimizu–Morioka chaotic system based on Numerical simulations and FPGA implementation. Firstly, the stability of three equilibrium points of the system is analyzed by using the Routh-Hurwitz criterion. Then the coexistence characteristics of the system are studied by using digital analysis methods such as the phase portrait diagram, Lyapunov exponents diagram and bifurcation diagram. Finally, the digital circuit of the system is realized by FPGA, and the feasibility of the system is illustrated.



OS15-7 Research on the Motion Track of High-speed Objects

Qianqian Zhang[1,2], Fengzhi Dai[1,3], Jichao Zhao[1], Haokang Wen[1], Hongbo Hao[1]
([1]Tianjin University of Science and Technology, China, [2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China, [3]Advanced Structural Integrity International Joint Research Centre, Tianjin University of Science and Technology, China)

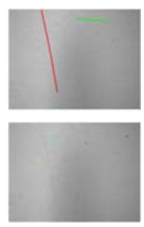
Measuring the location coordinates of high-speed moving objects is of great significance in the strike accuracy test. Currently, there are a variety of measurement methods, such as light screen target measurement and acoustic target measurement. The integrated control and verification workstation includes the calculation system and wireless transmission module, which is mainly responsible for the final modeling and analysis of the data collected by the system to obtain the calculation results and display them. "Acquisition station 1" and "acquisition station 2" are composed of acquisition computer, battery, DCDC module (secondary power supply), camera, trigger module and other parts. The main function of acquisition station is to analyze and process the data collected by the camera and upload them to the integrated control algorithm workstation. The purpose of this design is to develop a set of high speed object trajectory landing point prediction system, and realize the recording, processing, wireless uploading, human-computer interaction and other functions of high speed linear motion target video collected by 4 sets of high speed cameras set up at two points.



OS15-8 Research on Surface Defect Detection of Aluminum based on Image Processing

Xuemin Liu[1], Ce Bian[3], Di Yin[2], Yuxuan Zhu[2], Yasheng Yuan[2]
([1]China Petroleum Engineering and Construction Corporation, China,
[2] Tianjin University of Science and Technology, China,
[3]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China)

Aluminum material is relatively smooth. Aluminum surface engender scratches and bruises easily when collide with other metal. Surface defect detection of aluminum products is particularly important. It is very convenient to use machine vision method for defect detection. Defect contour extraction is an important part of machine vision for defect detection. The surface of aluminum metal is very reflective and shallow scratches are easily mistaken for defects. There are many kinds of filtering, such as the mean filtering, gauss filtering, median filtering and directed filtering. With the help of filtering, dynamic threshold can achieve a good effect. The severe scratch defect and the slight scratch can be clearly separated from the surface of the aluminum product.



OS15-9 Research on the Control of Multi Position Production Line based on PLC

Yong Hou[1], Hao Wang[1], Runhua Mao[2], Xuemin Liu[3]
([1]Tianjin University of Science and Technology, China, [2]Tianjin Tianke Intelligent and Manufacture Technology CO, LTD, China, [3]China Petroleum Engineering and Construction Corporation, China)

In view of the disadvantages and limitations of the traditional manual production line with the cylinder processing system as the main body, this paper proposes an intelligent production line control system based on PLC. The rearrangement of production line and automatic planning of production process are realized. The application of the system in practical production shows that the designed system can achieve 8 hours of unmanned automatic operation. It greatly improves production efficiency and product quality, reduces labor intensity and production cost, and makes the traditional production line intelligent and automatic.

OS15-10 Research on acoustic source localization system based on acoustic holography

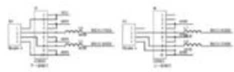
Xiuqing Wang, Xiaoyun Jia (Tianjin University of Science and Technology, China)

In order to solve the leakage problem caused by the rupture of oil and gas pipelines, a sound source localization algorithm based on acoustic holography is proposed. According to statistically optimal cylindrical near-field acoustic holography (SOCNAH), establishing the simulation analysis of simple sound source and multiple sound source. In this paper, the influence of noise on multi-frequency and multi-source sound signal is also studied, the window function is used to reduce the noise to acquire the best sound source signal. In addition, the near-field acoustic holography experiment of pipe cylinder model was established. The all-weather acoustic emission acquisition system was used to acquire and analyze acoustic emission signals, and the acquired acoustic emission signals were imported into the near-field acoustic holography positioning software to obtain the acoustic source positioning results. Simulation and experimental results show that the sound source localization method based on SOCNAH can realize the sound source localization of pipeline, determine the damaged state of pipeline, and carry out targeted remedy to ensure the safety of pipeline facilities.

OS15-11 Anti-interference Method of Electrical Fast Transient for Fire Alarm of Substation

Guo Wangyong, Ju Zhenfu, Chen Guang, Zhu Bo
(Beijing NARI Yihe Environmental Protection Technology Co., Ltd, China)

Aiming at the problem of electrical fast transient (EFT) in substation, the mechanism of EFT generation and electromagnetic compatibility is analyzed. In this paper, the anti-interference method of fire alarm is designed for complex EFT in substation. The test results show that the design method can suppress the influence of EFT on fire alarm in substation effectively, improve the immunity level of EFT of fire alarm in substation, and meet the requirements of level 4 in GB / t17626.4-2006 standard.

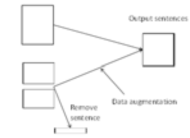


OS16 Media Information Processing and Artificial Intelligence (4)

OS16-1 Effectiveness of Data Augmentation in Pointer-Generator Model

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

We propose a new data augmentation method in automatic summarization system, especially Pointer-Generator model. A large corpus is required to create an automatic summarization system using deep learning. However, in the field of natural language processing, especially in the field of automatic summarization, there are not many data sets that are sufficient to train automatic summarization system. Therefore, we propose a new method of data augmentation. We use Pointer-Generator model. First, we determine the importance of each sentence in an article using topic model. In order to extend the data, we remove the least important sentence from an input article and use it as a new article. We examine the effectiveness of our proposed data augmentation method in automatic summarization system.



OS16-2 Mouse Cursor Control System Using Facial Movements

Masayoshi Tabuse[1], Manase Mizobe[2], Yasunari Yoshitomi[1], Taro Asada[1]
([1]Kyoto Prefectural University, Japan, [2]TORAY ENGINEERING Co.,Ltd, 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 facial movement. Recognition of facial movement of a person makes it possible to operate a computer. Furthermore without the adjustment for a user and adjustment for the distance from a user, it is possible to reduce the burden on a user. We developed a system to resolve these problems. In our system, a web camera and dlib C++ library are used to obtain the face direction and extract feature points of the face. Changing the face direction, we can move a mouse cursor. Recognizing an open mouth or closed eye, we can carry out an operation of mouse click. In this paper, we evaluated the effect on operability due to the face direction and recognition rate due to distance.

OS16-3 Facial Expression Synthesis Using Vowel Recognition for Synthesized Speech

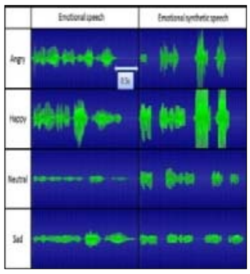
Taro Asada[1], Ruka Adachi[2], Syuhei Takada[3], Yasunari Yoshitomi[1], Masayoshi Tabuse[1]
([1]Kyoto Prefectural University, Japan, [2]Software Service, Inc., Japan, [3]Seika Town Hall, Japan)

We have developed a system for facial expression synthesis of an agent in making a speech using vowel recognition for synthesized speech generated for the agent. The speech is recognized using a speech recognition system called as Julius, followed by facial expression synthesis of the agent using preset parameters depending on the vowel each. For making the agent, we used MikuMikuDanceAgent (MMDAgent), which is a freeware animation program that allows users to create and animate movies, to create an agent. To produce the agent's voice, we used the speech synthesis function setting built into MMDAgent. The impression evaluation obtained from questionnaire survey indicates that our proposed system is more natural than an agent that using preset parameters manually decided for each speech.

OS16-4 Speech Synthesis of Emotions in a Sentence Using Vowel Features

Rintaro Makino[1], Yasunari Yoshitomi[2], Taro Asada[2], Masayoshi Tabuse[2]
([1]SoftBank Corp., Japan, [2]Kyoto Prefectural University, Japan)

Recently, methods for adding emotion to synthetic speech have received considerable attention in the field of speech synthesis research. We previously proposed a method for speech synthesis of emotions using vowel features of a speaker. In the previous study, as an initial investigation, we adopted the utterance of a Japanese name that is semantically neutral. In the present study, by using the proposed method, emotional synthetic speech in a sentence made from the emotional speech of one male subject was discriminable with a mean accuracy of 78.9% when 13 subjects listened to the emotional synthetic utterances of "angry," "happy," "neutral," or "sad" when the utterance was the Japanese sentence 'このぬいぐるみかわいくない' (in Japanese), which means, 'This stuffed toy is pretty, isn't it?"



OS17 Machine Learning Technologies for Human Understanding (2)

OS17-1 Customization of Contents for Acquisition of Skills of FPS without Trainer

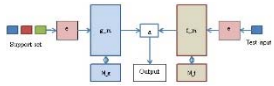
Masao Kubo, Takeshi Ueno, Hiroshi Sato (National Defense Academy, Japan)

In this paper, we conducted a study to develop self-study materials with the ability to provide more suitable materials for students' skills. A first person shooter (FPS) type simulator-based teaching material is created to acquire knowledge to escape from a building surrounded by a group of zombies. Conventional games resume from a predetermined scene regardless of trainee's skill when defeated by an enemy, but are often forced to play with this, which contributes to no increases in knowledge and skill. In this paper, we propose a re-spawning point suitable for each trainee by a recommendation algorithm which tries to find good game scenes by trial and error. A simple experiment is performed to confirm the effectiveness of the proposed content generation algorithm.

OS17-2 Visual Classification of Malware by Few-shot Learning

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

The threat of malware to modern computer systems has been growing. Millions of new malware are said to be created every day all over the world. Therefore, the same malware is rarely found in different organizations. In this situation, signature-based malware detection can not work effectively. This study focuses on a few-shot learning, which can learn the pattern from very few samples. We apply Memory Augmented Matching Network (MANN) to malware classification. We show that the MANN can classify unknown malware from its binary image. We also show that the combination of MANN and prototypical network increases the accuracy of classification.



OS18 Service Robotics (7)

OS18-1 Modelling Autonomous Parallel Parking Procedure for Car-like Robot Avrora Unior in Gazebo Simulator

Dinir Imameev[1], Ksenia Shabalina[1], Artur Sagitov[1], Kuo-Lan Su[2], Evgeni Magid[1]
([1]Kazan Federal University, Russia).
([2]National Yunlin University of Science and Technology, Taiwan)

This paper focuses on implementation of path planning and control for Avrora Unior robot car that enable autonomous parallel parking. Path planning is based on existing geometrical approach, which was modified to fit specific kinematics of the robot shape and control geometry. Geometry and parking space size determine path key points: steering and counter-steering points. We implemented and tested the algorithm using Avrora Unior robot model in Gazebo simulator.

OS18-2 Traffic Sign Recognition Algorithm for Car-like Robot Avrora Unior

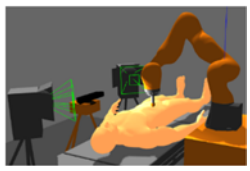
Nikita Nikiforov[1], Ksenia Shabalina[1], Artur Sagitov[1], Kuo-Hsien Hsia[2], Evgeni Magid[1]
([1]Kazan Federal University, Russia)
([2]National Yunlin University of Science and Technology, Taiwan)

Achieving high accuracy of traffic signs detection and recognition is difficult in real-time and is heavily influenced by non-ideal environment conditions. In this paper, we propose to combine a set of Haar cascades that had been trained on a large number of samples and could recognize different types of road signs in different positions and orientations. We use feature detection and feature matching in the process of traffic sign type identification. Our algorithm was validated on Avrora Unior robot model in a simulated environment within Gazebo.

OS18-3 Robotic palpation modeling for KUKA LBR IIWA using Gazebo Simulator

Artur Shafikov[1], Artur Sagitov[1], Hongbing Li[2], Natalia Schiefermeier-Mach[3], Evgeni Magid[1]
([1]Kazan Federal University, Russia), ([2]Shanghai Jiao Tong University, China)
([3]Health University of Applied Sciences, Austria)

Palpation is one of diagnostic methods being extensively used in medical practice. It is often used for tumor detection in cancer screening but its efficacy is highly dependent on examining physician's skill. Therefore, using a robotic tool could make this procedure more objective. In this paper, we present our control and perception modules of autonomous palpation robotic system. We have modelled KUKA LBR IIWA manipulator control using MoveIt motion planning in Robot Operating System and validated it in Gazebo simulator.



OS18-4 Remote Control Application for "Servosila Engineer" Robot on Android Mobile Devices

Daniel Kiryanov, Roman Lavrenov (Kazan Federal University, Russia)

Even though modern mobile robots' autonomous navigation capabilities rapidly increase, teleoperation mode is still an important tool, especially in critical domains like rescue or military robotics. This paper presents Android OS based teleoperator control tool for Russian crawler robot Servosila Engineer. We changed the way of data exchange between a robot and its operator, which allows using Wi-Fi data standards in order to simply data transfer from OCU process to vehicle process. Our application provides robot remote control and video data transfer from robot onboard cameras.

OS18-5 Network Failure Detection and Autonomous Return for PMB-2 mobile robot

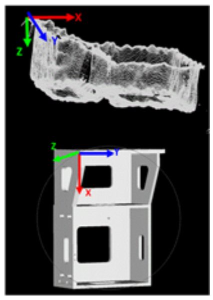
Dmitry Bereznikov, Aufar Zakiev (Kazan Federal University, Russia)

In real world teleoperated tasks a robot connection with its operator is not always stable, so it is important to increase the robot autonomy. This paper focuses on increasing robot autonomy through autonomous return and charging station docking in a case of connection loss. We integrated the algorithm into real robot control system or PAL Robotics PMB-2 robot and experimentally demonstrated its good efficiency. The algorithm analyzes network failure through incoming TCP/IP packets, uses Simultaneous Localization and Mapping (SLAM) and path planning algorithms for autonomous return, and dock station plugin for the robot docking and recharging, which continues until the connection to teleoperator station is restored.

OS18-6 Pick and Place of Large Object Based on 3D vision

Hsien-Huang P. Wu, Jia-Kun Xie
(National Yunlin University of Science & Technology, Taiwan)

Automation is a necessary tool to achieve unmanned factory, and machine vision plays a vital role for providing intelligent recognition in automation. In this study, technique of 3D camera is used for 3D image capturing and matching to identify, pick and place large objects automatically. The system uses a commercially available 3D stereo vision camera to build the image acquisition system, and reconstruct the large objects in 3D. This 3D image is used for classifying the objects with the 3D object recognition algorithm. After the object was identified and its 3d information was obtained, a robot arm integrated with the camera system can be used for grasping. Compared with traditional 2d image matching for 3d big object recognition, stable 2d image features are much harder to obtain due to the shadow. The 3D stereo vision camera does not require strict requirements for lighting control, and only needs stable ambient light. The reduction of the difficulty in building image acquisition environment for large 3d object and the cost of camera system provide a new option for applications that requires large object identification.



OS18-7 Application of the MyRIO Based Mobile Robot Using Vision System

Bo-Jun Yang[1], Kuo-Hsien Hsia[1], Kuo-Lan Su[1], Evgeni Magid[2]
([1]National Yunlin University of Science & Technology, Taiwan)
([2] Kazan Federal University, Russia)

The paper develops a MyRIO based mobile Robot with the vision system. The mobile robot contains a robot arm and a gripper based on the subject of the world skill competition project. The structure of the mobile platform uses the Matrix elements. The mobile platform integrates some sensors, four DC servomotors, three RC servomotors, a MyRIO based controller, and a vision system. The core of the MyRIO based controller is the NI-Single-Board RIO 9606 module. The vision system of the mobile Robot can search and recognize the assigned shape and color billiard ball to be fixed on the front side using Otsu algorithm. In the experimental results, the mobile robot moves to the assigned location from the start location, and uses the vision system to search the assigned billiard ball. Then the mobile robot uses the robot arm to catch the assigned billiard ball, and moves to the assigned position autonomously. The robot arm puts down the billiard ball. Finally, the mobile platform moves to the start location and stop.

OS19 Advanced Information Processing Applications (5)

OS19-1 Extraction of Irrelevant Sentences from Online Hotel Reviews

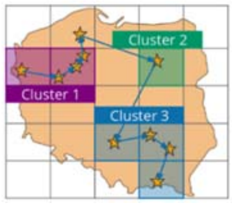
Shogo Watanabe, Masaharu Hirota, Tetsuya Oda (Okayama University of Science, Japan)

Many reviews of hotels have been posted on review sites such as TripAdvisor and Yelp. Many tourists select a hotel to reserve based on their ratings and reviews. Although containing useful information, those reviews may also contain useless information, which reduces their readability. Removing irrelevant sentences from those reviews can improve their readability. This paper proposes a method to extract irrelevant sentences from a review. Our approach uses a supervised learning method to classify the sentences into relevant and irrelevant. We demonstrate the performance of our proposed method by evaluation experiment using TripAdvisor dataset.

OS19-2 Inferring Home Location of Foreign Tourists Based on Travel Routes Extracted from Social Media Sites

Lugasi Chen, Masaharu Hirota (Okayama University of Science, Japan)

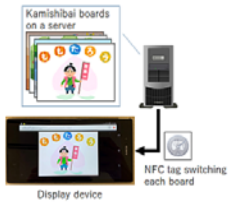
Tourists of certain regions have a tendency to visit certain places when travelling abroad. The availability of large amount of data from social media sites allows researchers to profile those tendencies, which could be useful for many applications. We propose a method to infer the home location of a tourist based on such tendency, using metadata annotated to contents from social media sites. Our approach uses Long Short-Term Memory neural network to learn the quantized travel route of each user, which is based on clusters extracted from the obtained data. We demonstrate the performance of the proposed method by evaluation experiments using photographs obtained from Flickr.



OS19-3 The IoT Solution to Archive and Play the Digital Library of Kamishibai

Motohide Yoshimura, Ayumi Eikawa (The University of Nagasaki, Japan)

We innovate an IoT solution to archive and play the digital library of Kamishibai. Kamishibai is a form of Japanese picture story show. It consists of a sets of illustrated boards placed in a miniature stage-like device and the narrator tells the story by switching each boards. Our solution has two aspects. The one is an educational tool for children and the other is an electric toy playing with children. As to the former, the children collect local mythology and tradition and make kamishibai boards by drawing software. As to the latter, children construct a miniature stage-like device by applying NFC tags for switching each boards. In this paper, we report findings through the construction of a kamishibai device which displays the kamishibai boards on a content server by using NFC reader and tags.



OS19-4 A Research on Prediction of Inter-firm Relationships

Takao Ito[1], Tsutomu Ito[2], Matsuno Seigo[3], Rajiv Mehta[4], Makoto Sakamoto[5]
([1]Hiroshima University, Japan, [2]Hino Motors Ltd, Japan, [3]Ube National College of Technology, Japan, [4]New Jersey Institute of Technology, USA, [5]University of Miyazaki, Japan)

Inter-firm relationship is one of the basic issues for corporate strategy formation. A plethora of researches of inter-firm measurement have been published, but studies on prediction of inter-firm relationships have been sparse. In order to discover a new approach of corporate strategy, the authors collected transaction and cross shareholding data sets from Yokokai, parts suppliers of Mazda, and developed a new model to forecast inter-firm relationship using graph theory. The main contribution of this research is that prediction of inter-firm relationship depends not only future trends, but also interactive firms' relationship.

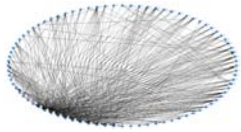


Fig. 1. Interactive transactional inter-firm relationship among Yokokai in2008.

OS19-5 Generation of Arbitrarily-Oriented Ripple Images Using Circular-Sector-Type Smoothing Filter and Inverse Filter

Toru Hiraoka (University of Nagasaki, Japan)

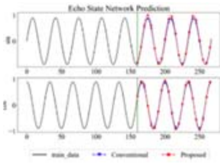
A non-photorealistic rendering method for generating ripple images from photographic Images has been proposed using intensity gradient. Ripple patterns imitate a wave on the water surface and are composed of continuous lines with fluctuations. Ripple images are expressed by superimposing ripple patterns on photographic images. However, the conventional method can only generate vertical and horizontal ripple patterns. Therefore, in this paper, we develop a method that can generate ripple patterns in any orientation. Ripple images generated by the proposed method are called arbitrarily-oriented ripple (AOR) images. The proposed method is executed by an iterative process using circular-sector-type smoothing filter and inverse filter. To verify the effectiveness of our method, we investigate the changes in AOR images by changing the values of the parameters, then we visually evaluate the appearance of these images.

OS20 Artificial Intelligence for Embedded Systems and Robotics (5)

OS20-1 A Hardware-Oriented Echo State Network for FPGA Implementation

Kentaro Honda, Hakaru Tamukoh (Kyushu Institute of Technology, Japan)

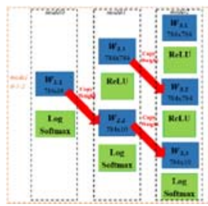
Recurrent neural network (RNN) is commonly used for applications such as voice recognition and stock prediction. This paper designs echo state network (ESN), a kind of RNN for field programmable gate arrays implementation. The proposed network is able to compute faster compare to CPU, and the circuit's resources are reduced by using fixed-point operation, quantization of weights which including accumulate operations and making the dataflow modules more efficient. The circuit is verified by the prediction of sine and cosine wave experiments, and through the results, promising performance is shown.



OS20-2 Network with Sub-Networks

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

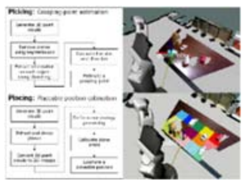
We introduce *network with sub-networks*, a neural network which it's weight layers can be detached into sub-neural networks during inference phase. To develop trainable parameters which can be inserted into both base- and sub-models, firstly, the parameters of sub-models are duplicated to base-model. Each model is forward-propagated separately. All models are grouped into pairs. Gradients from selected pairs of networks are averaged and updated both networks. With MNIST dataset, our base-model achieves the identical test-accuracy to the regularly trained models. In other hand, the sub-models are suffered an extend of loss in test-accuracy, nevertheless the sub-models provide alternative approaches to be deployed with less parameters compare to the regular model.



OS20-3 A Study on Fast Pick-and-Place Method for Home Service Robots using 3D point clouds

Tomohiro Ono, Hakaru Tamukoh (Kyushu Institute of Technology, Japan)

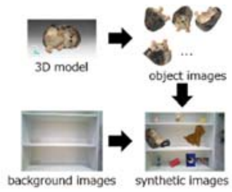
Home service robots have begun attracting attention due to decreasing birthrate and increasing aging population. The basic functions of home service robots are object recognition, picking and placing, recognition of people and environment, and interaction with peoples. In this paper, we focus on object picking and placing in the domestic environment. The pick-and-place task is a very important technique used for arrangement shelves and tidying up rooms. In order for the robot to operate smoothly, the movement to pick-and-place the object must be fast. Therefore, we develop a fast pick-and-place method using 3D point clouds. Regarding the picking, we describe a grasping-point estimation method. Regarding the placing, we describe a placeable position estimation method. These methods use at RoboCup@Home, an international competition aimed at the practical application of home service robots, and their effectiveness and validity are verified.



OS20-4 Acceleration of training dataset generation by 3D scanning of objects and Evaluation of recognition accuracy

Yushi Abe, Yutaro Ishida, Tomohiro Ono, Hakaru Tamukoh
(Kyushu Institute of Technology, Japan)

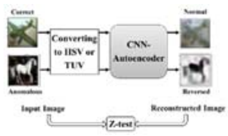
Object recognition is an important technology for Home Service Robot. When using object recognition in the home, a data set of the object matched to the environment is required. Generally, it takes a lot of time to create a data set, and there are concerns about human error in annotation work. As a previous method, there is an automatic annotation system, but it has been a problem that a lot of manual processing is required for dataset generation. In this paper, we use 3D scanning of an object to eliminate manual processing when creating a data set and further speed up. As a result, the proposed method can create datasets about 40 minutes faster than the previous method, and manual processing is not necessary. But, training YOLOv2 by using the dataset, and comparing the mean Average Precision, the accuracy was only dropped 4 points.



OS20-5 Anomaly Detection Using Autoencoder Trained with Reversed Color Models

Obada Al aama, Hakaru Tamukoh (Kyushu Institute of Technology, Japan)

Autoencoders (AEs) have been applied in several applications such as anomaly detectors and object recognition systems. However, although the recent neural networks have relatively high accuracy, but sometimes false detection may occur. This paper introduces AE as an anomaly detector. The proposed AE is trained using both normal and anomalous data based on convolutional neural network (CNN) with three different color models HSV, RGB and TUV (own model). As a result, the trained AE reconstruct the normal images without change, whereas the anomalous image would be reconstructed reversely. The training and testing of the autoencoder in case of RGB, HSV, and TUV color models were demonstrated and Cifar-10 dataset had been used for the evaluation process. It can be noticed that HSV color model has been more effective and achievable as an anomaly detector rather than other color models based on Z-test and F-test analyses.



OS21 Mathematical Informatics (4)

OS21-1 Hidden Surface Removal for Interactions between User's Bare Hands and Virtual Objects in Augmented Reality

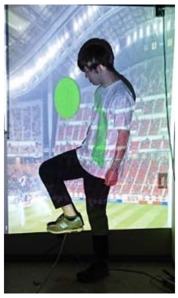
Takahiro Ishizu, Makoto Sakamoto, Kenji Sakoma, Takahiro Shinoda, Amane Takei
(University of Miyazaki, Japan), Takao Ito (Hiroshima University, Japan)

Augmented reality (AR) technology is a technique of superimposing information generated by a computer on perceptual information that we receive from real space.
Recently, much attention has been focused on interaction techniques between users and virtual objects, such as the user directly manipulating virtual objects with his/her bare hands. On the other hand, in AR technology, since the 3-dimensional (3D) model is superimposed on the image of the real space afterwards, it is always displayed on the front side than the hand. Thus, it becomes an unnatural scene in some cases (occlusion problem). In this study, this system considers the object-context relations between the user's hand and the virtual object by acquiring depth information of the user's finger. In the evaluation experiment, it is confirmed that the hidden surface removal in this study not only makes it possible to consider the object-context relations but also can distinguish between finger boundaries and to clarify and process finger contours.

OS21-2 Proposal of Interactive Projection Mapping using Human Detection by Machine Learning

Takahiro Shinoda, Makoto Sakamoto, Takahiro Ishizu, Kenji Sakoma, Amane Takei
(University of Miyazaki, Japan), Takao Ito (Hiroshima University, Japan)

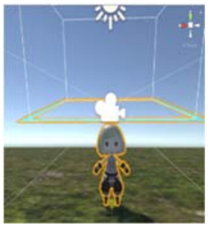
In recent years, “Entertainment Computing” (EC) has attracted attention and has become one of the major industries in Japan. “Projection mapping” is well known in this EC. Projection mapping is a video technology that creates a new space by synthesizing space and video using a projector. Among them, many people are fascinated by works that create a fantastic world by combining dancer performance and projection mapping. However, these works require the performer to accurately align with the coordinates of the image objects in the projection mapping, which is not easy for everyone. In this study, we aim to entertain not only the people who see the projection mapping but also performers. Therefore, we have prototyped an interactive projection mapping that changes according to user movement. This time, we focused on sports and projected the ball to the user to experience baseball pitching and soccer lifting. Furthermore, we conducted a questionnaire survey to evaluate the sense of use of this system, and the results showed that many people can enjoy by this projection mapping.



OS21-3 Fundamental Study on Control of CG Characters by Electroencephalography (EEG) Analysis

Kenji Sakoma, Makoto Sakamoto, Takahiro Ishizu, Takahiro Shinoda, Amane Takei
(University of Miyazaki, Japan), Takao Ito (Hiroshima University, Japan)

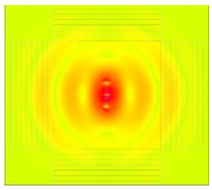
Virtual Reality (VR) technology is expected to develop in various fields such as medical, education, business and entertainment. In this study, we aim at more intuitive operation by focusing on troublesome mounting in VR. When using VR equipment, it is necessary to set up many cables and sensors. Also, when using it, there must always be a certain space around the user. This is because we always use the controller when operating in the world of VR, for reasons of care or for the safety of the user. For this reason, there arises a problem that the location must be selected when the VR device is used. As a method to solve this problem, we propose manipulation method of CG character by electroencephalography (EEG). In this paper, we ask five subjects to manipulate CG characters by EEG and evaluate this system by questionnaire.



OS21-4 Development of parallel microwave analysis code: ADVENTURE_Fullwave

Amane Takei (University of Miyazaki, Japan)

In this presentation, a parallel microwave analysis code based on an iterative domain decomposi-tion 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 decompo-sition method is introduced in subdomains. The simplified Berenger’s PML is in-troduced which these eight corners are given the average value of all PML’s layers.



OS22 Robot Competitions for Social Contribution (5)

OS22-1 Real-Time Self Localization for Autonomous Robot of RoboCup MSL

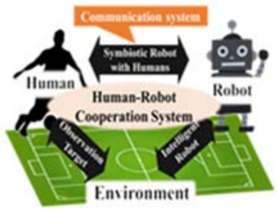
Kaori Watanabe, Yuehang Ma, Tetsuya Yoshida, Hidekazu Suzuki
(Tokyo Polytechnic University, Japan)

The main focus of the RoboCup competitions is the game of soccer, where the research goals concern cooperative multi-robot and multi-agent systems in dynamic environments. In the field of RoboCup, self-localization technique is important to estimate own position including goal and other robot positions and to decide strategy. This paper presents a self-localization technique using an omni-directional camera for an autonomous soccer robot. Then, we propose the self-localization method which generates the searching space based on a model-based matching with white line information of soccer field, and which recognizes the robot position by optimizing the fitness function using Genetic Algorithm. Moreover, we perform verification experiment of self-localization and verify the accuracy of the proposed method.

OS22-2 Behavior Selection System for Soccer Robot Using Neural Network

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

With the progress of technology, the realization of a symbiotic society with human beings and robots sharing the same environment has become an important subject. An example of this kind of systems is soccer game. Soccer is a multi-agent game that requires strategies by taking into account each member's position and actions. The behavior of a player changes depending on the game situation, such as the score gap, the remaining time and the importance of the match. Players may play offensively when their team is losing, or be defensive when their team is winning with a minimum score difference. In this paper, we discuss the results of the development of a learning system that uses SOM to select behaviors depending on the situation.



OS22-3 Development of the Auto Measurement System for Cedars in a Forest Using a Drone

Keiji Kamei, Masahiro Kaneoka, Ken Yanai, Masaya Umemoto,
Hiroki Yamaguchi, Kazuki Osawa (Nishinippon Institute of Technology, Japan)

Drones have been used in many purposes for a long time. Especially, development of the automatic observation systems such as measurement using drones for the primary sector of industry have been frequently researched in recent. The measurement of a tree growth in a forest is also one of the aim for a drone application. In this study, our aim is to develop the automatic measurement system for size of a tree in a forest. The difficulties are that a drone has to recognize trees, to create a map of a forest and to measure the size of trees from a front camera. To overcome those difficulties, we propose that a drone recognizes trees based on RCNN, creates a map from SLAM and measures a tree by SFM. Experimental results from the drone competition show that a drone has been able to recognize a tree and to fly safety.

OS22-4 Report on the 5th Tomato-harvesting Robot Competition

Yasunori Takemura[1], Takayuki Matsuo[2], Takashi Sonoda[1], Kazuo Ishii[3]
([1]Nishinippon Institute of Technology, Japan, [2]National Institute of Technology, Kitakyushu College, Japan, [3]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. With an aim to promote the automation of tomato harvesting, we have organized the tomato harvesting robot competition since 2014. In this paper, we report on the results of 5th tomato harvesting robot competition in 2018.

OS22-5 Reports on the 7th Underwater Robot Festival in Kitakyushu

Yuya Nishida[1], Takashi Sonoda[2], Takayuki Matsuo[3], Shinsuke Yasukawa[1], Masanori Sato[4], Yasunori Takemura[2], Kazuo Ishii[1]
([1]Kyushu Institute of Technology, Japan, [2]Nishinippon Institute of Technology, Japan, [3]National Institute of Technology, Kitakyushu College, Japan, [4]Nagasaki Institute of Applied Science, Japan)

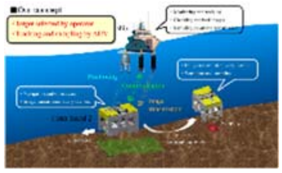
For enhancement of oceanic engineering technology and researchers, underwater robot competition has been held since 2016. Seventh competition in this year consists AUV league that university's vehicles automatically cruise at field and junior league that underwater craft is made. Six teams from nationwide university jointed in the AUV league in the competition and challenged the autonomous navigation in the field using developed vehicle. 14 teams from nearby high school jointed the junior league in the competition, learned to make underwater craft for two days, and try the competition using handmade craft. The paper reports competition regulation of AUV and junior league and results of the competition held in October 2019.

OS23 Advances in Field Robotics and Their Applications (6)

OS23-1 Sea Trials for Benthos Sampling Using Autonomous Underwater Vehicle

Yuya Nishida[1], Shinsuke Yasukawa[1], Takashi Sonoda[2], Keisuke Watanabe[3], Kazuo Ishii[1]
([1]Kyushu Institute of Technology, Japan, [2]Nishinippon Institute of Technology, Japan, [3]Tokai University, Japan)

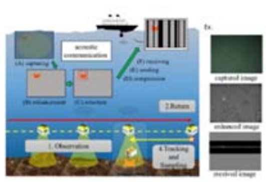
Autonomous underwater vehicles (AUVs), free from umbilical cable, can cruise a wide area and show good performances for scientific bio-resource surveys. In the surveys, AUVs take seafloor images for estimations of benthos biomass and distribution using mounted camera systems. The next requirement of the bio-resource survey using AUVs is to take back the samples for scientists. Although previous AUVs would capture the specific benthos which are decided as the target beforehand, scientists on the site can't chose the sampling targets during a mission. To realize an efficient bio-resources survey by AUVs, we had developed an sampling method that human in the loop and command the sampling targets. This paper explains our sampling method by the AUV Tuna-Sand2 and shows results of sea trials.



OS23-2 Field Experiments of Underwater Image Transmission for AUV

Shinsuke Yasuakwa[1], Yuya Nishida[1], Jonghyun Ahn[2], Takashi Sonoda[3], Kentaro Yanagise[1], Keisuke Watanabe[4], Kazuo Ishii[1]
([1]Kyushu Institute of Technology, Japan, [2]Hiroshima Institute of Technology, Japan, [3]Nishinippon Institute of Technology, Japan, [4]Tokai University, Japan)

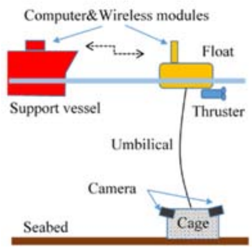
Autonomous Underwater Vehicle (AUV) has enough advantage in the task of observing a wide range area, however, it is difficult to achieve a task that requires high adaptability and recognition such as biological sampling. In order to improve the efficiency of the survey, it is necessary for AUV to report the state of the seafloor to the operators on board reasonably and in real time. In order to realize such new seafloor observation, the seafloor image transmission using the underwater acoustic communication device is the one of the solutions. We have been developing seafloor image selection and image compression technology for acoustic transmission. In this paper, we report the results of underwater image selection and transmission in biological sampling experiments conducted in November 2019 off the coast of Suruga-bay, Shizuoka, Japan. The image selection algorithm selected an image of a benthos, while an image of a marine snow was also selected. We also report the results of acoustic communications.



OS23-3 Development of Subsea Creature Monitoring Station for AUV Exploration Assistance

Keisuke Watanabe[1], Koshi Utsunomiya[1], Amir Sadiq[1], Daichi Hiramaki[1], Kyoko Takashima[1], Kazuo Ishii[2] ([1]Tokai University, Japan, [2]Kyushu Institute of Technology, Japan)

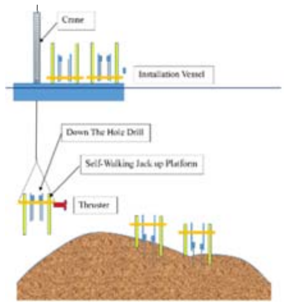
We are developing a set of sea creature investigation systems which consists of AUVs, a support vessel and subsea monitoring stations. In this paper, we introduce our recent result on developing a low cost monitoring station. This station is intended to monitor underwater images of subsea creatures wirelessly from a boat in advance before an exploration AUV is put into water. As AUV images cannot be checked until it returns, we need information where we should put an AUV. The station consists of a cage on the seabed, a float which relays underwater images to a support vessel and an umbilical cable from the cage to the float. We describe the concept, the design and system integration of the station in the paper. We also present an experimental result that was carried out to test the implemented functions at Suruga bay at 100m water depth.



OS23-4 Consideration on Installation Method of In-situ Drilling Platform through Simulations (withdraw)

Keisuke Watanabe (Tokai University, Japan)

Drastic cost reduction for subsea drilling is desired for scientific research and subsea mining field exploration. The traditional drilling method needs a special drilling platform which equips massive drilling facilities with many trained crews. Low cost in-situ drilling platform is one of the possible alternatives and we are studying about a self-walking jack up platform with a down the hole drill. In this paper, I focus on consideration on installation method to reduce its installation cost. The platform is suspended from a vessel and its motions in the horizontal plane are controlled by thrusters attached to the platform. The dynamics of a suspended platform with thrusters is formulated and a simulation program to estimate position accuracy is developed. Through simulations the effectiveness of the method is confirmed.



OS23-5 Environment Map Generation in Forest Using Field Robot

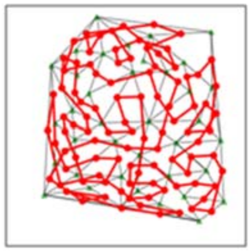
Noboru Takegami, Eiji Hayashi, Ryusuke Fujisawa (Kyushu Institute of Technology, Japan)

We are developing an autonomous field robot to save labor in forest operation. About half of Japan's artificial forest area is already available as wood. However, trees are not harvested and forest resources are not effectively used, because the labor and costs are not sufficient. The employment rate of young people in forestry tends to decline, and the unmanaged forest area is expected to increase in the future. Therefore, in our laboratory we propose an autonomous field robot with all terrain vehicles that focuses on the automation of work. The robot automates weeding and observation in the forest. In this research, we introduced Robot Operating System (ROS) to this robot. In addition, we observed trees by generating an environmental map in the forest using Simultaneous Localization and Mapping (SLAM).

OS23-6 Graph-Based Path Generation for Area Coverage

Ayumu Tominaga, Eiji Hayashi, Ryusuke Fujisawa (Kyushu Institute of Technology, Japan)

This research addresses the trajectory generation in 2D Euclidean space for navigation of an autonomous land vehicle (ALV). The main aim of the approach presented is to give the trajectory in order to cover the given work space to the ALV for maximize effect of the ALV. In this work, we propose graph-based offline trajectory generation method. Here, a Hamiltonian path would be find using 2D coordinates of environmental landmarks as nodes in the graph. The Hamiltonian path contains nodes that will express a midpoint between a pair of landmarks, it could be treated as global trajectory. We applied the method to an actual artificial forest with treating cultivated trees as nodes in the graph.



OS24 Robot Intelligence and Factory Automation (6)

OS24-1 Deep Object 6-Dof Pose Estimation Using Semantic Pixel-wise Segmentation

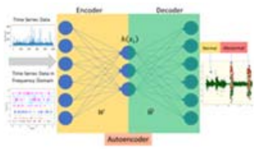
Victor Pujolle, Eiji Hayashi (Kyushu Institute of Technology, Japan)

Pose estimation algorithms’goal is to find the position and the orientation of an object in space, given only an image. This task is complex, especially in an uncontrolled environment with several parameters that can vary, like the object texture, the background or the lightning conditions. Most algorithms performing pose estimation use deep learning methods. But it is hard to create dataset to train such kind of models. In this paper I developed a new algorithm robust to a high variability of conditions using semantic segmentation of the image and trainable on a virtual dataset. This system should perform keypoints based pose estimation without considering background, lighting or texture changes on the object.

OS24-2 Autoencoder with Spiking in Frequency Domain for Anomaly Detection of Uncertainty Event

Umaporn Yokkampon, Sakmongkon Chumkamon, Eiji Hayashi
(Kyushu Institute of Technology, Japan)

Since the research of big data and the internet of things are broadly and facilitate human life which could let the human know their information to predict or plan their activities. However, in uncertainty real-world data, there are the anomaly cases sometimes occur in all data because of the obtaining data process or obtaining devices. In this paper, we propose the autoencoder method with spiking raw data to the frequency domain to analyze and predict the anomaly case among the standard data set. Moreover, we evaluate and performance of the autoencoder method with various activation functions and loss functions. In this paper, we propose the new idea to utilize the Root Mean Square Standard Deviation for the loss function to improve the reconstruction process of the method. Finally, we evaluate and discuss the characteristic of Autoencoder performance.



OS24-3 Evaluation of the Relationships Between Saliency Maps and Keypoints

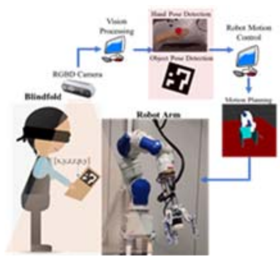
Ryuugo Mochizuki, Kazuo Ishii (Kyushu Institute of Technology, Japan)

Attention is a selective action of focusing on particular objects, which is essential for environmental recognition and categorized into bottom-up attention driven by locally outstanding image region and top-down attention like searching. Itti et. al. proposed the saliency map based on intensity, color and edge orientation. However, the saliency depends on spatial frequency because fixed size Gaussian filters is applied for Itti's saliency map. Generally, scale-invariant keypoints in image feature extraction are generated by similar manner with saliency map and often applied for image matching and localization problems. One of desirable property of keypoints is invariant to photographed conditions such as lighting, orientation, distance, etc. In this paper, the relationships between saliency maps and keypoints are evaluated.

OS24-4 Robot Motion and Grasping for Blindfold Handover

Jiraphan Inthiam, Sackmongkon Chumkamon, Umaporn Yokkampon, Eiji Hayashi
(Kyushu Institute of Technology, Japan)

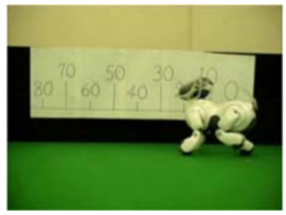
Autonomous robots in human-robot interaction (HRI) recently are becoming part of human life as the number of service or personal robots increasingly used in our home. In order to fulfill the gap of HRI merit, we would like to propose a system of the autonomous robot motion and grasping creation for assisting the disabled person such as the blind people for handover tasks to help blind people in pick and place tasks. In this paper, we develop the robot motion to receive the object by handing over from the blindfold human to represent the blindness. To determine the target of the object and human hand, we implement the 6DOF pose detection using a marker and hand detection using the Single Shot Detection model in Deep learning for planning motion using 9DOF arm robot with hand. We finally experiment and evaluate the tasks from blindfold-robot handover tasks.



OS24-5 Gait Learning Method for Quadrupedal Robot Using Chaos Time-series Analysis

Yuehang Ma, Kaori Watanabe, Hidekazu Suzuki (Tokyo Polytechnic University, Japan)

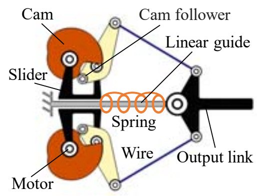
In the field of pet robots and robot-assisted therapy (RAT), characterization of animal motion is important for the development of robots resembling various animals. This paper presents a method for the generation of animal gait in quadrupedal robots. In this study, we employed AIBO as an experimental quadrupedal robot and generated the gait of the robot on the basis of an animal's gait. In the previous study, we optimized the mono-leg orbit, which can efficiently output a propulsive force, by imitating a dog's gait using a genetic algorithm. Moreover, we generated the quadrupedal gait of AIBO using both the optimum orbit of the mono-leg and an animal's gait, classified as the gait of a walking dog based on zoology. In this report, minor deviation of parameters for each joint was corrected to realize the stable gait on the ground.



OS24-6 Development of Antagonistic High Power Joint Mechanism with Cams

Katsuaki Suzuki[1], Yuya Nishida[1], Takashi Sonoda[2], Kazuo Ishii[1]
([1]Kyushu Institute of Technology, Japan, [2]Nishinippon Institute of Technology, Japan)

Acquiring flexible and agile behaviors as seen in biological systems, the robot can achieve acrobatic movements such as jumping and throwing. These acrobatic movements are expected to extend the range of robot activity. In this research, we propose a special mechanism using a pair of motors, springs and cams, which has three functions: normal operation, instantaneous operation and variable rigidity. Then, we derived a mathematical model of the mechanism and discussed the input-output characteristics of the mechanism by changing the design parameters, and analyzed the difference between the theoretical and measured results.



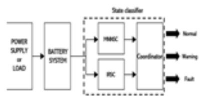
GS abstracts

GS1 Neural Networks (6)

GS1-1 Neural Network and Internal Resistance based SOH classification for lithium battery

Jong-Hyun Lee[1], Hyun-Sil Kim[2], In-Soo Lee[1] ([1]Kyungpook National University, Korea)
([2]Naval Combat Systems PMO Agency For Defense Development, Korea)

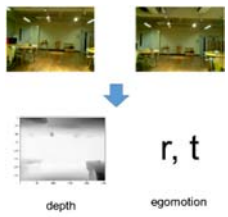
Today, lithium battery is used in various fields. Therefore, for stable use, it is important for the device and the system quickly to detect the defect occurring in the battery and diagnosis the fault accurately. Battery fault can be diagnosis by measuring the state of health (SOH) of the battery, and SOH is changed by various operating conditions. In this paper, the battery SOH monitoring system was implemented to diagnosis the fault of a battery cell through MNNSC (Multilayer Neural Network State Classifier) and IRSC (Internal Resistance State Classifier). In this method, MNNSC utilized the discharge voltage data that was obtained by operating the lithium battery cell at high temperature. On the other method, IRSC uses open circuit voltage, terminal voltage and current to calculate internal resistance. From the experiment results, we know that the proposed battery SOH monitoring method diagnosed the state of battery very well.



GS1-2 Estimation of Self-Posture of a Pedestrian Using MY VISION and Deep Learning

Tomoyuki Kurosaki, Joo Kooi Tan (Kyusyu Institute of Technology, Japan)

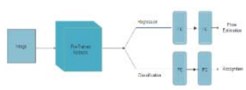
In this paper, we aim at developing a system that can be used easily even by elderly people; thus the use of a single wearable camera, resulting in an inexpensive system. Walking posture of a user equipped with a self-mounted camera is estimated using a pair of images obtained from the camera. Because the parallax of the image pair obtained from the chest-mounted camera is normally small when walking straight, it is difficult to estimate camera motion. So we employ a convolutional neural network in this research. It consists of three networks; a bootstrap net, an iterative net and a refinement net. In the bootstrap net and the iterative net, optical flows, camera motion and a depth image are alternately estimated to improve accuracy. In the refinement net, the resolution of the depth image is increased.



GS1-3 Simultaneous Space Object Recognition and Pose Estimation by Convolutional Neural Network

Roya Afshar, Zhongyi Chu (Beihang University, China)
Shuai Lu (Beijing University of Chemical Technology, China)

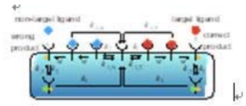
The increasing population of orbital debris is considered as a growing threat to space missions. For this purpose, Convolutional Neural Network was implemented based on transfer learning and data augmentation in order to conduct satellite classification and pose regression. In addition, the effects of un-centered and noisy images as well as different illumination conditions were analyzed by implementing different pre-trained networks. Based on the results, the present method could identify satellites and evaluate their poses against different space conditions effectively.



GS1-4 An error correction mechanism for reliable chemical communication systems

Masashi K. Kajita (The University of Tokyo, Japan)

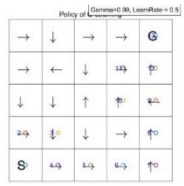
Chemical communication systems, such as bio-inspired chemical sensory systems or biological cells, sense the environment by detecting target ligand molecules, which convey environmental information. However, non-target ligands, similar to the target ones, are ubiquitous in the environment and can hamper accurate information transmission. In this work, we investigate an error correction mechanism for reliable chemical communication and find an intuitive understanding of how the mechanism can amplify the small difference between the target and non-target ligands. We also demonstrate that the mechanism can balance accuracy and output intensity. Our approach may provide a method to design a reliable chemical communication system.



GS1-5 A Reinforcement Learning-Based Path Planning Considering Degree of Observability

Yong Hyeon Cho, Chan Gook Park (Seoul National University, Republic of Korea)

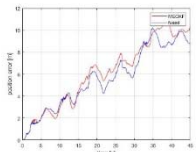
This paper presents a novel way to find a path using degree of observability as a reward for the reinforcement learning in the INS/GNSS loosely coupled system. In the proposed algorithm, an agent follows the Dubin's car model in a grid map, using a degree of observability to update Q-value. Various ratios between the penalty and reward at each step show different trajectories, and the specific state's degree of observability at the endpoint of a grid map is compared with one another. The progress is shown by a computer simulation.



GS1-6 A Performance Analysis of Pose Estimation Based on Two-View Tracking and Multi-State Constraint Kalman Filter Fusion

Tae Ihn Kim, Jae Hyung Jung, Chan Gook Park (Seoul National University, Republic of Korea)

This paper presents a performance analysis of two-view tracking and Multi-State Constraint Kalman Filter (MSCKF) fusion for a pose estimation. The system and measurement model of both two-view tracking and MSCKF are derived based on the fusion condition. The simulation result of the fused algorithm using the Drone Racing dataset, collected from an aggressive flight of micro aerial vehicle (MAV), shows the performance improvement of both attitude and position estimation compared to the performance of MSCKF.

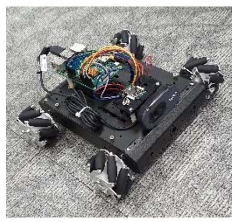


GS2 Control Techniques (5)

GS2-1 ORB-SLAM based Sensor Fusion Algorithm for Real-Time Precision Driving

Yong·Jin Ock, Zhan·Ming Gu, Jang·Myung Lee (Pusan National University, Korea)

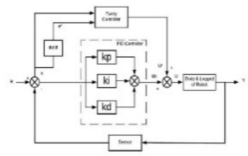
There was a problem that it was not possible to determine the exact posture and position with a single image alone. To compensate for this, additional IMU sensor and encoder sensor should be installed and calibrated. At this time, the encoder sensor acquires information about the distance traveled and the attitude of the mobile robot. The IMU sensor measures the attitude error caused by the sliding and friction of the mobile robot and acquires the slope information of the current terrain. As a result, by combining the location information acquired by using the SLAM and the complex location information of the IMU sensor and the encoder sensor, precise position control is possible even in a space without many feature points.



GS2-2 Gait Control of A Four-Legged Robot with Fuzzy-PID Controller

Arphakorn Kunha, Amornphun Phunopas, Wisanu Jitviriya
(King Mongkut's University of Technology North Bangkok, Thailand)

The mobile robot has developed as the legged mobile robot. Which has a high locomotion performance on smooth and non-smooth surface Meanwhile, the legged mobile robot movement should be have algorithm for control on the legged. Because the legged robot has required self-balance and gait that imitated from the locomotion behavior of four-legged animal, such as dog cat horses, etc. Therefore, this paper aims to describes the control system design of the four-legged robot through the PID that it is a controlling of gait. The fuzzy model has used as a determined a range of error in system. But the experiment has emphasized on walking-trot in the gait of animal



GS2-3 Hybrid Force/Position Teaching and Control Method for 6 DoF Manipulator utilizing f-PAWTED

Quang-Trung Chu, Hiroki Tanaka, Hideki Inuzuka, Yoshifumi Morita (NITech, Japan)
Masao Sakai (Aichi Pref., Japan)

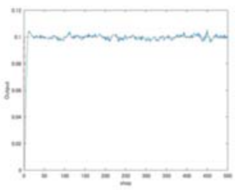
We have proposed a direct teaching method coupling with a specialized device, Parallel Wire Teaching Device (PAWTED) with a force/torque sensor, named f-PAWTED so that operator can easily teach any desired trajectories to the 6 DoF robot manipulator by using a human hand instead of a teaching pendant. By applying the f-PAWTED and hybrid position/force control method, the robot can accurately reproduce the teaching trajectory in both position and direction, as well as force and torque in 3-dimensional space without consuming huge amount of time. In this paper we propose a teaching system based on hybrid position/force control measure to teach robot in both position, orientation and force. The experimental results proved that we can precisely playback desired trajectories by utilizing f-PAWTED.

f-PAWTED

GS2-4 A Study on Generalized Predictive Control in Consideration of Noise

Akira Yanou (Kawasaki University of Medical Welfare, Japan)

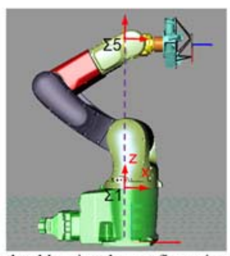
Generalized Predictive Control (GPC) is one of the model-based control methods. The control law is derived through the performance index based on the sum about the squares of control input and the squares of the error between reference signal and output prediction. Although coprime factorization approach has been used in order to extend the conventional control law in the previous researches, there has been a possibility that the order of the derived control law becomes high. Therefore, this paper extends GPC through newly defined output prediction and proposes the method to re-design the control law or the characteristic from noise to output with keeping the closed-loop transfer function. Numerical example is shown to check the characteristic of the proposed method.



GS2-5 Simulation Study on Emergency-Stopping Avoidance Control due to Singularity During Teaching Operation with Parallel Wire-Type Teaching Device

Hideki Inuzuka (NITech, Japan), Masao Sakai (Aichi Pref., Japan), Yoshifumi Morita (NITech, Japan)

It is possible to easily teach complex motions to industrial robots by the direct teaching system using our developed teaching device (PAWTED). However, practical experiments have shown that there are still some problems. One of them is the interruption of teaching operation caused by the emergency stop feature of the robot due to singular configuration during teaching operation. In order to solve this problem, we proposed a control method to avoid the emergency stop of the robot due to shoulder singular configuration by correcting robot motion using a unique and flexible characteristic of the PAWTED. Specifically, we considered correcting the robot motion by changing the target value that the robot follows the PAWTED. We confirmed the effectiveness of the proposed control method by conducting simulation study to verify that the emergency stop does not occur even in a singular configuration or its vicinity.



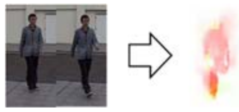
shoulder singular configuration

GS3 Vision & Image Processing (4)

GS3-1 Detecting Pedestrians and Moving Directions by a MY VISION System

Kenta Hori, Seiji Ishikawa, Joo Kooi Tan (Kyushu Institute of Technology, Fukuoka, Japan)

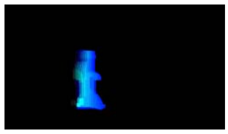
We propose a method for detecting visually impaired (user) pedestrians using images provided by a camera attached to the visually impaired (head). In the proposed method, a flow area different from the flow generated by the movement of the camera is extracted from the camera image. Pedestrians are detected by calculating the characteristics of a multi-scale cell HOG (MSC-HOG) that places cells along the outline of a person in the area. At the same time, a histogram is created in the flow area of the pedestrian, and the direction is recognized by comparing it with the flow histogram showing the walking movements learned in advance. This is to know the direction of the pedestrian toward the user wearing the camera. We experimented with the proposed method and showed its effectiveness.



GS3-2 Human Motion Recognition Using TMRIs

Cao Jing, Youtaro Yamashita, Joo Kooi Tan (Kyusyu Institute of Technology, Japan)

With the aggravation of the aging trend in Japan, the number of elderly people living alone has gradually increased, and the development of elderly people care system has begun to receive attention. Considering the importance of introducing an intelligent robot to the care of those elderly people in near future, this paper concentrates on automatic human motion recognition. This paper proposes a MHI(Motion History Image)-based method called TMRIs (Triplet Motion Representation Images) that solve in recognition, a self-occlusive motion problem particularly in the depth direction using a single camera. The performance and effectiveness of the method are verified by experiments.



GS3-3 Automatic Extraction of Abnormalities on Temporal CT Subtraction Images Using Sparse Coding and 3D-CNN

Yuichiro Koizumi, Noriaki Miyake, Huimin Lu, Hyoungseop Kim
(Kyushu Institute of Technology, Japan)
Takatoshi Aoki (University of Occupational and Environmental Health, Japan)
Shoji Kido (Osaka University)

In recent years, the rate of death from cancer has tended to increase in Japan, and the number of deaths from lung cancer in particular has been increased. CT device is effective for early detection of lung cancer. However, there is a concern that increasing the performance of CT will increase the burden on doctors. Therefore, by presenting a "second opinion" in the CAD system, the burden on the doctor is reduced. In this paper, we develop a CAD system for automatic detection of lung cancer from 3D CT images. Our proposed method using temporal subtraction technique, sparse coding and 3D convolutional neural network. Also, to confirm effective sparse coding parameters in CNN, an image for each parameter was used as an input image. We applied our method to 31 cases and obtain the result that sparse level contributed most to the score.

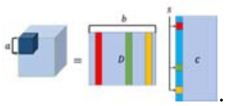
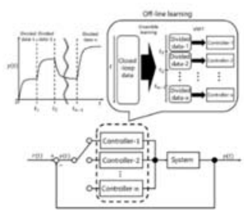


Fig. 1 Sample of image reconstruction by sparse coding. Combining base size a, base number b, and sparse level s.

GS3-4 Design of a Data-Driven Multi Controllers Using VRFT and Ensemble Learning

Takuya Kinoshita, Yuma Morota, Toru Yamamoto (Hiroshima University, Japan)

Data-driven control has been focused as a scheme of directly designing a controller from experimental data, and virtual reference feedback tuning (VRFT) was proposed for linear time-invariant systems. In the field of machine learning, the ensemble learning was proposed to improve the accuracy of prediction by using multiple learners. In this study, a design scheme of data-driven controllers using ensemble learning and VRFT is newly proposed for linear time-varying systems. The ensemble learning can divide the linear time-varying system into some sections that can be regarded locally as linear time-invariant systems.



GS4 Robotics 1 (4)

GS4-1 Design and Modeling of an Automatic Cartesian Farming Robot

Wisanu Jitviriya, Amornphun Phunopas
(King Mongkut's University of Technology North Bangkok, Thailand)
Eiji Hayashi (Kyushu Institute of Technology, Japan)

Currently, farming robots have become increasingly for households. The combination of modern technology and agriculture that is automatic system design, it is extremely accurate but also cost less than hiring human labor to work. All processes of the farming robot must be organic. The development of Cartesian Coordinate Robot (CCR) can operate in the multi-function such as tillage, applying fertilizer, sowing, and watering plants. In this research, which focuses on image processing to detect the plants and eliminate the weeds. The OpenCV library was used to detect the color green from the plant leaves, and from that, the program will do the rest. The elimination part of the system was constructed based on a brand-new idea that had never been done before until now. From all the experiments done, the conclusion looked promising, the blades were able to function at 80 percent efficiency

GS4-2 K-APF Algorithm to Avoid Obstacles in Path Planning

Dong-Kyo Jeong (Pusan National University, Korea)
Jang-Myung Lee (Pusan National University, Korea)

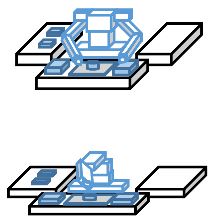
In this paper, a K-APF algorithm has been proposed to resolve the local minima and unstable path problems that occur when the conventional APF (Artificial Potential Field) algorithm is used for path planning in a real environment, which is the most popular path planning algorithm for robot manipulators and mobile robots, To avoid the local minima with the conventional APF, repulsive coefficients have been added to the potential field and the unstable path has been smoothed by the Kalman filter. In order to demonstrate the performance of the proposed algorithm, it is compared to A * star and Dijkstra algorithms through the real experiments in terms of accuracy and speed.

GS4-3 Selection System of Robot type for cell assembly production (Production efficiency comparison of single arm robot and double arm robot)

Takahiro Watanabe, Hidehiko Yamamoto, Takayoshi Yamada (Gifu University, Japan)

The purpose of this research is to compare the production efficiency of single arm robot and double arm robot. This research determines which robot is suitable for an assembly cell production, a single arm robot or a double arm robot. First, we develop the system to use a double arm robot to determine the best parts location by genetic algorithm (GA). The system consists of two modules, the conditions module and GA module. The conditions module reads work data for robot and sets up various parameters required for GA. The GA module decides the efficient arrangement place of parts, jigs and robot hands by GA and outputs the acquired arrangement visual images. Next, we use a single arm robot at the same parts location and compare the working time with a double arm robot. Finally, we conclude which robot is suitable for an assembly cell judging from the results of a total robot working time.

Cell type assembly machine



GS4-4 Curvature Surface Magnetic Wheel Climbing Robot with Adaptive Electromagnetic Adhesive Force

Richit Palangwatanakul, Apisit Thungsang, Kaned Thungod (Mahasarakham University, Thailand)
Arsit Boonyaprapasorn, Wanachart Borisut (Chulachomklao Royal Military Academy, Thailand)
Thavida Maneewarn, Suriya Natsupakpong (King Mongkut's University of Technology Thonburi, Thailand)
Thunyaseth Sethaput (Sirindhorn International Institute of Technology, Thammasat University, Thailand)

Various industrial structures or machines mostly consist of different shapes of ferromagnetic curvature surfaces. The magnetic wheel climbing robot is the suitable approach for achieving both adhesion and locomotion of the inspection robot. However, the adjustable magnetic force for robot adhesion is necessary, especially when the thickness of the surface is not uniform or the variation of the air gap between the magnetic adhesion units caused by the curvature of the surface. This can lead to the insufficient adhesive force. Furthermore, unnecessary driving torque of the motor to actuate the climbing robot from the over design of the magnetic adhesive force from the magnetic wheels can be avoided. Due to the level of the adaptive adhesive force is necessary to be considered, we designed the adaptive electromagnetic adhesive force mechanism for the curvature surface climbing robot with magnetic wheels. The PID controller was employed to control the electromagnetic force, and the adhesive force was measured by a load cell. This measurement signal was used as a feedback signal. In the paper, we investigated the capability of this adjustable magnetic force system. Five aspects of experimentation were implemented. It was clear that the light weight electromagnetic force adjustment mechanism could provide the flexibility to regulate the adhesive force for the magnetic robot while traveling on the ferromagnetic curvature surface.

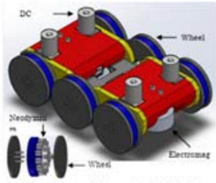


Fig. 1. Schematic of the magnetic wheel climbing

GS5 Robotics 2 (4)

GS5-1 Intention Classification of a User of a Walking Assist Cart by Using Support Vector Machine

Noritaka Sato, Tomoki Yokotani, Yoshifumi Morita (Nagoya Institute of Technology, Japan)

As the number of elderly people increases, the demand for walking assist cart is increasing. To develop better assist function for the cart, we focused on the prediction of the intention of a user of the cart (Intention classification of the user), when the state of the cart changes. As the first step of the study, the forces and torques to the cart from the user's hands, and the rotational velocities of the wheels are sensing. And the Support Vector Machine, which is one of the most popular and effective techniques in the artificial intelligence is used for intention classification. As a result of the experiment with five subjects, we confirmed that our method was able to predict the intention of the user with enough accuracy.

GS5-2 Development of an Environmentally Adaptable Autonomous Mobile Robot

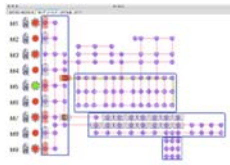
Naokazu Iwata, Joo Kooi Tan (Kyushu Institute of Technology, Japan)

In recent years, much attention has been focused on service robots for supporting a human life. These kinds of robots are necessary to adapt flexibly to various kinds of usage, scenes, or even to environmental changes. This paper proposes an autonomous mobile robot system that adapts to various environments and travel routes based on the image features obtained from a RGB-D camera. To realize the robot system, techniques are proposed for estimating self-position of a robot using local features of images, segmenting an image for finding a movable area, and planning a route for finding the destination. In the route planning, junctions are labelled numbers for the robot to travel in the order. The developed mobile robot travels to the destination employing the information on the estimated road region, its self-position and the planned route. Experimental results, traveling two routes in a real environment, show effectiveness of the proposed system.

GS5-3 Automated Task and Path Management for Industrial AGVs in Foam Manufacturing Plant

Amornphun Phunopas[1], Wisanu Jitviriya[2], Noppadol Pudchuen[3], Songklod Tunsiri[4], Eiji Hayashi[5] ([1,2,3]King Mongkut's University of Technology North Bangkok, Thailand, [4]Urban Community Development College, Thailand [5]Kyushu Institute of Technology, Japan)

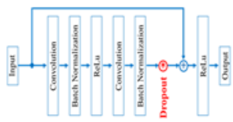
AGVs are increasingly used in the automated warehouse with a high demand for changing traditional workflow management to industrial 4.0. The heart of the computerized system is the central software that can distribute work functions from the queues and manage the AGVs' traffic. On the 2D floor plant layout, the girds are initially from marked points or the place that AGVs have to transit to do an assigned task. This research proposes autonomously generating paths via four nearest grids and path switching scenarios. The results show the generated paths with sequential tasks concurrently in random conditions. The task management method can prevent the AGVs' crash and bottleneck from the operation of nine machines in the foam manufacturing plant.



GS5-4 Object Tracking Method Considering Time Series Information Using Re3 with Stochastic Depth

Taichi Kitayama, Hyoungseop Kim (Kyushu Institute of Technology, Japan)

Container transportation plays an important role for international cargo transportation. The amount of container cargo movement in the world has increased over the years. In recent years, container terminals are required to automate cargo handling operations from the viewpoint of improving the working environment and ensuring the safety of employees. However, the automation of cargo handling work has not yet been realized in Japan. One of the reasons is that the technology for measuring the relative position between the container and the suspension has not been established. In this study, we propose a tracking method of container gripping area using deep learning for the purpose of relative position measurement. A hybrid tracker called Re3 using a convolutional neural network (CNN) was used. However, this model alone has a problem that global features cannot be extracted well. To solve this problem, this method incorporates a model called Stochastic Depth.

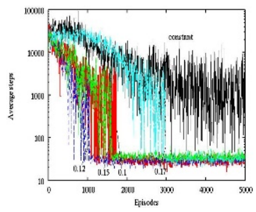


GS6 Application Techniques (9)

GS6-1 A Promoting Method of Role Differentiation Using the Learning Rate Getting Periodically Negative Value in Multi-agent Reinforcement Learning

Masato Nagayoshi, Simon Elderton (Niigata College of Nursing), Hisashi Tamaki (Kobe Univ.)

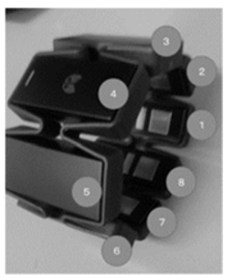
There have been many studies on the multi-agent reinforcement learning (MARL) in which each autonomous agent obtains its own control rule by RL. Here, it is considered that different agents having individuality is more effective than uniform agents in terms of role differentiation in MARL. In this paper, we propose a promoting method of role differentiation using a wave-form changing parameter in MARL. Then we confirm the effectiveness of role differentiation by the learning rate getting periodically negative value through computational experiments using "Pursuit Game" as one of multi-agent tasks.



GS6-2 Verification of a Combination of Gestures Accurately Recognized by Myo Using Learning Curves

Kengo Kitakura, Hideyuki Tanaka (Hiroshima University, Japan)

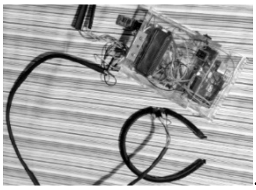
A technique for hand gestures recognition from sEMG (surface ElectroMyoGraphy) is useful for extending means of human communication. Recently, the Myo armband is one of the most popular sEMG acquisition systems, because it is relatively inexpensive and easy to remove. This paper studies verification of a combination of hand gestures recognized by using the Myo armband as an input device. To this end, relationship between data distribution and learning curves is investigated for binary classification problem and multi-class classification problems. A verification method is then proposed for finding a combination of gestures accurately classified. Since data distribution depends on removal of Myo, three cases of experiments are carried out: For the first case, Myo is not removed between training and test data acquisition. For the second and third case, Myo is removed; there is a little time or several days between data acquisition.



GS6-3 Augmentative and Alternative Communication Device Based on Head Movement to Aid Paralyzed Victims with Speech Disabilities

Vihanga Ashinsana Wijayasekara, Torin Wirasingha
(Informatics Institute of Technology, Sri Lanka)

Augmentative and Alternative Communication (AAC) devices have been proven to be as an alternative to help people, who are having communication difficulties. In this paper, high tech AAC device focusing on paralyzed victims or other motor disabilities which allow them to control with their head gestures is proposed. The proposed system consists of a mobile app, a controller and a head tracker. Easy to familiar with and less material cost are some of the key advantages of this device.



GS6-4 The research about editing system of performance information for player piano. -Make inferences about whole musical composition by using DP matching system-

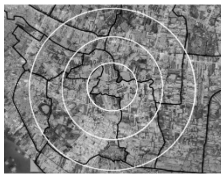
Mami Ezaki, Eiji Hayashi (Kyushu Institute of Technology, Japan)

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. In this paper, we developed a system that automatically estimates the performance expression of unedited music using edited performance data and score data. We described a phrase search using DP matching and a method for selecting an optimal phrase for inference, and evaluated the entire song.

GS6-5 Evaluating Public Perception using Fuzzy Logic: A case study of Praeksa Mai dumpsite in Samut Prakan, Thailand (withdraw)

Sun Olapiriyakul, Khemika Kongpetch
(Sirindhorn International Institute of Technology, Thammasat University, Thailand)

Particularly in developing countries, the issue of poor public perception towards solid waste management systems is commonly observed across rapidly urbanizing regions. The evaluation of the negative impact a waste site on public perception is the first step towards improving the well-being of local communities. This research intends to resolve the negative perception issues by developing a constructive impact assessment methodology for a dumpsite in Samut Prakan, Thailand. The evaluation of public perception impact is made, considering the factors affecting public concerns over environmental and health risks and population exposure. The fuzzy logic concept is used to account for public perception variability across different impact distances. The calculated impact scores of the dumpsite are useful for problem prioritization and benchmarking purposes.



GS6-6 A Development of a Model CubeSat with an Amateur Radio Transceiver for Education on Satellite Communication

Masahiro Tokumitsu, Kentarou Konishi (National Institute of Technology, Yonago College, Japan), Taku Takada, (National Institute of Technology, Kochi College, Japan), Fumio Asai (Member of the Radio Amateur Satellite Corporation), Makoto Wakabayashi (National Institute of Technology, Niihama College, Japan)

We report the development of a Model CubeSat for education on satellite communication. The communication between satellites and ground stations is an essential part of satellite operations. As a critical part of the satellite systems, we focused on satellite communication to train the student so that they can develop and operate the CubeSats. The Model CubeSat equips with fundamental components that are necessary as the satellites such as an onboard computer, sensors, communication, and power supply. The software programs of the Model CubeSat provide two functions: controlling and monitoring satellites, and data transmission of telemetry data and image data. The students can study the communication equipment and operations of satellite communication by our Model CubeSat.

GS6-7 Performance Evaluations on Data Estimation Technique with Statistical Properties of Telemetry Data for Corrupted Data in Amateur Satellite Communication

Yusuke Teranishi, Masahiro Tokumitsu (National Institute of Technology, Yonago College, Japan), Taku Takada, (National Institute of Technology, Kochi College, Japan), Fumio Asai (Member of the Radio Amateur Satellite Corporation), Makoto Wakabayashi (National Institute of Technology, Niihama College, Japan)

We propose a novel technique on a data estimation for satellite communication and report performance evaluations of the proposed technique. The proposed technique estimates the original data from the incorrect data, error detection codes, and supplemental information. Firstly, we investigated the possibility of data estimation for long length data by a brute-force technique. The first simulation results showed that the brute force technique succeeded in estimating the original data with the length 1024 bit from the incorrect data. Next, we considered the novel technique on the data estimation for the incorrect data. The second simulation results showed that the proposed technique succeeded in estimating the original data with the length 1024 bit from the incorrect data.

GS6-8 Heritage Building Design Properties: Development of As Built Drawing by UAV Application via 3D Laser Scanner

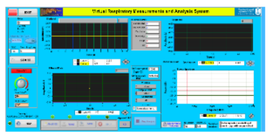
Hazry. D, Azizan. M. A, Noriman. N.Z, Romeli. N (Universiti Malaysia Perlis)

Inadequate information that represent the existing figure of the building after transmission of years and the eagerness of the Malaysian’s architect to develop frantically modern development design that gradually neglecting the essence of cultural perseverance of the heritage building has been brawled excessively. Taking a human – factor constraint to conduct the physical survey, the technologies such as computers and digital media have given beneficial value that are time saving and less risky. Hence, the research aim for the development of the as-built drawing for Masjid Alwi, Kangar, Perlis has been completed via UAV Application and 3D Laser Scanner, and results were recorded on the structural and architectural properties of Masjid Alwi and analyzed accordingly. Therefore, the needs of having an adequate as-built drawing due to the depletion of the ancient documentation and as a baseline for the designer maintaining the cultural identity in a way to design new development.

GS6-9 A Virtual System for Measurements and Analysis of the Respiratory Sounds for Diagnosis of Respiratory System

Ali S. AlMejrad (King Saud University, Kingdom of Saudi Arabia)

Respiratory problem is one of the most common health problems occurring in Saudi Arabia due to the continuous changes of the weather in addition to surface winds that cause dust during all seasons in the year. This problem affects most people and especially children and elderly in addition to the adults who have respiratory problems such as asthma. This is one of the problem that need emergency care as soon as it occurred. The goal of our research is to develop a compact respiratory diagnostic system using advanced signal processing that can be used remotely via the virtual instrumentation technology to help accurately diagnosis at early stages of respiratory diseases. It can also overcome the lack of expert physicians in rural regions and some urban clinics or health centers. The proposed system will be implemented using virtual instrumentation (VI) that consists of computer, microphone with simple analog circuit, digitizer and LabVIEW software (National Instruments Inc.). VI has been designed for easy measurement and analysis. In addition to that, it has features and ability to control the whole system of acquisition, play, display, processing, advanced analysis, display and report printing of the different acquired respiratory signals and diagnosis.. The developed system has been tested with real respiratory signals of normal and abnormal cases that proved to be efficient system while dealing with many respiratory problems conditions.



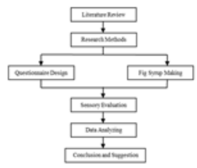
GUI and Front Panel of Developed Virtual Respiratory Measurements and Analysis System

GS7 Poster (4)

POS7-1 The Development and Evaluation of Fig's Leaf Syrup

Shang-Hui Li, Pei Hi Zheng, I Chih Chiang, Yu Ting Su, Syue Sheng Lin (Far East University, Taiwan)

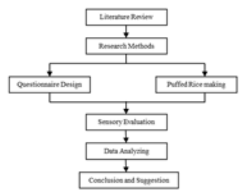
The figs sold in the market are only fruits and their processed products, the use of fig's leaves are relatively rare. But, the fig's leaves contain furanocoumarins, flavonoids, pectins, resins, sugar and vitamin C (Meng et al., 1996). Perhaps the special aroma of its leaves, or perhaps the astringency of the leaves, is less acceptable to consumers. This study used fig's leaves as the main ingredient with crystal sugar and different kinds of vanilla formula to developed if more consumers willing to accept fig's leaves syrup or not. Sensory evaluation was processed after the finished product was diluted. The obtained results was statistically analyzed by SPSS12.0, and analyzed by single factor variance and number allocation table analysis.110 students and teachers of Far East University, Tainan were selected for convenience sampling methods and 100 valid questionnaires were collected. The content of the evaluation was based on the appearance, aroma concentration, aroma preference, sweetness concentration, sweetness preference, astringency, aftertaste, overall acceptability, and willingness etc. And find out the most suitable formula of the fig's leaves syrup. In terms of appearance, the fig's leaf syrup with other vanilla is more acceptable than the original fig's leaf syrup. As for lemon fig's leaf syrup, although the aroma is the strongest, but the overall acceptance and purchase intention is the lowest due to its high astringency. Mint fig's leaf syrup, its aroma is accepted with the highest level. The results of the comprehensive data analysis found that the most suitable formula is fig's leaf 60gs, water 400gs, crystal sugar 200gs and mint 15gs.



POS7-2 The Research and Development of Fruit Puffed Rice

Shang-Hui Li, Yi-Ru Wang, Yi Ting Liu, En-Yi Lu, Cheng Han Li, Fang Quan Zhang
(Far East University, Taiwan)

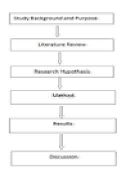
In recent years, Agriculture and Food Agency, FDA, Council of Agriculture, Taiwan, has vigorously promoted rice food products such as rice bread, rice noodles, rice seed strips, rice instant noodles, rice cakes, rice puffs and other products. Puffed rice is one of the most common traditional snacks in Taiwan. The ingredients are mostly rice, maltose, brown sugar, water and salad oil. But snacks on the market have gradually diversified due to the advancement of the times, moreover, the commercially available puffed rice flavor is too monotonous and generally not popular with young people. In this study, to make puffed rice based on natural freshly squeezed fruit juice with four different formulas of original flavor, pineapple, lemon and grape for sensory evaluation. 100 consumers were selected from Far East University, Tainan, by means of a convenient sampling method for formal favorite sensory evaluation. 95 valid questionnaires were collected. The evaluation included rice flavor, sweetness, acidity, degree of adhesion, brittleness and overall acceptance, and the results were statistically analyzed with SPSS12.0. The study results showed that pineapple-scented puffed rice had the highest level on crispness, the overall acceptance and the willingness to purchase. However, lemon- scented puffed rice had a low scent, low overall aroma, less fragile, high acidity, and low on overall acceptance and willingness to purchase. The most suitable formula for this study was rice scent 300g, pineapple juice 90g, salt 3g, maltose 90g, salad oil 24g, and brown sugar 160g.



POS7-3 The Research of Heath Western Cuisine – A Study of Aloe in Cooking

HsiYing Hsieh (Far East University, ROC Taiwan)

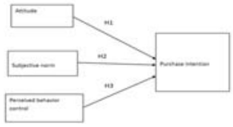
Aloe vera, is widely used in food, medicine and cosmetic products. With the improvement of the system of health care food, more and more consumers are willing to buy health care food and cosines. The study aims to understand the purchase intention of Aloe Vera in Cooking for heath western cuisine. The questionnaire will be surveyed in Tainan City. Data analysis of description, T-test, and Anova will be conducted by SPSS software. Based on the results of study, the future development and utilization of aloe vera cuisine are also discussed.



POS7-4 The Influence of Attitude, Subjective Norm, Perceived Behavior Control on Purchase Intention – A Study of the Green Restaurants in Taiwan

HsiYing Hsieh (Far East University, ROC Taiwan)

Green food is widely promoted for personal health and environmental benefits. Thus, decision making on consumers' purchase intention towards green restaurants in is a critical issue. The purpose of study is to understand how attitude, subjective norm and perceived behavioral control drive consumer influence purchase intentions of green restaurant. The questionnaire surveys were conducted to customers have consumed at green restaurants in Taiwan. In total, 250 valid questionnaires were obtained. Data analyses use the SPSS software to test hypotheses. Based the finding attitude, subjective norm and perceived behavioral control were significantly affect purchase intention for green restaurants. Implication and suggestion are then discussed.

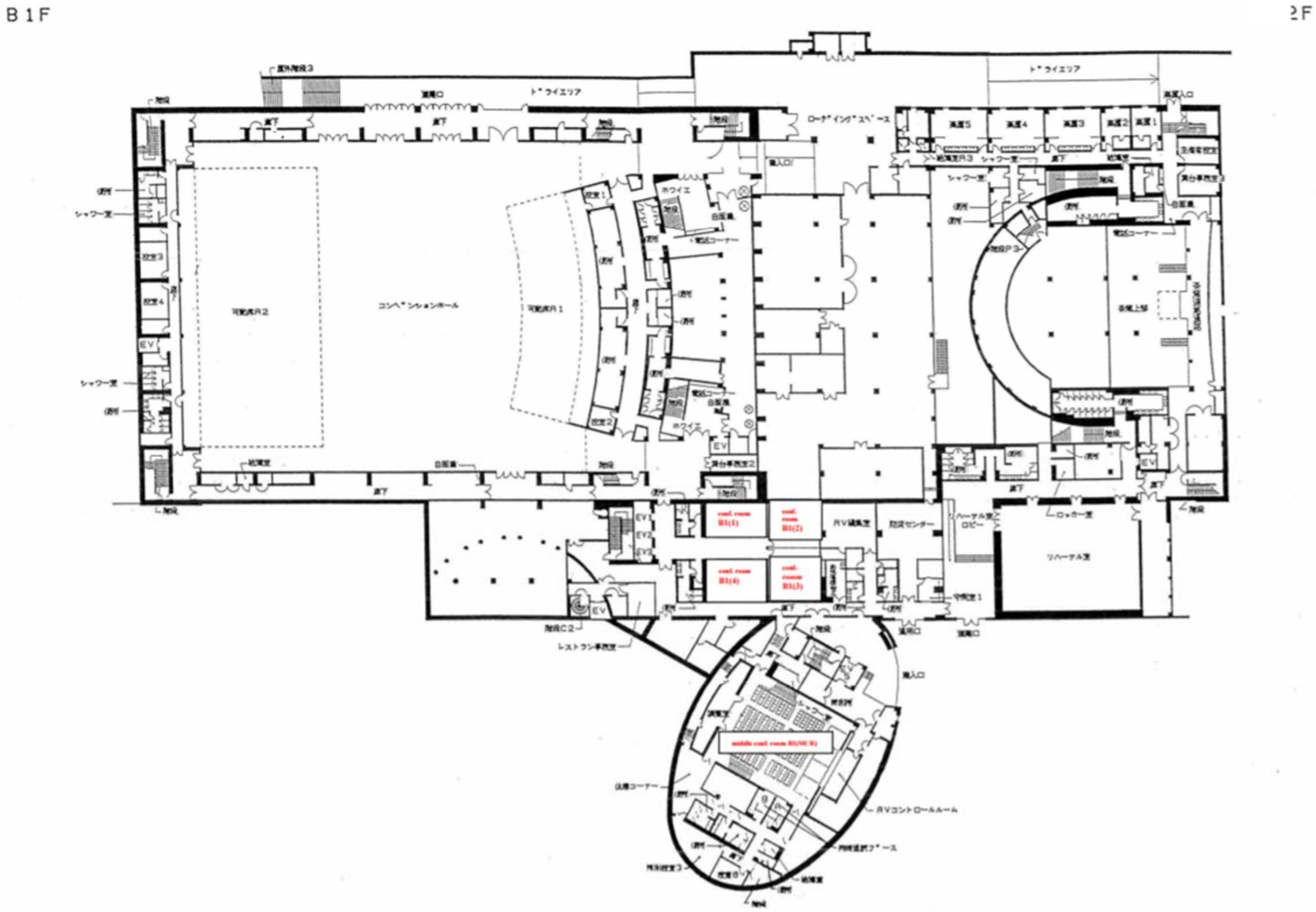


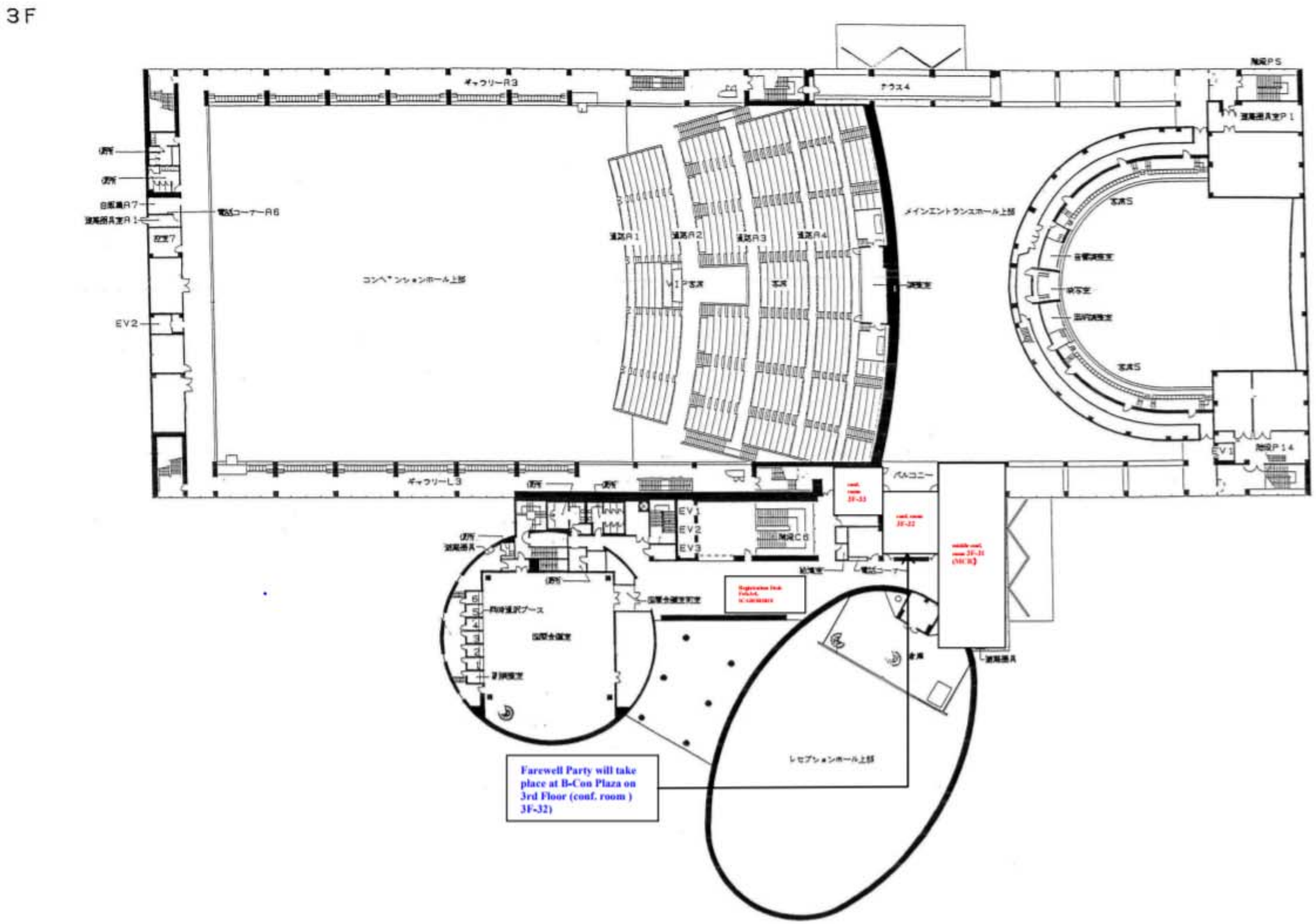
AUTHORS INDEX

Notation of session name

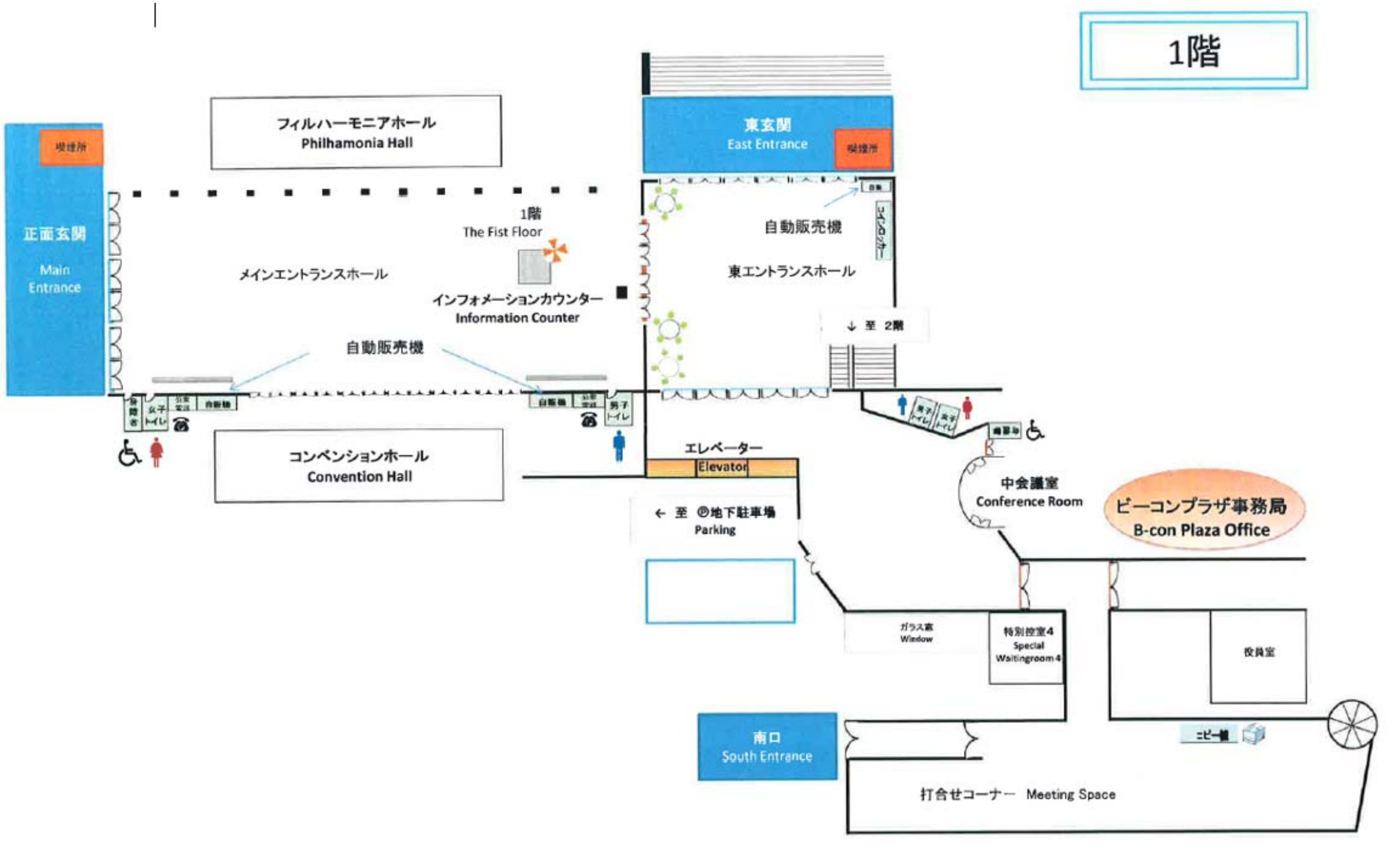
PS: Plenary Session IS: Invited Session, OS: Organized Session, GS: General Session,

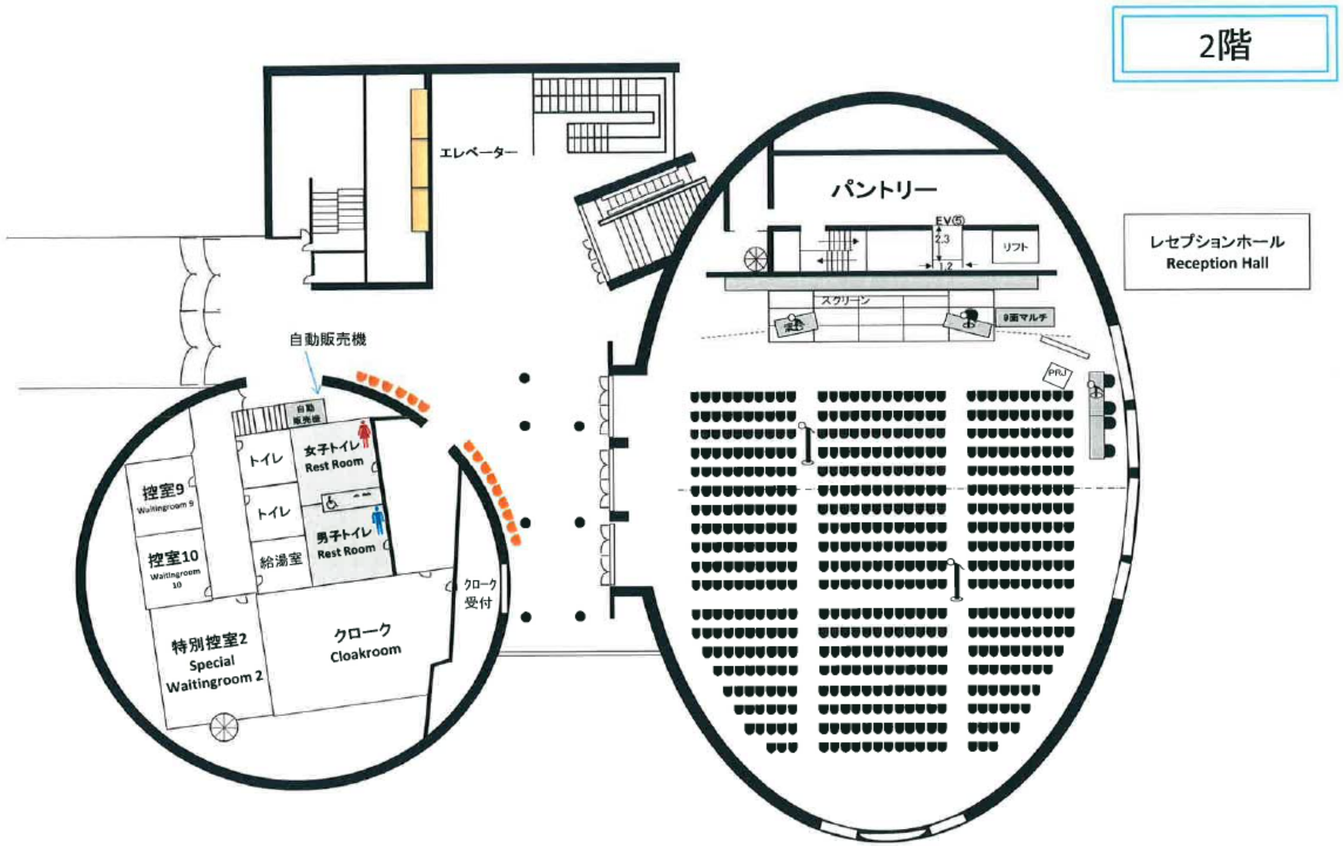
Note: 33/90 = (page no. in Technical Paper Index) / (page no. in Abstracts)

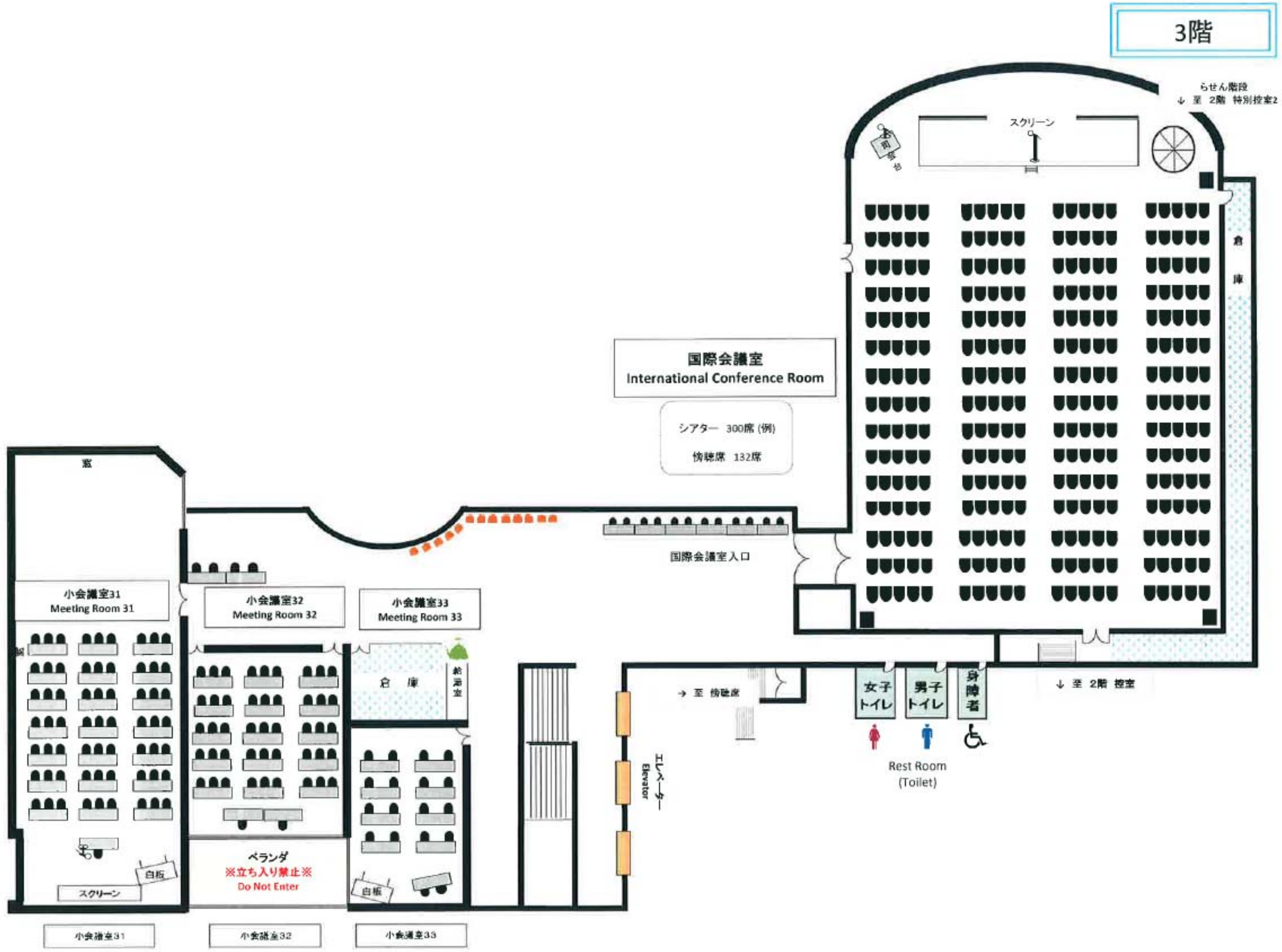

B1F
conf. room B1(1)
conf. room B1(2)
conf. room B1(4)
conf. room B1(3)
コンベンションホール
RV調整室
防災センター
リハーサル室


3F
バルコニー
EV2
EV1
EV2
EV3
Farewell Party will take place at B-Con Plaza on 3rd Floor (conf. room) 3F-32)


地下1階
一 般 道
General Road
芝広場
Grasses
私 道
Private Road
中会議室
Conference Room
ステージ
Stage
キッズスペース
Room for Kids
ベビールーム
Room for Kids
台車置き場 Hand carts
(荷捌き所) Room for Disposal of goods
館 外
Outside
ガラス窓
Window
リハーサル室
Rehearsal Hall
ロビー
守衛室
Security
Office
救護室
First-Aid
Office
小会議室3
Meeting Room 3
小会議室4
Meeting Room 4
防災センター
Disaster Prevention
Center
AV機械室
小会議室2
Meeting Room 2
小会議室1
Meeting Room 1
エレベーター
Elevator
予備：机、椅子置き場
男子トイレ
女子トイレ
倉 庫
← 至 フィルハーモニアホール 楽屋
↓ 至 B1F コンベンションホール


1階
フィルハーモニアホール
Philhamonia Hall
東玄関
East Entrance
正面玄関
Main Entrance
1階
The Fist Floor
メインエントランスホール
インフォメーションカウンター
Information Counter
自動販売機
自動販売機
東エントランスホール
↓ 至 2階
コンベンションホール
Convention Hall
エレベーター
Elevator
← 至 Ⓟ地下駐車場
Parking
中会議室
Conference Room
ビーコンプラザ事務局
B-con Plaza Office
特別控室4
Special Waitingroom4
南口
South Entrance
打合せコーナー Meeting Space


2階
エレベーター
パントリー
リフト
スクリーン
レセプションホール
Reception Hall
自動販売機
トイレ
女子トイレ
Rest Room
トイレ
男子トイレ
Rest Room
給湯室
控室9
控室10
特別控室2
Special
Waitingroom 2
クローク
Cloakroom
クローク
受付



3階
らせん階段
↓ 至 2階 特別控室2
スクリーン
倉
庫
国際会議室
International Conference Room
シアター 300席 (例)
傍聴席 132席
国際会議室入口
小会議室31
Meeting Room 31
小会議室32
Meeting Room 32
小会議室33
Meeting Room 33
倉 庫
給湯室
ベランダ
※立ち入り禁止※
Do Not Enter
白板
スクリーン
白板
小会議室31
小会議室32
小会議室33
→ 至 傍聴席
エレベーター
Elevator
女子トイレ
男子トイレ
身障者
Rest Room
(Toilet)
↓ 至 2階 控室

Towards Neuromimetic Computing

Takashi Kohno
Institute of Industrial Science, University of Tokyo, Komaba, Meguro-ku, Tokyo 153-8505, Japan

E-mail: kohno@sat.t.u-tokyo.ac.jp, http://www.sat.t.u-tokyo.ac.jp/~kohno

Abstract

Deep neural network brought a new era of neuro-inspired machine learning. Its splendid performance in pattern recognition is attracting both scientific and industrial resources and accelerating the field of "AI". In these days, neuromimetic computing is developing its presence in this field as a potential approach to the post-deep neural network, the next-generation AI. Silicon neuronal network (SNN) is an approach to the neuromimetic computing, which tries to construct an electronic circuit version of the nervous system. Its strong potential to innovate the computing paradigm from digital computing to "brain-compatible" computing facilitated IBM and Intel to develop SNN chips, but it still stays in the embryonic stage because of great hurdles. In this talk, we overview and discuss about trials in this field.

Keywords: neuromimetic computing, neuromorphic systems, silicon neuronal networks, artificial intelligence

1. Introduction

Alexnet[1] proved the artificial neural networks' potential to process huge-scale complex data more efficiently than the rule-based, theory-based, and other conventional machine learning algorithms. Deep neural network is the key technology for the modern artificial intelligence (AI) which is attracting increasing amount of resources from both scientific and industrial communities. This makes neuromorphic systems, ones inspired by or mimic the nervous system, recognized as a mostly potential approach to the AI.

Silicon neuronal network (SNN) is an artificial neuronal cells' network composed of electronic circuit versions of neuronal cells and synapses. It is a trial to develop an electronic circuit system that is compatible with the nervous system by reproducing its electrophysiological activities. It is not only a neuromimetic approach to the AI that has compatibility with the brain but also approaches to biomedical devices that complement lost brain functions[2,3] and high-speed brain simulators[4]. Ultralow-power analog circuit technology is applicable to SNNs because the timescale of neuronal activities is in the order of milliseconds which is far slower than the clock speed of digital computers. Field-effect transistors (FETs) in their subthreshold domain make it possible to realize a neuron circuit with the power consumption as low as less than 10 nW to several hundred pW[5,6,7]. In spite of its strong potential to innovate the computing paradigm from digital computing to "brain-compatible" computing, no analog SNN chip capable of information processing comparable to brain functions has appeared. How we can realize such SNNs?

2. Towards brain-compatible analog SNN

2.1. *Implementation Technology*

From the viewpoint of circuit implementation technology, integrating a huge number of nonlinear analog circuit blocks, the neuronal and synaptic circuits, is a big challenge because ultralow-power analog circuit is particularly susceptible to fabrication mismatch of the transistors. One of the most large-scale analog SNN chips in these days comprises 256 neurons and 128K synaptic inputs[8]. Their circuitry is designed so that their ability to produce and transmit neuronal spikes is as independent as possible of each transistor's nonlinear characteristics. This is achieved by adopting a simple neuronal model based on Integrate-and-fire (I&F)-based models and synaptic model that assume linear summation of synaptic currents. It makes possible to suppress the number of synaptic circuits by sharing one with multiple synaptic inputs. One of the bottlenecks in larger-scale integration is the capacity and power consumption of digital memory used for storing the synaptic efficacy. Analog nonvolatile memory devices such as memristors are the emerging technologies that are expected to contribute efficient implementation of neuromorphic systems. Because most of them have relatively low resistance, developing ultralow-power devices will be a key for the large-scale analog SNNs. Such devices are also expected to facilitate development of the SNNs that try to reproduce more complex neuronal activities. In such circuits, the effect of fabrication mismatch has to be compensated by applying appropriate bias voltages or currents to more than 10 parameter input lines[9]. Nonvolatile memory devices can provide compact and ultralow-power storage for these parameters.

Ultralow-voltage circuitry is an important technology for improving power efficiency. Because techniques of qualitative modeling make it possible to describe the neuronal models using arbitrary sigmoidal functions, finding ultralow-voltage circuits whose input-output characteristics are a sigmoidal curve is a key to improve the power efficiency. It was proved that ordinary 2-transistor inverter circuit can be used as an ultralow-voltage sigmoidal circuit[6]. Another work[10] proposed to utilize stochastic resonance in bi-stable circuits to produce sigmoidal characteristics. Because the sigmoidal characteristics in the ionic channels are also produced by stochastic resonance, this approach may enhance the mimetics of the SNNs to the neuronal cells' networks.

Another key technology is the ultralow-power neuronal spike delivery. TrueNorth chip[11] showed that a spike can be delivered by 2 to 20 pJ on a million-scale network chip by using asynchronous transmission technology. Inter-chip signal transmission generally requires far higher power, so the emerging inter-chip connection technology such as through silicon via (TSV) and near-field coupling are more important in these days.

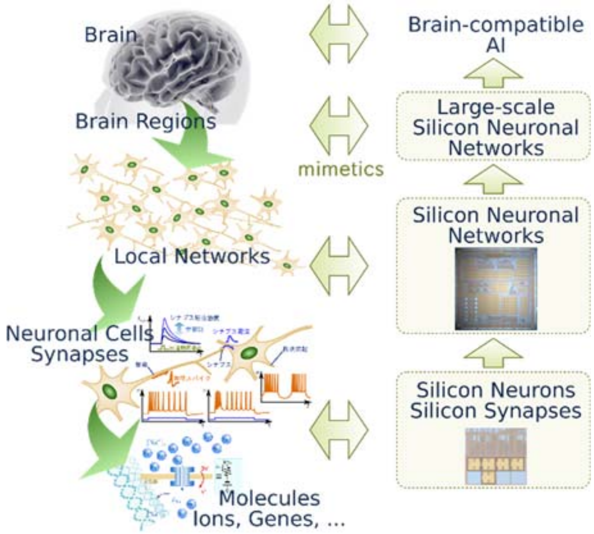


Fig. 1. Hierarchy in the brain and SNN

2.2. *Neuronal Modeling*

SNN is a bottom-up approach to the brain-compatible information processing systems. To reach its goal, the brain's deep hierarchy has to be followed step by step (Fig.1). The mechanisms and characteristics of the neuronal spike generation and transmission are being elucidated considerably by electrophysiological studies and the knowledge on high order brain function is being accumulated by cutting edge measurement technologies. However, it is left almost unknown how the collective activities of the neuronal spikes realize the information processing in the microcircuits in the brain. This is a most difficult question in the neuroscience, but its answer is imperative for SNNs to take the first step in the brain hierarchy: reproducing the information processing in the microcircuits.

As well as the theoretical studies on spiking neural network models, SNNs themselves are able to contribute by "analysis by synthesis" approach. Particularly, this approach matches spiking neuronal network models that focus on noise-dependent information processing such as a spatiotemporal spike pattern detection model[12] proposed by Masquelier et al. This model suggests that simple feedforward network with lateral inhibition can learn repeatedly given spatiotemporal spike patterns by the spike-timing-dependent plasticity (STDP) rule. Intrinsic noise in the analog SNNs helps verification and analysis of such biologically meaningful models.

References

1. A. Krizhevsky, ImageNet Classification with Deep Convolutional Neural Networks, *Advances in Neural Information Processing Systems* **25**, pp. 1097–1195, 2012.
2. F. Broccard, et al., Neuromorphic neural interfaces: from neurophysiological inspiration to biohybrid coupling with nervous systems, *J. Neural Eng.* **14**(4):041002, 2017.
3. G. Le Masson et al., Feedback inhibition controls spike transfer in hybrid thalamic circuits, *Nature* **417**(6891), pp. 854–858, 2002.
4. J. Schemmel, et al., A Wafer-Scale Neuromorphic Hardware System for Large-Scale Neural Modeling, *Proc. 2010 IEEE Int. Symp. on Circuits and Systems*, pp. 1947–1950, 2010.
5. N. Qiao, et al., Analog Circuits for Mixed-Signal Neuromorphic Computing Architectures in 28 nm FD-SOI Technology, *Proc. IEEE SOI-3D-Subthreshold Microelectronics Technology Unified Conference (S3S)*, 10.1109/S3S.2017.8309203, 2018.
6. I. Sourikopoulos, et al., A 4-fJ/Spike Artificial Neuron in 65 nm CMOS Technology, *Frontiers in Neuroscience* **11**(123), 2017.
7. T. Levi, et al., Development and applications of biomimetic neuronal networks toward brain-morphic artificial intelligence, *IEEE Trans. Circuits and Systems II: Express Briefs* **65**(5), pp. 577–581, 2018.
8. N. Qiao, et al., A reconfigurable on-line learning spiking neuromorphic processor comprising 256 neurons and 128K synapses, *Frontiers in Neuroscience* **9**(141), 2015.
9. T. Kohno and K. Aihara, A three-variable ultralow-power analog silicon neuron circuit, *Proc. 2016 International Symposium on Nonlinear Theory and Its Application*, pp. 190–193, 2016.
10. Y. Sakemi, et al., Noise-driven stochastic bistable circuits for brain-morphic systems, *Proc. 2018 International Symposium on Nonlinear Theory and Its Applications*, pp. 666–669, 2018.
11. P. Merola, et al., A million spiking-neuron integrated circuit with a scalable communication network and interface, *Science* **345**(6197), pp. 668–673, 2014.
12. T. Masquelier, et al., Competitive STDP-based spike pattern learning, *Neural computation*, **21**(5), pp. 1259–1276, 2009.

Innovative Robot - Robot/AI for Factory Automation-

Eiji Hayashi
Mechanical Information Science and Technology, Kyushu Institute of Technology
680-4, Kawazu, Iizuka, Fukuoka, 820-8502, Japan

Abstract

We have engaged in a regional revitalization project during the period from 2018 to 2022 as a subsidy grant which the Japan's government are promoting, by Yaskawa Electric Corporation, City of Kitakyushu and Kyushu Institute of Technology. In this project has been developed innovative robots combined Robot / Artificial Intelligence technologies.

Keywords: Factory Automation, AI/Robotics, Regional revitalization

1. Introduction

We think that AI and robots are already everywhere, especially not engineers due to the influence of the media. Certainly, their technologies have made enormous advances and progress in recent years like Siri, self-driving car system and so on as well as deep-learning neural network on image processing. The innovative robot for FA robot which means a revolution in productivity into next generation is described about what such a robot needs.

This paper reports on Robot/AI for factory automation about

2. Regional Revitalization Project in Kitakyushu

To develop and create an innovative robot with Yaskawa Electric Corporation, the project has been divided and corroborated issues into three groups with "Yaskawa Electric Corporation"(hereinafter, referred to as "Yaskawa") is shown in Fig.1. Until 1st target an innovative robot technologies and systems for 3C assembly (Computer, Communication and Consumer electronics), and a food industry are established.

Fig. 1. Research topics

2.1. *Talented persons in and out of Japan*

In this project has been invited 4 talented persons in and out of Japan in Table 1, to gain stronger supported on research and development of AI, factory automation and formation of robotics in city of Kitakyushu.

Table 1. Inviting Researcher in and out of Japan

Name	Affiliation
Hiroshi Hirukawa, Ph.D, Director	Robot Innovation Research Center, National Institute of Advanced Industrial Science and Technology （AIST）
Rolf Pfeifer, Ph.D., Professor emeritus	University of Zurich
Steve Cousins, Ph.D., Founder, CEO	Savioke
Yoji Yamada, Dr., Professor	Nagoya University

2.2. *Global cooperation*

To make a robotics base in Kitakyushu, new personnel are secured, continuous human resources development is needed with R&D. Therefore, we have built up effective relationship in oversea as is shown in Fig.2.

3. Example of system

The proposal system in Teaching less group is showed in Fig.3. This system has to synthesize various technologies using AI, vison and ontological Engineering and so on, and then is very complicated to control FA robot.

4. Conclusion

Now we have just started to develop an innovative robot for FA, to learn FA robot, and there are many issues in progress. We will accomplish our task with Yaskawa.

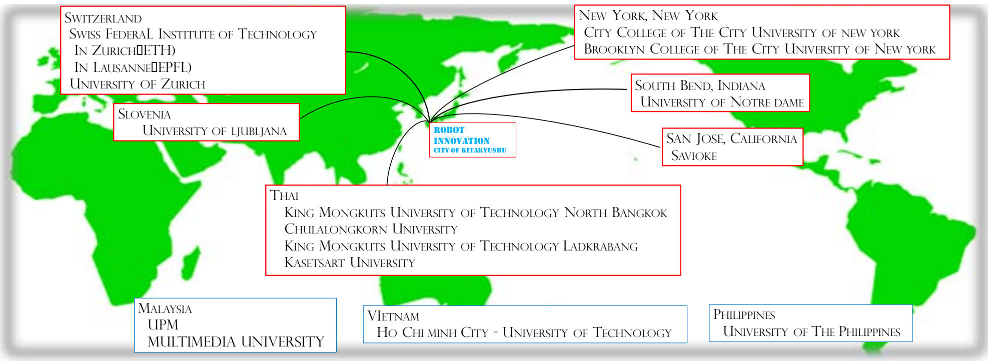


Fig. 2. Global cooperation

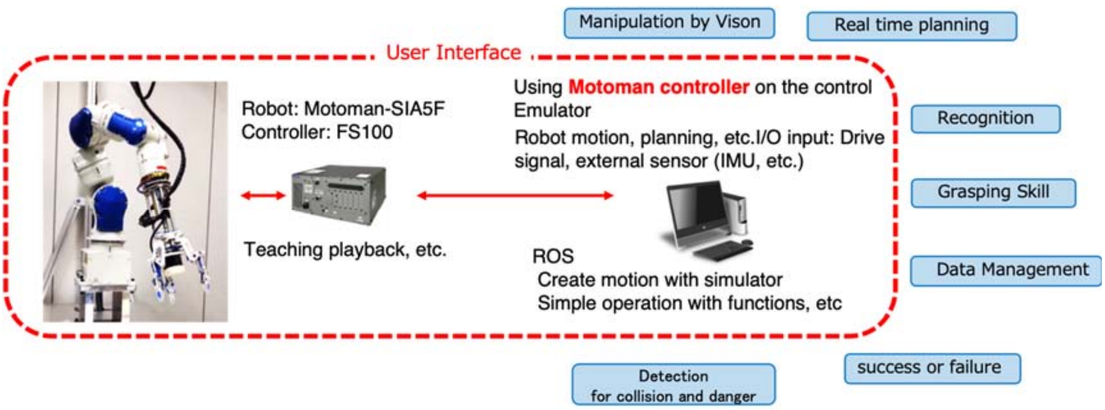


Fig. 3. Example of system in Teaching less

Body and Brain Training with Big Data and AI

Henrik Hautop Lund, Yan-Xin Liu, Massimiliano Leggieri
Center for Playware, Technical University of Denmark, Building 326,
2800 Kgs. Lyngby, Denmark
E-mail: hhl@playware.dtu.dk, maxleggieri@gmail.com
www.dtu.dk

Abstract

Utilizing a Big Data and AI approach, we developed a novel playful method for screening people for potential physical and cognitive shortages. The method creates a body and brain performance map for each individual, and the Big Data analysis provides a basis for automatically identifying the particular abilities, which may be underperforming in an individual. Further, several studies including randomized controlled trials with the Moto Tiles system have shown that particular Moto Tiles game play will increase performance of particular abilities, even after short-term play. Thereby, the proposed system can automatically generate personalized training protocols for the individual by selecting and providing the right Moto Tiles games for the individual to play to improve the underperforming abilities. The suitability of the method was tested in a small effect test with seniors with mild dementia at a care institution in Denmark. The results show that the seniors with dementia who were screened to be at high risk of falling, within the short period of training with the automatically generated personalized protocol increased their skills to no longer be at risk of falling.

Keywords: Playware, cognition, aging, game.

1. Introduction

With the demographic development, the population worldwide is changing towards having a composition of more senior citizens. In Japan in 2016, 27.3% of the population was aged over 65 years old [1], and many countries are moving towards a similar demography with many seniors [2].

As the percentage of seniors is rising, the absolute number of age-related health issues in the population will also rise. This is a challenge both for the individual well-being, and for the society economics. The increased public expenditure to age-related health issues may become a major expenditure for the society, unless some solutions are found to reduce these age-related health issues or to reduce the costs related to these issues. Some of the most common health issues, which will increase with the increased number of seniors include hypertension, coronary heart disease, type 2 diabetes, osteoarthritis, falls-related injuries, stroke, and dementia.

Technological development may help confronting this challenge. This can be achieved by combining a number of recent advances. On one side, playware development provide the possibility to create intelligent hardware and software that mediates playful actions among users of all ages [3-4]. The advantage of providing such playware within health may be that users will be intrinsically motivated to perform actions with the playware, and these intrinsically motivated actions may be of value to their health. Secondly, as a playware health technology, the Moto Tiles [5] were developed and proved to significantly improve the balance of older adults with less training period than ordinary training [6-8]. Also, Liu et al. [9] presented a Moto Tiles training regime, which could enhance cognitive abilities such as the reaction speed of the seniors. Thirdly, the development of the field of Big Data may allow the development of an AI expert system, which can potentially recognize health patterns of the population and the individual. In this work, we will combine the playware technology, health technology, Big Data and AI to provide a system, which can screen for specific health related challenges for the individual senior, and at the same time automatically generate a personalized playful training for the individual to improve on those identified challenges.

2. Big Data approach for Body & Brain Age Test

As described in Liu et al. [10], we developed a Moto Tiles Body & Brain Age Test based on four Moto Tiles games, each evaluating some body or brain abilities. Each of these games called Color Race, Special One, Final Countdown, and Remember (Simon Says in a newer version) lasts for 30 seconds making the test last for just 2 minutes in total. The physical layout of the Moto Tiles is 6 tiles in a horseshoe shape as shown in Fig. 1.

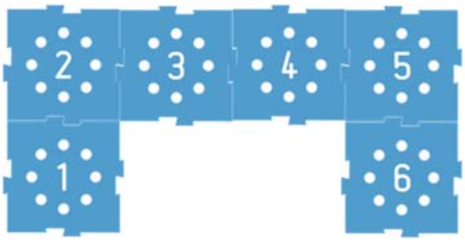


Fig. 1. The Moto Tiles layout for the Body & Brain Age Test.

Expanding on the approach from Liu et al. [10], using a Big Data approach, we collected data from 379 individuals aged 4-97 and analyzed the data. As an example, the collected data from the Moto Tiles Body & Brain Test can be visualized as in Figure 2 for the four test games.

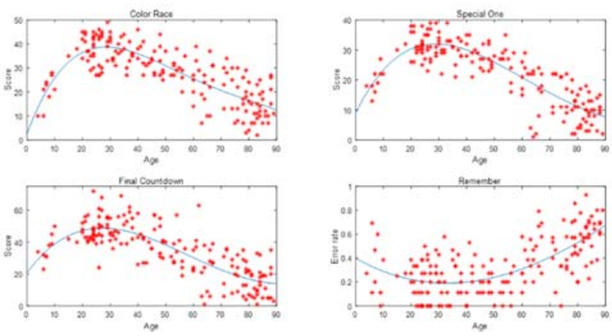


Fig. 2. The age-score model.

Analyzing the data, we can obtain the normative scores of different ages, e.g. as the average score in a game for each particular age, or by polynomial fitting using the procedure described in Liu et al. [10]. The line in Fig. 2 shows the polynomial fitting for each of the four games in the Moto Tiles Body & Brain Test.

3. Falls Risk Screening

Previous work (Liu et al. [11]) has shown that there is a strong correlation between Moto Tiles game tests and standard health tests such as Timed Up & Go (TUG) [12], Chair Stand (CS) [13], and Four Square Step Test (FSST) [14]. The TUG test measures the time that a subject takes to rise from a chair, walk three meters, turn around, walk back to the chair, and sit down. Typical cut point of the TUG test is 14 seconds [12]. The CS test measures the time that a subject takes to stand up from a chair and sit down five times. The cut point of the CS test for fall risk is 15 seconds [15]. In the FSST test, subjects are required to step as fast as possible into each square with both feet in a sequence clockwise and counterclockwise. The cut point of the FSST test for fall risk is 15 seconds [14].

Based on the data of the 51 participants (avg. 80.1 years old) who took both the Moto Tiles physical tests and Body & Brain Age Tests in Liu et al [10-11] tests, we analyzed the correlation between their performances of the two tests, which is shown in Table 1. It can be seen that the Color Race game had strong correlations with the TUG and FSST test ($|r_s| > 0.7$), and the Special One game also had a strong correlation with the FSST test. All the rest cases were moderate correlation.

Table 1. Spearman's coefficients between the physical tests and the Moto Tile games.

Correlation	Color Race	Special One	Final Countdown
TUG	-0.70611	-0.60695	-0.61745
CS	-0.51659	-0.42619	-0.45566
FSST	-0.80849	-0.70577	-0.65266

The cut points of the TUG, CS, and FSST tests for fall risk assessment are 14s, 15s, and 15s respectively. Linear regression models of the games and tests were built in order to transfer the standardized cut points to the games. Table 2 shows the cut points of the three games calculated from the three tests. The results calculated from different physical tests are consistent with each other, which indicates a high reliability of the games. By taking the average of the results, we could recommend the following Moto Tiles game cut points: Color Race = 19; Special One = 13; Final Countdown = 20.

Table 2. The cut points of the Moto Tile games

Cut point	Color Race	Special One	Final Countdown
TUG	19 (18.56)	13 (12.58)	20 (19.75)
CS	19 (19.26)	13 (12.92)	20 (20.32)
FSST	19 (18.69)	13 (12.66)	20 (19.81)

With these correlations verified, the perspective of the Moto Tiles Body & Brain Age Test is that this test can be used as a screening of the population. In this case, with the correlation with the well-established cut-points for falls risk, the Moto Tiles Body & Brain Test is a valid detector of fall risks. In this case, the scoring in the Moto Tiles games Color Race < 19 points, or Special One < 13 points, or Final Countdown <20 points indicates a high risk of multiple falls.

4. Physical and Cognitive Abilities

After being able to perform test and risk analysis with the Moto Tiles Body & Brain Age Test, it is possible to develop a system, which automatically recommend and generates personalized training protocols.

Firstly, the age-score model is integrated into the Moto Tiles Body & Brain Age Test app, which allows users to compare their scores with the normative score at their age. The comparison result is presented as a percentage difference as follows:

$$PD = (score - nv)/nv * 100\%$$

Where *PD* is the percentage difference and *nv* is the normative value.

In the Moto Tiles Body & Brain Age Test, four games are employed and the test results are presented as a spider web with 7 abilities (physical and cognitive abilities). Based on the relation between the abilities and the games, the ability scores (0-100) are calculated as follows:

$$Speed = PD_{FC} + 80$$
$$Agility = PD_{CR} + 80$$
$$Endurance = 100 - 10 * (FCscore_{2-9} - FCscore_{23-30})$$
$$Balance = (PD_{CR} + PD_{SO} + PD_{FC}) / 3 + 80$$
$$Memory = PD_{Remember} + 80$$
$$Concentration = PD_{SO} + 80$$
$$Orientation = PD_{CR} + PD_{SO} + 80$$

The algorithm maps the normative values to 80 points. Note that the endurance score is not calculated from the age-score model. Instead it measures if a player performs worse at the end of the game (23-30 seconds) than at the beginning (2-9 seconds) of the game.

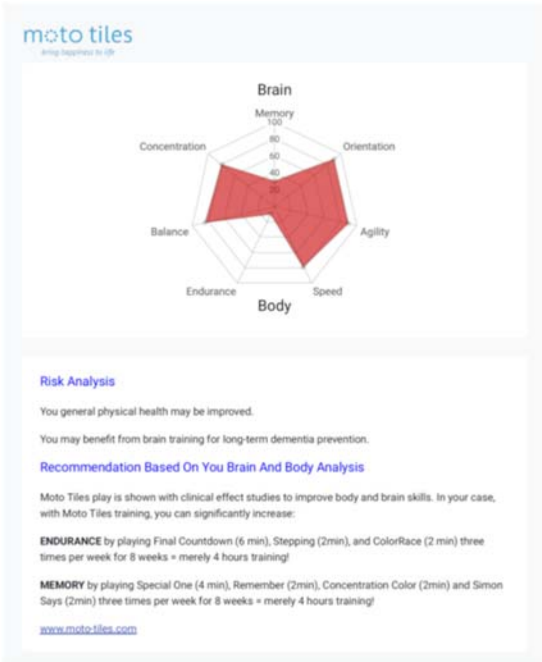


Fig. 3. Screen-shot of the spider web showing the user's score on the seven physical and cognitive abilities.

5. Automatically generated personalized protocol

The personalized training protocol is generated based on the spider web. Seven games are chosen for the protocol: Color Race, Special One, Final Countdown, Reach, Hit the Target, Simon Says, and Concentration. In order to choose appropriate games for the training protocol, each game is firstly assigned with a weight according to the spider web:

$$w_{cr} = 400 - (speed + agility + balance + orientation)$$
$$w_{fc} = 200 - (speed + endurance)$$
$$w_{htt} = 200 - (speed + agility)$$
$$w_{reach} = 200 - (agility + balance)$$

$$w_{so} = 300 - (orientation + concentration + balance)$$
$$w_{ss} = 200 - (memory * 2)$$
$$w_{conc} = 200 - (memory + concentration)$$

The weights of most games are associated with the sum of two abilities or a double of one ability, except for Color Race with four and Special One with three. This is because Color Race and Special One are two fundamental games which are relatively easy to be understood, and they therefore have more chance to appear in the training protocol.

Then the number of games *N* and total training time *T* are defined. In the first trial, *N = 2* and *T = 7* minutes. In later weeks of the training, N is raised to 3. The first N games with greatest weights are selected. The training lengths of the selected games are calculated as below:

$$T_i = \frac{w_i}{\sum_{j=1}^{N} w_j} * T$$

6. Discussion

The Moto Tiles games can be used as health tests for the seniors, and that scoring Color Race < 19 points, or Special One < 13 points, or Final Countdown <20 points indicates a high risk of multiple falls. Further, it shows that the Moto Tiles game is a very consistent and robust test.

This means that just by seniors playing and having fun on the Moto Tiles, it is possible to automatically at the same time obtain an indication of their health status and their risk of falls. Further, other scientific effect studies show how balancing skills are improved dramatically while playing on the Moto Tiles, so the Moto Tiles system also presents a tool to lower such risk of falls by playing certain games on the Moto Tiles.

Additionally, the health tests may show other indications. For instance, FSST times of 11 seconds are able to differentiate between persons with chronic stroke and healthy adults older than 50 years, so this can also be recognized simply by playing on the Moto Tiles. Similarly, for CS a score >13.6 seconds is associated with increased disability and morbidity (Guralnik et al, 2000), and metaanalysis results “demonstrated that individuals with times for 5 repetitions of this test exceeding the following can be considered to have worse than average performance” (Bohannon, 2006):

60 - 69 y/o 11.4 s; 70 - 79 y/o 12.6 s; 80 - 89 y/o 14.8 s

Also, TUG is a predictor for global mortality (also after presence of cardiovascular disease was taken into account). Hence, different cut-off points in the Moto Tiles games can be used to show various health related issues (diagnoses) when people are simply playing the games and are having fun.

References

1. *Japan Statistical Yearbook 2018*, Statistics Bureau, Ministry of Internal Affairs and Communications, 2018.
2. D.E. Bloom, A. Boersch-Supan, P. McGee, A. Seike, Population aging: facts, challenges, and responses, *Benefits and compensation International*, 41 (2011) 22.
3. H. H. Lund, and C. Jessen, “Playware - Intelligent technology for children’s play,” Technical Report TR-2005-1, June, Maersk Institute, University of Southern Denmark, 2005.
4. H.H. Lund, T. Klitbo, and C. Jessen, “Playware Technology for Physically Activating Play”, Artificial Life and Robotics Journal, 9:4, 165-174, 2005
5. http://www.moto-tiles.com/
6. H.H. Lund, Play for the elderly-effect studies of playful technology, in *International Conference on Human Aspects of IT for the Aged Population* (2015), pp. 500-511.
7. H.H. Lund, J.D. Jessen, Effects of short-term training of community-dwelling elderly with modular interactive tiles, *GAMES FOR HEALTH: Research, Development, and Clinical Applications*, 3 (2014) 277-283.
8. H.H. Lund, J.D. Jessen, Effect of playful balancing training-a pilot randomized controlled trial, *2013 ISAROB*, (2013).
9. Y. Liu, H.H. Lund, L. Wu, Playful Cognitive Training with Physical Interactive Tiles for Elderly, in *2018 International Conference on Information and Communication Technology Robotics (ICT-ROBOT)* (2018), pp. 1-4.
10. Y. Liu, M. Leggieri, H. H. Lund, Playful Body and Brain Test with the Moto Tiles, in Proceedings of ICAROB 2019, Beppu, Oita, Japan, 2019
11. Y. Liu, M. Leggieri, H. H. Lund, Electronic Measurement and Gamification of Balance Tests, in Proceedings of ICAROB 2019, Beppu, Oita, Japan, 2019
12. A. Shumway-Cook, S. Brauer, and M. Woollacott, "Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go Test," *Physical therapy,* vol. 80, pp. 896-903, 2000.
13. .J. M. Guralnik, E. M. Simonsick, L. Ferrucci, R. J. Glynn, L. F. Berkman, D. G. Blazer, *et al.*, "A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission," *Journal of gerontology,* vol. 49, pp. M85-M94, 1994.
14. W. Dite and V. A. Temple, "A clinical test of stepping and change of direction to identify multiple falling older adults," *Archives of physical medicine and rehabilitation,* vol. 83, pp. 1566-1571, 2002. 15. S. L. Whitney, D. M. Wrisley, G. F. Marchetti, M. A. Gee, M. S. Redfern, and J. M. Furman, "Clinical measurement of sit-to-stand performance in people with balance disorders: validity of data for the Five-Times-Sit-to-Stand Test," *Physical therapy,* vol. 85, pp. 1034-1045, 2005.

Body and Brain Training with Big Data and AI 2 – A Pilot Test of Falls Prevention

Henrik Hautop Lund, Yan-Xin Liu, Massimiliano Leggieri

Center for Playware, Technical University of Denmark, Building 326,
2800 Kgs. Lyngby, Denmark
E-mail: hhl@playware.dtu.dk, maxleggieri@gmail.com
www.dtu.dk

Abstract

The paper presents the pilot test of the Moto Tiles Body & Brain Test combined with automatically generated individualized training protocols. The test is performed with seniors (avg. 80 years old) with mild dementia in senior care centers in Roskilde, Denmark. The seniors performed pre- and posttest on tests correlated with timed-up-and-go (TUG), chair-to-stand (CS), and four square stepping test (FSST). During the intervention, the dementia-affected seniors performed their individualized training on the Moto Tiles games over 3 weeks with an average total of 9 sessions of 7 minutes each. On average, the seniors improve 41%, 31%, 42% and 3% on the four tests. In terms of cut-points for multiple falls risk, the seniors are at high risk at pre-test, while they are at low risk at post-test. This indicates that the very short time (1h over 3 weeks) playful training on Moto Tiles can potentially bring seniors from high risk of falling to a much lower risk of falling. The present study is a small pilot test, and results should be verified in randomized controlled trial(s).

Keywords: Playware, cognition, aging, game.

1. Introduction

Playful training can have a dramatic and fast effect on the health of seniors, as has been shown in [1-4]. The high effect of play on the senior health seems to arise as a collateral health effect of the play. Play acts as an intrinsic motivation, and a playware tool such as the Moto Tiles mediates a playful physical interaction of the seniors.

Play is a free and voluntary activity that we do for no other purpose than play itself. When in play, we can often do more than we would normally be able to do, because we get into a play dynamics in which we forget about time and place. In the play dynamics, we forget about our fears and limitations. A play force – such as the playware tool Moto Tiles – pushes the seniors into the play dymanics, and when fear and limitations are forgotten, the seniors are performing playful movements that they would otherwise seldom make. This seems to be the qualititative explanation of the high and fast effects seniors experience on physical functional abilities and cognitive abilities when playing with the Moto Tiles.

Also, as described in part 1 [5]-ibid, the play on the Moto Tiles can be used as a valid senior health screening tool, e.g. for falls risks. Falls is a major concern for seniors health. In the United States, about 40% of the population age ≥65 living at home will fall at least once each year, and the rates are much higher for persons aged >75 years and persons living in long-term care institutions [6]. Part 1 [5] showed that there are validated cut-points for multiple falls risks in the scores on the Moto Tiles games. As mentioned there, "the scoring in the Moto Tiles games Color Race < 19 points, or Special One < 13 points, or Final Countdown <20 points indicates a high risk of multiple falls." [5]

Importantly, the screening of seniors health by playing the Moto Tiles games not only provides a mapping of the individual physical and cognitive abilities, but can also provide an automatically generated personalized training protocol to follow after the screening, so that the right kind of abilities are addressed in the play on the Moto Tiles for each individual senior.

In this work, we will look at a first pilot test of such screening and playful Moto Tiles training with automatically generated personalized training protocols.

2. Intervention

The pilot test was performed with 12 seniors with mild dementia from Seniorhøjskolen in Roskilde, Denmark. They were average 80 years old. They were pre- and posttested with the Moto Tiles Body & Brain Age Test (see [5]), and had an intervention of 3 weeks, during which they played average 9 sessions of Moto Tiles training. The games and duration of games were automatically generated for each individual senior. They performed 3 cycles of the loop in Figure 1.

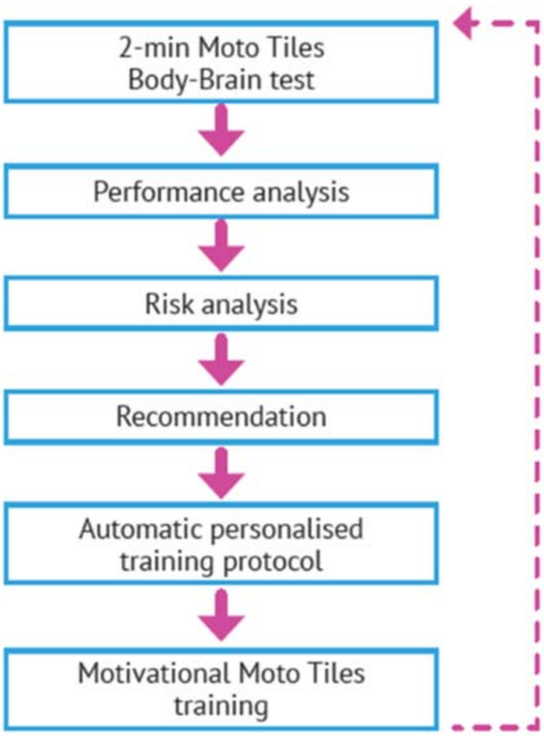


Fig. 1. The intervention loop with the Moto Tiles Body & Brain Age Test, risk analysis, and automatically personalized training protocol. Seniors perform the Moto Tiles training for a number of sessions before they again take the Moto Tiles Body & Brain Age Test.

Accordingly, first each senior performed the 2 minutes Moto Tiles Body & Brain Age Test composed of four test games of 30 seconds each. Then the system makes a performance and risk analysis based on the score in the four games. The analyses is presented to the senior, and the system uses the analysis to provide recommendation and to automatically create a personalized training protocol for the senior (see [5] ibid). Then the senior performs a number of training sessions with this protocol. In the present pilot test, the senior performs 3 sessions over the following week. After those sessions, the senior will again perform the 2 minutes Moto Tiles Body & Brain Age Test, and a new personalized training protocol is generated and will be used for the sessions in the following week. And so the cycle continues (see Fig. 1).

3. Results

Table 1 shows the scores of the 7 participants who performed all tests (pre-test, mid-test, post-test). It shows the average score on each of the four test games and the percentage change from the pre-test. As can be seen, on average the seniors improve their test scores with 41%, 31%, 42%, and 3% on the four test games.

Table 1. Score on the test games based on the 7 participants who completed all the three tests.

	Color Race		
Avg	13.6	14	19.1
Changes		0.4	5.6
Changes (%)		3.2	41

	Special One		
Avg	10.6	10.6	13.9
Changes		0	3.3
Changes (%)		0	31

	Final Countdown		
Avg	13.3	13.4	18.9
Changes		0.1	5.6
Changes (%)		1	42

	Remember error rate		
Avg	0.6	0.5	0.5
Changes		-0.02	-0.02
Changes (%)		-3	-3

Also, in terms of the actual score on the test games, it can be seen that in pre-tests, the seniors score well below the cut-point for multiple falls risk, i.e. Color Race < 19, Special One < 13, Final Countdown < 20, so they are at high risk of falling. At the post-test, in both Color Race and Special One, the seniors are scoring above the cut-point, indicating that they have improved their skills to a level where they are no longer at the high risk of falling.

In order to better understand how the personalized training protocol is generated automatically for the individual senior based upon the score in the 2 minutes Moto Tiles Body & Brain Age Test, let us take a look at two example.

The first example is a male aged 67 years of age. His test scores are shown in Table 2. At the first test, he scores 25, 17, 26 and 0.556 on the four test games. Based on the performance in the first test, the spider web showing his score on the different physical and cognitive abilities is generated – see the blue dots in Figure 2.

As can be seen, he scores comparably low in the memory ability and a little low on the speed ability. This results in the automatically generated personalized training protocol being generated to become as follows for the first week of training:

First training protocol: Final Countdown * 4, Simon Says *3

Indeed, these two games put specific focus on speed and memory, and he would be playing these 4 minutes Final Countdown and 3 minutes Simon Says at each session over the first week of training.

Table 2. Score on the test games for a male participant.

	Test 1	Test 2	Test 3
Color Race	25	25	32
Special One	17	17	18
Final Countdown	26	24	42
Remember	0.556	0.333	0.385

Then, at the second test he scores 25, 17, 24, 0.333 (Table 2), and the red dots in Figure 2 shows how he has improved in the memory ability. So the new automatically generated personalized training protocol for the following week puts more focus on speed, agility and balancing:

Second training protocol: Color Race * 2, Final Countdown * 4, Special One * 1;

At the third test, he scores 32, 18, 42, 0.385 (Table 2), and as can be seen on the spider web (Figure 2), he scores very well in all abilities.

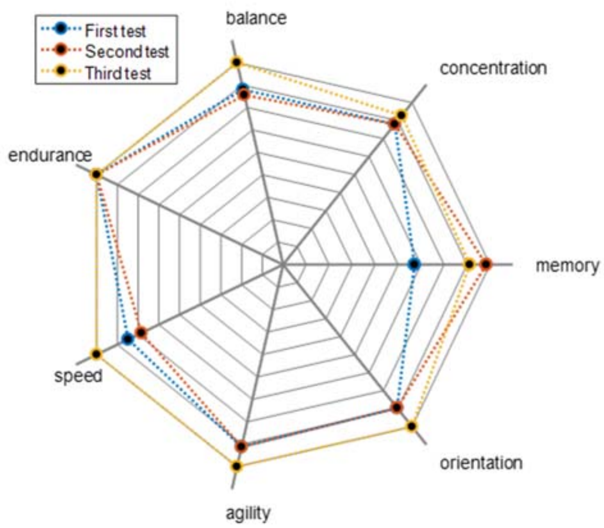


Fig. 2. The performance of a male participant in three tests.

As a second example, let us look at a female participant aged 65 years of age. Her test scores are shown in Table 3. In the first test, she scores low with 13, 10, 4, 0.778 in the four tests. Figure 3 shows the performance of the first test with the blue dots in the spider web. As can be seen, the balancing, speed and memory abilities are very low. So the first training protocol is generated to put focus on games training these abilities in combination:

First training protocol: Color Race * 4, Special One *3

Table 3. Score on the test games for a female participant.

	Test 1	Test 2	Test 3
Color Race	13	12	16
Special One	10	9	12
Final Countdown	4	5	9
Remember	0.778	0.563	0.5

In the second test, she scores 12, 9, 5, 0.563, so the memory ability has improved (see Figure 3), and therefore the next automatically generated personalized training protocol puts less emphasis on games for memory abilities, and more focus on games for balancing, speed and agility:

Second training protocol: Color Race * 3, Final Countdown * 2, Special One * 2;

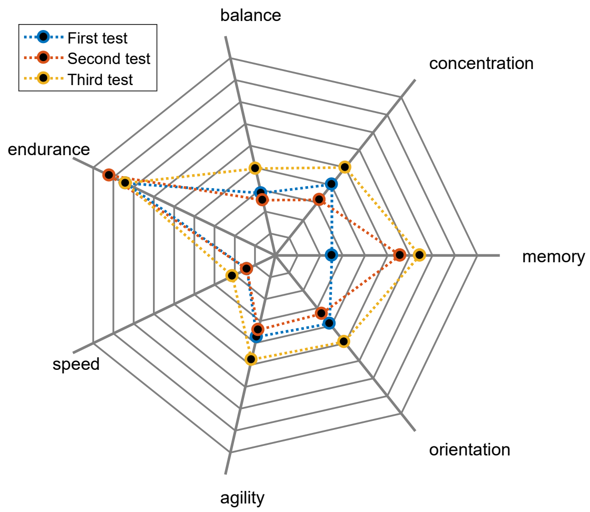


Fig. 3. The performance of a female participant in three tests.

6. Discussion

As has been shown, the Moto Tiles games can be used as health tests for the seniors, and scoring Color Race < 19 points, or Special One < 13 points, or Final Countdown <20 points indicates a high risk of multiple falls. In the case of the pilot tests, the average score of the seniors after training with the personalized protocols took the average from indicating high risk of falling in the Color Race and Special One score to a level above the cut-points indicating lower risk of falling.

In the qualitative evaluation based on interview, the therapist evaluating the pilot test expressed that "the customized training is perfect because it targets to the individual weakness".

It is observed that in the Remember game, there were some participants who got a higher error rate. This can be because their memory did not change over the short intervention period (still can not remember many colors), but their mobility was improved, so they could have more guesses (but most were wrong), and this resulted in a higher error rate. In future, we will be using another memory game (Simon Says).

The high increase, which is noticed (31-42%) may contain some test-retest effect. Also, the pilot test did not include a control group to compare. Hence, one should view the present tests merely as a pilot test. Randomized controlled trial(s) is needed to confirm the indications of the pilot test.

Acknowledgements

The authors would like to thank all participants of this experiment. Part of the work was produced as part of the EU Horizon 2020 project REACH under grant agreement no. 690425.

References

1. H.H. Lund, Play for the elderly-effect studies of playful technology, in International Conference on Human Aspects of IT for the Aged Population (2015), pp. 500-511.
2. H.H. Lund, J.D. Jessen, Effects of short-term training of community-dwelling elderly with modular interactive tiles, GAMES FOR HEALTH: Research, Development, and Clinical Applications, 3 (2014) 277-283.
3. H.H. Lund, J.D. Jessen, Effect of playful balancing training-a pilot randomized controlled trial, 2013 ISAROB, (2013).
4. Y. Liu, H.H. Lund, L. Wu, Playful Cognitive Training with Physical Interactive Tiles for Elderly, in 2018 International Conference on Information and Communication Technology Robotics (ICT-ROBOT) (2018), pp. 1-4. [4]
5. H.H. Lund, Y. Liu, M. Leggieri, Body and Brain Training with Big Data and AI. In Proceedings of ICAROB 2020, Beppu, Oita, Japan. Ibid.
6. L.Z. Rubinstein, L. Z. Falls in older people: epidemiology, risk factors and strategies for prevention. Age and Ageing 2006; 35-S2: ii37–ii41.

Playware Ball – Initial Development Impressions of an Intelligent Ball

Morten Roed Frederiksen, Massimiliano Leggieri, Henrik Hautop Lund
Center for Playware, Technical University of Denmark, Building 326, 2800 Kgs. Lyngby, Denmark

E-mail: hhl@playware.dtu.dk
www.playware.elektro.dtu.dk

Abstract

This paper investigates the use of barometric pressure sensors in combination with accelerometer data as a viable way to track the activity level of players in a fast-paced exercise game using a football to motivate the participants. A fitness game was developed along with a prototype intelligent ball with the aim of tracking both the skill level and the level of activity for the users. The initial findings indicate that the sensor fusion of barometric pressure data and accelerometer data proves reliable in the fast-paced setting of football games. Furthermore, the constructed prototype foam exercise ball proved a reliable container to house both the sensor and computational devices in a safe manner and could be useful for further tests on other aspects of combining gaming and physical exercise.

Keywords: Playful technology, Playware, Personal Health Technology, Adaptive Games, Football.

1. Introduction

Increasing the number of daily physical leisure activities performed by the elderly in nursing homes could prevent several diseases including dementia [1]. It is vital that no one spends too much time sitting down as research shows inactivity has a dramatic impact on potential risk factors such as cardiovascular disease, type 2 diabetes, metabolic syndrome, and obesity [2]. The simple solution to avoid those risk factors is to increase the daily level of activity. However, there is resistance from some of the elderly to sustain a healthy balance of daily routines with enough exercise as some psychological barriers hinder the elderly to join in. The natural physical limitations of old age can sometimes make them feel inadequate and prevent them from wanting to exercise. The mundane nature of repeating exercises can also be a factor that prevents them from raising their activity levels.

In an attempt to increase the overall activity level of the elderly generation in a fun and accessible manner, we developed playware [3, 4] focusing on play with intelligent artefacts to engage the users. Indeed, other playware such as Moto Tiles have shown to have a significant effect on the physical health of seniors [5-7]. According to the Playware ABC [8, 9], the hardware was developed to be usable in multiple contexts spanning various usage scenarios and developed for different user types with a large variety in their amount of previous experience with digital entertainment (see [10] for other playful rehabilitation use). In this work, the hardware was developed as a semi-autonomous robotic toy that could be self-explanatory in its intended use and that would require little to no instructions before being adopted. As a way to measure the activity and skill level of the users of the toy ball, the ball was equipped with a barometric pressure sensor and an accelerometer to provide reliable data for performing adequate movement analysis. The robotic exercise ball was developed to be expressive and fun. It was not developed using any specific expressive robotic models but analyzing it after its constructions phase using the MOAM model of affective expression modalities results in a point distribution of 3 for the sound modality which provides a coordinated effort to respond to user input [11].

The playware ball was preliminarily tested under different context circumstances to investigate how well the sensor combination of barometric pressure sensors and accelerometer sensors functioned with different pressure changes and different types of user handling. The form factor and sensor choice for the exercise tool proved a sensible combination usable for a large variety of purposes.

2. Physical construction and software setup

The physical object was a 20cm diameter, foam-body, semi-autonomous exercise ball. The outer layer of foam was selected for the following main reasons:

- It contained a penetrable surface that allowed air to travel freely to and from the barometric sensor so that it would provide robust readings.
- The foam protected the sensors in a flexible enclosure.
- The elastic body of the foam enclosure worked as a shock absorber to prevent damage as the ball would receive hard kicks during exercise games.
- The foam ball invited energetic activities as its morphology would resemble a non-harmful toy that you were not afraid of kicking around with the risk of hitting indoor furniture or other people in the vicinity.

Fig 1. The Playware Ball containing the electronics inside.

Inside the foam outer layer was a small embedded computer platform. The main system on a chip was an ESP8266 based platform with onboard wifi to provide connectivity and data transfer to and from a mobile application. To the SOC, the accelerometer and barometric pressure sensor were attached. To facilitate communication with the Arduino based ESP8266 platform, the I2C communication protocol was used for both sensors. To power the device a dual cell lithium-ion battery was used. Sufficient charging ports were added to allow a swift and easy charging process between the exercise games. The constructed morphology aimed to entice physical activity and to present the hardware in a reliable package. Figure 1 shows the complete form factor of the exercise ball.

The sensor processing and game initiation controls were controlled by a companion mobile application. The game would connect to the ball via a shared wifi connection and would receive sensor data with a refresh rate of 200hz. The refresh rate was chosen as the best compromise between high granularity of sensor measurements and high power consumption for both transmitting the data as well as post-processing the data on a mobile device. Figure 2 shows the application's main window and activity selection screen. To make sure the height estimations were as accurate as possible, the game would auto-calibrate the height measurement base of the barometric pressure at the start of each physical activity game. The software would assume that the ball was placed on the ground as the game was initiated to establish the lowest possible readings within each game. To process the data the following post-processing filters were applied to the stream of incoming data. All data points were averaged and the current height estimation was only affected through a low pass filter with high perseverance of the existing value. To focus the data measurements and filter unnecessary readings from the stream, the accelerometer data was used to detect when the ball was kicked and when the ball was descending towards the ground with detected expected gravity speeds. The accelerometer data was also employed to detect when people kicked the ball and whether the ball was moving or was kept still. The audio-based expression modality of the physical setup was also provided through the app and the audio would automatically select speaker mode and increase the volume to ensure the users could hear the intended audio cue throughout the exercise sessions.

3. Fun in the game

The games developed for the exercise ball was created with the intention of being easily understandable, fun to play and also fun to watch. The games should be explainable in a single sentence to ensure the users maintain a flow in the exercise routines by avoiding that anyone would stop in the middle of an exercise confused by the rules of the game. The games were challenging and fun as a means to keep the users coming back for repeated exercise sessions. The reason for making the games entertaining to watch as well was done in an effort to entice new players to join in.

From a health perspective, the games focused on two different types of activities with an emphasis on different body aspects to strengthen:

- The cardiovascular system.
- The fine muscle movement and eye-leg coordination.

Fig 2. The interface on the tablet when playing the rehab game for upper extremity training with two playware modules.

The first game, which emphasized cardio strengthening exercises, was a game that enticed the participants to keep moving at a steady pace throughout the exercise session. The premise of the game was that the exercise ball would sense whether it was moving or standing still and communicated those different states to the users through each exercise session. To force the users to move during the game, the software would also keep track of each time the ball was kicked sufficiently hard. A hard kick meant that the users would have to chase after the ball and prompt further cardio strengthening activities. If a certain playing time had passed with no one kicking the ball (e.g. 4 seconds), the ball would play an explosion sound to convey that the game had ended. The premise was to keep the ball moving for as long as possible. High-scores would be gained if you could top the previous players' overall exercise time.

The second game attempted to improve the finer muscle movements in the user's legs and to strengthen the eye to leg coordination. The premise and goal of the game was to keep the ball in the air as much time as possible and avoid it hitting the ground. The sensors would register whether the ball was kept still, to make sure the users were not resting the ball on their feet during an exercise session. The barometric pressure was used to ensure that the ball would not stay beneath a threshold user-adjustable height for too long. If that happened, a sound would convey to the user that the game had ended.

4. Prototype verification

The developed prototype hardware and game software were tested by 5 faculty staff members of the playware lab at the Technical University of Denmark. The early testing with the foam ball indicated that the users often accidentally stepped directly on the critical center parts of the construction. To ensure this part would work in extended exercise sessions we attempted to perform tests where we put some physical strain on the equipment. This included the following tests:

- Hard-kick tests, in which we performed multiple subsequent hard kicks directly aimed at the center mass of the ball where the electronics were housed.
- Step-on tests, in which we stepped on the ball from several approaching angles.
- Drop tests, in which we dropped the ball from different high altitudes between first to the third floor of the university building.

It was important to make sure the ball would work across several different contexts. The final product was aimed at working both outside and inside, in high humidity, and warm and cold weather. The foam enclosure needed for the barometric pressure sensor to work excluded the construction from working in scenarios where it would get wet so it was only tested in dry conditions. To ensure the barometric sensor worked in less than ideal conditions, we attempted to introduce some sensor noise to occlude the sensor data during test exercise games. This included changing the local pressure in the room by opening a window during a game session and by moving the ball while it was calibrating the low height measure during game initiation. Different heights were tested and a different pace of height changes was tested to investigate the limits of these sensors in this specific setting.

Overall the combination of barometric pressure and accelerometer sensor data worked very well to determine the different states needed in the exercise games. They determined the altitude and movement amount with a high enough confidence level to be usable even in this fast-paced exercise context. The sensors often showed some erratic behavior where it would return outlier data such as jumping to a very high or low altitude level during the games but using basic filtering methods these wrong measurements were easily discarded. Once in a while, the local pressure would change in a given room. This changed the minimum height threshold defined in the game. By changing the calibration to occur at the start of each game we countered this effect with some success. However, this required the users to keep the ball still at the ground at the start of each game, which sometimes proved difficult as the test users had a hard time leaving the ball alone. We see that as a positive indication that the chosen morphology invites physical interaction.

5. Discussion - improvements

Our testing phase showed that the next prototype could benefit from including more onboard expressive features. This includes both more expressive features employed mid-game and also pre-game explanations and during game selection. Although the mobile application is a great platform to deliver written instructions and to provide processing power to analyze the sensor data, using it subtracts from the experience of having an intelligent ball that works alone without the need for tethered connections.

The expressive sound used throughout the exercise games makes the ball seem alive and should be coming from the ball, not the app. It was a clear indication through the tests that the users wanted to forget the app and just focus on the ball. This resulted in users kicking the ball to a distance about ten meters from the app, at which point it proved difficult for the data connection to remain active and which also made it difficult for the users to hear any game-related audio cues coming from the app. Moving to a standalone platform in the ball would also make it easier for people to interpret the exercise ball in an anthropomorphic way which might improve how people interacted with it. If people would see it as a viable partner during the exercises it could make them more eager to exercise alone with the ball. However, moving the electronics and adding necessary processing power to facilitate a standalone decide might be difficult as doing so will be a tradeoff between having the robustness and reliability of using simple electronics inside the ball vs using a more complex setup using powered speakers and more processing.

In the preliminary testing phase, there was a tendency for people to view the playware ball as an intelligent entity. (E.g. some people expressed that the ball tried to trick them). This anthropomorphic interpretation may be utilized in further prototypes. This could maybe amount to making the ball laugh when people kick it or otherwise taunt them to interact with it. The responsive nature of the interactive exercise ball seemed to fit the scenario of enticing exercise well and it made people see the ball as a tool that was both energetic and alive.

Further improvements in future prototypes might also benefit from including better haptic feedback to the users. This is needed to make the ball an untethered standalone device. The haptic feedback could provide a way to provide feedback when the users initiate games or attempt to stop current activities. E.g. it might vibrate to give feedback or the users could shake it to change activity. Our current knowledge from developing this and other playware devices tell us the more simplistic we make the device, the more it will entice people to use it.

Using an app with the ball might also make people scared of trying it as the mobile device itself represents high levels of complexity and demands a steep learning curve from people who are not already familiar with smart devices. In development phases for new technology, we cannot per default assume that the elderly generation is comfortable with using mobile platforms or computers. Our preliminary testing shows that simplicity might be a viable way to appeal to a large variety of ages - when combined with a device that resembles a common object such as a football it makes it easy for new users to approach and start using it.

6. Conclusion

The sensors and the hardware proved reliable during our preliminary testing phases. The playware ball itself was a successful prototype that gave us knowledge on the type of material and enclosure needed to proceed in building further prototypes. The morphology proved enticing for people to pick up and start using and the expressive nature of the audio cues was interpreted as fun and informative during the exercise games. This was a preliminary investigation into the development of an intelligent piece of a playware exercise equipment to facilitate more enjoyable physical activities for the elderly generation and as such the prototype accomplished a lot with simple sensors in a complex scenario. The prototype platform managed to get reliable information on all the discrete ball states needed to provide a good experience for the test users and the expressive features of the playware ball made the ball fun to use. The platform proved a viable foundation to continue development for further testing of combining gaming and physical exercise.

References

1. J. Verghese, R. B. Lipton, M. J. Katz, C. B. Hall, C. A. Derby, G. Kuslansky, A. F. Ambrose, M. J. Sliwinski, and H. Buschke "Leisure activities and the risk of dementia in the elderly", The 2003 New England journal of medicine, volume 348.25, page 2508-16.
2. M. T. Hamilton, D. G. Hamilton, and T. W. Zderic, "Role of low energy expenditure and sitting in obesity, metabolic syndrome, type 2 diabetes, and cardiovascular disease", Diabetes Journal volume 56.11, pages 2655-67, 2007
3. H. H. Lund, and C. Jessen, "Playware - Intelligent technology for children's play," Technical Report TR-2005-1, June, Maersk Institute, University of Southern Denmark, 2005.
4. H.H. Lund, T. Klitbo, and C. Jessen, "Playware Technology for Physically Activating Play", *Artificial Life and Robotics Journal*, 9:4, 165-174, 2005
5. H. H. Lund, "Play for the Elderly - Effect Studies of Playful Technology," in Human Aspects of IT for the Aged Population. Design for Everyday Life. (LNCS Vol. 9194, pp 500-511, Springer-Verlag, 2015)
6. H. H. Lund, and J. D. Jessen, "Effects of short-term training of community-dwelling elderly with modular interactive tiles," *GAMES FOR HEALTH: Research, Development, and Clinical Applications*, 3(5), 277-283, 2014.
7. http://www.moto-tiles.com (checked: 4/12/2019)
8. H. H. Lund, H.H. "Playware ABC: Engineering Play for Everybody." *Journal of Robotics Networks and Artificial Life*, *3*(4), 2017
9. H. H. Lund. "Playware Research–Methodological Considerations". *Journal of Robotics, Networks and Artificial Life*, *1*, pp.23-27, 2014
10. H. H. Lund, M. R. Frederiksen, and M. Leggieri, "Engineering Modular Playware", in Proceedings of the 2019 International Conference on Artificial Life and Robotics. ALife Robotics Co, Ltd., Vol. 24. p. 9-13 (Proceedings of International Conference on Artificial Life and Robotics), 2019.
11. M. R. Frederiksen and K. Støy, "A systematic comparison of affective robot expression modalities", in Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, 2019

PixelBeing – An Eco-sustainable Approach to Robotics and AI

Luigi Pagliarini[1,2,3], Henrik Hautop Lund[1]

[1] Centre for Playware, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
[2] Academy of Fine Arts of Macerata, Via Berardi, 6, 405111 Macerata, Italy
[2] ISIA Design, Via C. Battisti, 198, 65123 Pescara, Italy.

luigipagliarini@gmail.com

Abstract

In this paper, we describe and define the range of possible applications and the technical contours of a robotic biotechnological system to be worn on the body for playful interactions. Moving from earlier works on Wearable and Modular Robotics, we describe how, by using modular robotics for creating wearable, it is possible to obtain a self-sustainable and flexible wearable system, consisting of freely inter-changeable input/output modules that through the use of solar, mechanical, and other sources of renewable energy are able to suit some specific tasks. Here, we drive the attention on early prototypes to show the potentialities of such an approach, and focus on depicting possible application in the future of electronics domain. Indeed, our artistic experiment is a clear example of how to scale down electronics to an eco-sustainable level, which can still create playful and useful interactions for many application domains.

Keywords: Eco- technology, Playware, User-friendly robotics, Artifcial Intelligence, Sustainable Technologies, Eco-friendly Hight-Tech.

1. Introduction

The inevitable destiny of all of the future technologies is to align to the process of lessening the environmental impact both in terms of production, materials, and consumption. Therefore, electronics, computing, and robotics too should try to reduce their environmental effect from the design phase. To work on design decisions to improve sustainability, designers need clear information on how to keep the whole process manageable, sustainable and doable. As a consequence, such product design consists both of choosing between an enormous array of options and, vice versa, a very limited subset of possibilities.

Therefore, the main intention of the PixelBeing project is to design an artifact able to provide product designers with actionable insights into the main triggers of environmental impacts, so they can change their design conception to be more environmentally friendly. We do so either to support eco-design strategies and to create necessarily low-cost and eco-friendly products, while keeping in mind that any product still needs its robustness, its flexibility and the usual support provided by any developer. Our decision was to realize a prototype that, in part, is a tool for product designers with which they are able to quickly see, compare and assess design variations without needing to go through the whole experimental process, building on the playware research methodology [1]. Indeed, eco-friendly design factors, such as mass, energy use and transport volume, etc. are very many and difficult to ensemble together in the same production, and our experimentation shapes some knowledges that might provide easy entry points to make design improvements and allows designers to develop product-specific guidelines.

On one side, our research piece aspires to be part of that larger investigation that is facing a concrete challenge, a challenge that might lead towards an epochal change while, on the other hand, the PixelBeing project being a handmade art-oriented research piece (building on previous robot art investigations [2]) carries along with it a relatively low-demanding output in terms of tangible industrial production, by now, and on the opposite an extra request to fulfill a serious request in terms of innovative aesthetical results.

The PixelBeing thinking has meant trying to revisit High-Tech production of AI and Robotics artifacts under a more ethical light, looking for the best ways to analyze hotspots and opportunities in the production life cycles under an economical point of view, without sacrificing efficient, accurate and transparent results. A first example regards materials. In PixelBeing it has been a search on the most available and accessible, easy and effortless production methods that avoid the use of plastic or other toxic and high-impact materials as much as possible. Another clear example are methods. In PixelBeing we tried to minimize either the size and the value or the implant difficulty, in terms of speed and effort, of any given functional component, as well as, we focused on technologies where the lifespan of the whole structure and replacement of broken parts is an essential target. A final example regards energy. We focused our research on the most obvious energy impact of the technology in use. In PixelBeing, indeed, we mostly make use of solar panel circuits and we are actually immerged in the process of implementing mechanical energy ones in the form of kinetic energy, in which objects have the ability to do work when they move, and potential energy, in which objects have the ability to do work due to its own position.

In short, this electronic art project aims at an initial renovation of the idea of robotics so that it will be able, in the near future, to reach the design, production and consumption of objects that are ethically more adequate in the sense of eco-sustainability.

2. The PixelBeing Project

PixelBeing is a project for a theatrical character that consists of a robotic system to be worn on the body (derived from previous work [3, 4]). It is meant to produce aesthetical and playful interactions, and it is built using concepts derived from modular robotics and modular playware [5, 6]. It foresees a full body suit made of a mask, gloves, shoes and a flexible wearable processing system, where freely interchangeable input/output modules can be positioned on the body suit in accordance with the aesthetical demand and tasks, at hand. The idea is to implement both a virtual (sensors-based) interaction and a more physical one to reach a wider range of possible outcomes, behaviorally and aesthetically. Therefore, the basic challenge is to design a general interface that focuses on the users' body interaction with the real world, and possibly with a social environment. At the currents state we have tested the general principles and have developed the mask as well as started assembling the gloves and the body costume.

Fig. 1. The PixelBeing mask. Lateral view.

This project mainly focuses on eco-friendly technologies and processes and tries to exploit the sustainability of any electronic (art) tools. It makes use of recycled components, biomaterials and supports, solar panels and all of the possible sources of renewable energies, and it also pay attention to the principle that procedures and resources should be as natural and easy to access, as much as possible.

The basic idea is a suit that ends up to be the playfield where to arrange different mechanical, electronic and electromechanical modules according to any need. The modules we use on it can be either isolated or interconnected and should be made so that they can be easily and quickly relocated, therefore, the way the resulting suit configuration will perform a task will depend both on the modules' specific functionalities and on their physical (or their geographical) displacement on the suit itself. Modularity is essential, and that is because besides of offering a larger variety of possible configurations and activities it is a pretty eco-friendly approach since it allows the less expensive procedure when a subpart of the artifact gets out of work. Of course, such What&Where System can be applied to a wide and complex number of situations and tested on many potentials and can be used to create body interactions in several application domains.

Therefore, it becomes crucial to experiment the possible "definition" and implementation of the idea of module and of module's functionality. We framed a number of modules general characteristics as: A) Each module-circuit is fully autonomous energetically and electronically, although a circuit can be thought as 'eventually' connected physically or virtually to others modules, to a font of energy, or to any other computer interface; B) Although there can be exceptions, each single circuit is conceived independently from its final location; C) A module should be applicable to any geolocation and should be thought as for a general purpose, not limited to any single and specific application.

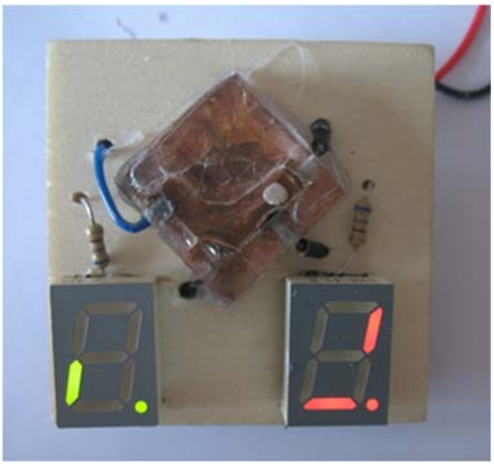


Fig. 2. A wooden (industrial production) made PixelBeing's module. First prototype, circuit with switch, sensor and led-based actuators.

As said above, modules can be thought as either isolated or interconnected, and in the latter case we think to a communication paradigm where modules can communicate: A) Locally: Neighbor-to-Neighbor communication (wired or wireless); B) Globally: From one module to a module far away (wireless); C) One to one communication system; D) one to many communication systems; E) Many to many communication system. Our research, at the moment, can be considered primitive. We aim to evolve it in many different ways and a grant and/or a team experienced in electronics, robotics, informatics and AI would be ideal to such a goal. Indeed, we wish to implement modules that have a highly significant input set (a large number of sensors, including biofeedback and neurofeedback ones), a clever input pre/processing, and a large set of articulated outputs (including motor actions).

There are few more levels of research embodied into the PixelBeing.

The first one explores the use renewable energies. We already furnished PixelBeing with accumulators modules as well as displaced on his head different solar panels modules.

The two interconnected modules are able to sustain wee electronic circuits and if perfected should be able in the early future to supply enough power to larger circuits.

We are now in the process of exploiting the use of mechanical energy modules of two different kinds. The first one that makes use of the kinetic energy produced by the body of the actor wearing the suit, and the second one which simply uses the ability of taking advantage to work thanks to gravitational due to its own position in three-dimensional space.

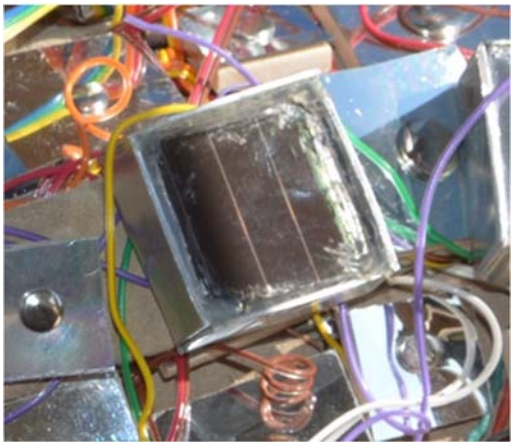


Fig. 3. Detail of the PixelBeing mask. A Solar-panels based circuit module mounted on PixelBeing's helmet

The second level of investigation is digging in the potentialities of the use of biomaterials when building and assembling modules. We are making a real effort trying to integrate electronic circuits within the most natural material. Of course, there are many elementary problems, especially related to the moisture potential, we are slowly overcoming.

Meanwhile we basically aim at reaching such a level of integration where electronics co-exists with alive materials and, probably a dream, circuits are somehow partially fed by the growth such processes like the chlorophyll photosynthesis produced by any plant.

Fig. 4. A detail of a sense-and-actuate module to be mounted on PixelBeing' suit. Later prototype. A biomaterial-based circuit made out of natural wood and moss.

A third level of investigation is the one that focuses on targeting the minimal size of the modules since, often, the smaller modules are the lesser they consume, the more adaptable is to implement the whole structure, and the cheaper to repair/replace broken units.

The fourth level of investigation regards the energy flow. Since the PixelBeing design is a design that fully relies on renewable energies it has to rethink the flow of energy in terms of appropriate modules location. The "electronics space" must be thought in profitable geometrical terms, so that acquiring and distributing energy itself has to be done in most convenient way.

3. Discussion and Conclusion

In this paper we described and defined a full range of possible implications and the technical contours for a modern robotic biotechnological system. Moving a general overview on eco-friendly design we described how, by using modular robotics for creating wearable, it is possible to approach a self-sustainable and flexible wearable system, consisting of freely inter- changeable input/output modules that through the use of renewable energy we were able to reach some important goals as well as targets, as for example introduce natural materials and renewable energy in such delicate field as wearable robotics design. We described the progresses of our first prototypes to show the potentialities of such an approach, and focus on depicting possible future applications in all of the electronics domain. Indeed, our artistic experiment is a clear example of how to scale down electronics to an eco-sustainable level, which can still create functional, playful and useful interactions for many applicative fields.

Acknowledgements

The authors wish to thank Demian Battisti for the help in developing the first prototype system.

References

1. H. H. Lund. "Playware Research–Methodological Considerations". *Journal of Robotics, Networks and Artificial Life*, *1*, pp.23-27, 2014
2. L. Pagliarini, and H. H. Lund, The development of robot art. *Artificial Life and Robotics, 13*, 401-405, 2008.
3. H. H. Lund, and L. Pagliarini, Robotic Art for Wearable. In *Proceedings of EUROSIAM: European Conference for the Applied Mathematics and Informatics.* 2010
4. Pagliarini, L., & Lund, H. H. (2010). Wearing the playware. *Artificial Life and Robotics*, *15*(4), 381-384.
5. Lund, H. H., & 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.
6. H. H. Lund, H.H. "Playware ABC: Engineering Play for Everybody." *Journal of Robotics Networks and Artificial Life*, *3*(4), 2017

Social Play with Modular Playware

Luigi Pagliarini[1,2,3], Henrik Hautop Lund[1]

[1] Centre for Playware, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
[2] Academy of Fine Arts of Macerata, Via Berardi, 6, 405111 Macerata, Italy
[2] ISIA Design, Via C. Battisti, 198, 65123 Pescara, Italy.

luigipagliarini@gmail.com

Abstract

In this paper, we describe the development of social playware for the Moto Tiles. We exemplify the development of social playware with the design of a number of social games that people can play together on one set of Moto Tiles. The design envisions play that mediates coordination among participants, play that is physically challenging, play that is cognitive challenging, play that is changed with the different layouts of the tiles, play that is enriched by different sensor modalities, etc. In all cases, the game and interaction design is guided by the Playware ABC concept of creating these social play opportunities for Anybody, Anywhere, Anytime, by Building Bodies and Brains, and thereby facilitating that therapists and care workers can easily Construct, Combine, and Create new social play interactions to fit their particular users.

Keywords: Playware, Social Play, Playware ABC, Moto Tiles, Games.

1. Introduction

There is increasing concern about the growth and, oppositely, the declining of mental and physical health amongst people, especially children and elderly, in the modern scientific world and many theorists have begun to build theory linking both these situations to either increasing or decreasing opportunities to engage in play. A broad area of researches encompassing psychology, anthropology, education, sociology, and physiotherapy offer a new blend of practical and theoretical perspectives that may shed further light upon this topic. One thing seems to emerge from all of the studies, that free play [1,2,3,4] in children accelerates the growth of mental abilities and that social play in adults decelerates the degeneration of them [5] besides also improving their relationships, mood and physical strength. Indeed, there are a large variety of researches that demonstrate how play in adults and elderly citizens relieves stress, improve brain functions, stimulate the mind and boost creativity, improve relationships and your connection to others, let people feeling younger and energetic, besides healing physical and emotional wounds. Despite of that, sometimes, to get elderly people engaging in social play can be difficult, and we believe some special settings and that some specific tools could greatly help in reaching such target. For this reason, in the last decade we have progressively increased our research on how to build play and play tools that facilitate social play by increasing the level of communication between participants, and mediating coordination among players. All of that it is not obvious, especially if such games and such tools have to engage people from any age and still require to be physically and mentally challenging.

In this paper we report about of our experience run with the Moto Tiles platform [6,7], a platform that we constantly evolved either in terms of hardware, software and gaming in the last twenty years, and that is in turn giving magnificent result in therapeutic and social terms.

Finally, we also describe some of the most effective social games obtained.

2. Social Games

Groups-games run on Moto Tiles, can be of two specific kind:

1) *Cooperative Group Games* (**SGG**), where a cluster of persons pushes together towards a common target;

2) *Competitive Group Games* (**CGG**), where two or more groups compete to reach a wider, larger, bigger or faster goal in a specific situation (i.e. usually, scoring points).

The two kind of games have their specific psychological, sociological, pedagogical, and therapeutic functions and impact. We will not discuss that aspect in this paper although when considering social games, most likely, it is the most important factor to analyze.

Further, the above distinction is mostly theoretical since, sometimes, the very same game can be interpreted/played both in a cooperative and/or a competitive way.

As a final premise, it is to be underlined that, normally, in order to get Competitive Group Games up and running, with Moto Tiles, we often need to define how a group can be identified on the playfield. There are many possibilities, but we usually use colors, meaning that each group is represented by a specific color.

Smash Group Game.

The Smash Group Game is a simple game that can be played both as aSGG or a CGG.

As a SGG, once defined the number of participants (i.e. 3 to max 8 for each set of tiles) it will be played with a number of tiles that has to be larger than the number of players. Tiles can be positioned on the playfield both in an ordered and contiguous way or randomly. Once the participants are ready, the game is started and at each cycle the master-tablet will activate (i.e. turn on) a subset of tiles, equivalent to the number of players. To accomplish the sub-task given in every cycle, each participant has to jump on a tile and keep pressing the central button until all of the players are on all of the active tiles. Every single cycle performance, if properly accomplished, will end when all the active tiles are being pressed (i.e. an active tile shows a continuous light, while an active tile when pressed shows an intermitting light) or, if there are some unpressed active tiles, when the cycle times out. As you can easily imagine, the group has to be as fast and as accurate as possible. The average response of the group can be calculated over time and the master-tablet can provide feedback for the group speed/accuracy, as a single parameter to refer to. If appropriately executed, the game itself can be very playful, funny and has a great call upon social interaction.

The very same game can be interpreted in a CGG way. The only real difference is the master-tablet has to be predisposed for running a number of groups (i.e. 2 or 3) with a specific number of participants (i.e. not necessarily homogeneous). Once tiles and people are appropriately disposed in the room, the game is started. Of course, each sub-group will be identified by a specific color. The game remains the same although the results will be, in this case, interpreted in terms of competition amongst groups, and the tiles feedback and analysis will be presented so that there is a winner for each single cycle and a sort of score for the whole “tournament”.

Colors Group Game.

The Colors Group Game is another simple game that can be run both on SGG or CGG modality, although in our opinion the Cooperative one is the most appropriate.

In this game, each single player gets as identity a single color, which will represent her/him both as a colored tile-target and as a colored jacket or a handkerchief, or headscarf to be somehow worn.

For this game we envision 3 to max 10 players – accordingly with the number of people in the group/s, the number of tiles at hand, and with the number of distinguishable colors that can be expressed (unfortunately colors perception also depends on lighting context of the playfield/room).

At each game cycle, the master-tablet randomly distribute a number of lit up tiles so to cover all of the participants/colors range. At that point every player has to run to its tile-target and disactivate it as fast as possible. Essentially, this is a multiplayer version of the well-known Color Race game on the Moto Tiles.

The feedback provided by the game depends on the chosen modality. In the Cooperative/SGG one we show the average result of the group (i.e. in terms of timing), while in the Competitive/CGG we prefer to show the ranking of each group/color.

We consider this application at its start since there are many possible variations of it to test, measure and eventually apply.

Patterns Group Game.

The Patterns Group Game is one more game that can be used in both SGG or CGG situations.

It is quite dissimilar to the two games presented above since it does not produce stimuli for the groups simultaneously, but sequentially. As for the Colors Group Games each group is identified by a specific code that, in this case, is not a color but a specific pattern shown through the tiles' leds. A pattern, of course, is a specific geometrical configuration of the lit up eight-leds embodied in each tile, which can resemble a typical shape (i.e. circle, square, triangle, etc.).

The tiles can be freely distributed on the floor and the master-tablet only needs to know how many groups are there and show, for each of them, a specific geometrical pattern that will be the target for each individual/group.

Each run/cycle consists of a number of reiterations during which only a single tile, amongst all, will be lighted up with one specific pattern at time. All the patterns will be presented, in turn, but not necessarily in order, one after another. For example, if we have ten tiles, two group (i.e. **A** and **B**) with two patterns (i.e. **a** and **b**) there might be ten reiterations where the sequence **5 times a** and **5 times b** is randomly chosen. (e.g. AABABBBAB, or ABBAAABAB, etc.).

At every reiteration, each group will have two target tiles. One is the matching-tile and the other one is the target-tile. The matching-tile is a tile proper of each group that must be specified at beginning of the game and is positioned close to one player, the matching-player.

Therefore, for every reiteration of the game, one target-tile is lit up with a specific target, while all the matching tiles of all groups are lit up.

The first player/group which select its own matching-tile takes the turn. Indeed, the matching-tile once pushed will attribute, right or wrong, the specific reiteration to an unambiguous pattern/group.

Only after that smash/activation the target pattern tile can be switched off. Consequentially, a reiteration can be considered finished only when both a matching-tile, first, and the target-tile, secondly, are pressed.

The largest difference between the two former games and the Pattern Group Game is that cooperation and attentional demand in the latter one is much higher either in terms of visual perception, or motor reaction as well as group coordination.

The outcome of the game can be presented in different ways, therefore either as absolute hit/errors points scored, or as average reaction times for hit, or as a mixed component of all of them.

Other social play games include:

Jump Jump

The Jump Jump Game for more players, where two tiles light up and two players have to hit them at the same time. For instance, with 10 players each player has one tile, and with 5 players, each player may have two tiles. (See Fig. 1)

Piano

The Piano Game where each tile with a distinctive color represents a note. In the game a tune is played and each player has to hit the tile in front of him/her to make the appropriate note. (See Fig. 2)

Color Race

The Color Race Game where up to four players are competing on catching their color which jumps around on the tiles, and the player who first catches ten wins the game. (See Fig. 3)

Fig. 1. The JumpJump game where players can both sit down and stand up while playing together.

Fig. 2. The Piano game amongst seniors and therapist in Finland.

Fig. 3. Social intergenerational play on Moto Tiles by Culture Homes in Hong Kong.

3. Discussion and Conclusion

In this paper we described an initial set of six games to be run on our playware platform, that point directly at eliciting a deep social interaction. Indeed, in our point of view, it is important to expand the playware research to focus on the social interaction, so that the starting point for the research and development becomes the *social interaction* mediated by the technology rather than the individual interaction mediated by the technology. We can define social playware as follows: *social playware is playware which aims at creating playful social interaction between several users* [8].

We focus on such social playware because we believe, as almost all the scientific community seems to be aware now, that social play is fundamental for both mental and physical growth and maintenance. Indeed, there are many studies [1-5] carried on children and elderly population that show the multiple benefits of such an activity, improvements that go from the not concretely defined "wellness" to the more specific results obtained in the muscular, skeletal, cardio-circulatory, respiratory, metabolic systems, as well as the nervous one in terms of cognitive and emotive apparats. Therefore, working on a technology that stimulate gaming and social interaction and that parallelly produce a deterministic level of automated measurement in terms of behavioral response, it seems to be the right way to spend a scientific effort.

Acknowledgements

The authors wish to thank Massimiliano Leggieri for the help in developing the first prototype system.

References

1. P. Jarvisa, S. Newmanb, L. Swiniarskic. "On 'becoming social': the importance of collaborative free play in childhood". *International Journal of Play, 2014 Vol. 3, No. 1, 53–68, 2013*
2. R. Holmes, L. Romeo, S. Ciraola, & M. Grushko. The relationship between creativity, social play, and children's language abilities. Early Child Development and Care, 185(7), 1180-1197. 2015
3. Pyle, A., & Danniels, E. (2017). A Continuum of Play-Based Learning: The Role of the Teacher in Play-Based Pedagogy and the Fear of Hijacking Play. Early Education and Development, 0(0), 1–16. http://doi.org/10.1080/10409289.2016.1220771
4. Pellegrini, A. D., Dupuis, D., & Smith, P. K. (2007). Play in evolution and development. Developmental Review, 27(2), 261–276. http:// doi.org/10.1016/j.dr.2006.09.001
5. Ghazi, Hasanain & Sutan, Rosnah & Elnajeh, Maged & Abdalqader, Mohammed & Baobaid, Mohammed. The importance of social participation and networking among elderly people: short review. Journal of Management and Science. 15. 2017
6. Lund, H. H., & 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.
7. H. H. Lund, H.H. "Playware ABC: Engineering Play for Everybody." *Journal of Robotics Networks and Artificial Life*, *3*(4), 2017
8. H. H. Lund and T. Thorsteinsson. "Social playware for mediating tele-play interaction over distance", *Artificial Life and Robotics*, vol. 16, no. 4, pp. 435-440, 2012.

An FSK based industrial analog signal transmission

Po-Yun Shih[1], Chung-Wen Hung*[,1], Chau-Chung Song[2]
[1]*Department of Electrical Engineering, National Yunlin University of Science &Technology ,123 University Road, Section 3, Douliou Yunlin 64002, Taiwan, R.O.C*

[2]*Department of Aeronautical Engineering, National Formosa University, 64 Culture Road, Huwei Yunlin 63201, Taiwan, R.O.C*

*E-mail: wenhung@yuntech.edu.tw**

Abstract

An FSK (Frequency-Shift Keying) based industrial analog signal transmission is proposed in this paper. Due to the advantage of digital transmission, such as noise immunity and error check, the digital transmission is more and more popular. The FSK is a simple kind of basic digital transmission. Although the analog signal modulation is already well-known and mature, but it is easily interfered by the noise, especially in an industrial environment. When the carried frequency is occupied, the alternative frequency is necessary. However, it is a tough work for sure. The FSK based wireless transmission is used to perform a wireless transmission of analog signal in the industry. It also has good effects of removing the wiring from device to device, and breaks the restriction of the device movement. Experimental results show that the proposed method can work well, and the wireless transmission bandwidth achieves 16kHz.

Keywords: FSK, Wireless, Industrial communication, Active filter

1. Introduction

In recent years, the IoT (Internet of thing) concept is more and more concerned by every citizen. It is not only used in a residential environment, but can also be utilized for both agricultural and industrial application. The “IIoT” (Industrial Internet of thing) will also be proposed by this trend. The concept of IIoT focus more and emphasizes the communication of machine to machine, and machine to user, so interface between the facility’s sensor is the key point of this concept [1]-[2].
The connection between the machines always uses the physical wire in the traditional way. However, the defect is the signal type and transmission way might be limited by the distance [3]. It also causes the system couldn’t change the transmission dynamically at times when the emergency events occur. [4] also uses the MCU (Micro Control Unit) which carry the RTOS(Real-time operating system) to collect multiple sensors data in the industrial environments.
In this paper, the wireless connection is concerned to replace the physical wire between the machine to machine. Although the analog signal modulation has been used for a long while, [5] mentions that digital modulation has advantage of better noise immunity, robustness to channel impairments, easily multiplexing of various forms of information and detect and correct transmission errors by accommodated digital error-control codes.
Because of the reasons stated above, the goal on this paper is put the eye on transmit the analog signal by digital modulation. The paper is focused on the analog signal because there are many continuous analog signal on a machine, such as voltage, current or temperature of

the machine. [6] collects all the information to monitor by digital wireless transmission, then the emergency events are detected by comparing the original waveform. In [7], the MCU is used to sample the analog signal from sensors in the industrial environment, and the sampled data are transmitted by the wireless signal to the master for monitoring. However, both of two papers never discussed about reconstruction of signal.

2. System Architecture

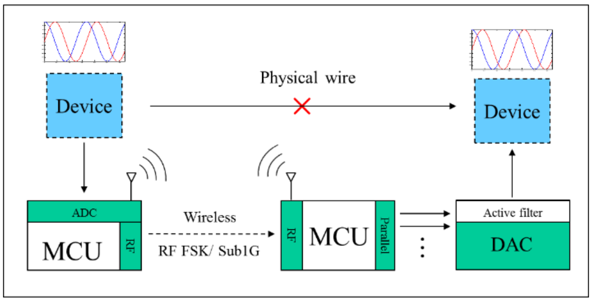


Fig. 1. The figure of system architecture

Fig. 1 shows the proposed system architecture in this paper. It consists of two MCUs. These two MCUs not only response for digital data transformation, the most important thing is to fulfill the wireless communication. Due to the setting in this paper is the unidirectional transmission, the following section is divide into two parts with transmitter and receiver to introduce in detail.

2.1 *Transmitter*

In the transmitter, the MCU enables a timer to trigger ADC (analog-digital converter) to sample the analog signal. When the sampling amount reaches the goal of RF (radio frequency) packet length, the MCU will turn the RF transmission on to transmit the whole packet to the receiver. Since the hardware blocks of RF and ADC are independent on the MCU. When the MCU performs a RF transmission, the ADC can also keep sampling to the signal, so that the data never lost.

2.2 *Receiver*

The receiver has the same design with the transmitter. In addition to enabling a timer to turn on the RF reception within a fixed period. The parallel transmission is used to send the digital data from the receiver to the external DAC (digital-analog converter). By the repeatedly receiving and converting, the analog output signal will last continuous conversion.

2.3 *Parallel Digital Analog Converter circuit*

The DAC7821 IC is adopted in the DAC circuit, which is produced by the Texas Instrument. The reason to use the parallel transmission from MCU to the external DAC circuit is the transfer speed. Due to the requirement for real time convert of the analog signal, the serious communication might cause further delay to the system which is called the latency.

2.4 *Active filter circuit*

Due to that the system in this paper uses the digital modulation of wireless communication to transmit, the analog signal data will be sampled to non-continuously digital data. Because of the DAC performs the zero-order hold for discrete-time signal, means that the data will also become unsmooth caused by the reconstruction on DAC. The active filter is designed to resolve this problem. The low pass active filter is implemented with a 4 order Butterworth filter. The goal expects that the analog signal within the setting passband can be send wirelessly and successfully and smoothly output results.

3. Experimental Results

In the experiment process, the function generator is used to produce standard sinusoidal waveform. The MCU samples this signal by a fixed period. The transmitter sends the sampling data to the receiver by RF, then the receiver can transmit the data to the DAC circuit. After the conversions on DAC circuit, an oscilloscope is used to observe and compare the output result of the active filter circuit and the input source signal.

3.1 *The maximum Bandwidth of system*

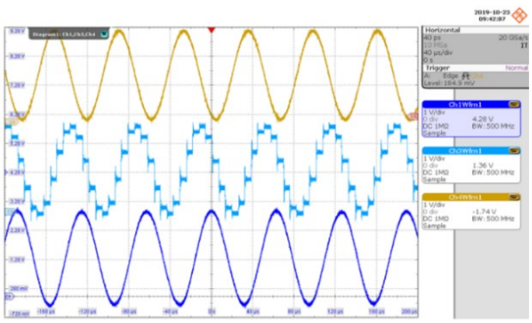


Fig. 2. The comparison of whole system output and input result

The highest sampling rate of the ADC and DAC in the paper is set on 160kS/s. The maximum passband width is set in 16 kHz because the expectation of this paper is that the input signal to be at least consistent with 10 digital data to be reconstructed. In the oscilloscope waveform of Fig. 2, Ch4 (brown curve) is the waveform produced by the function generator; Ch3 (light blue curve) is the output after DAC; and Ch1 (indigo curve) is the result of the filter which is also the signal after the reconstruction. The Fig. 2 shows that the 16 kHz sinusoidal wave input signal can be reconstructed by 10 digital data.

3.2 *Latency*

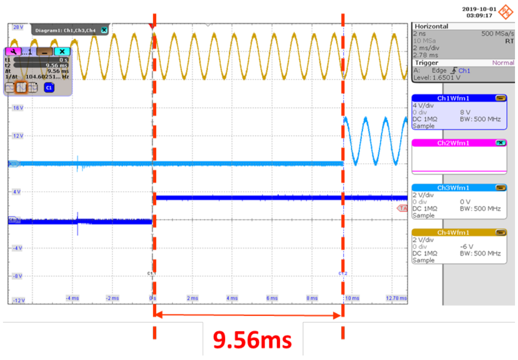


Fig. 3. The latency of input to output

The latency on the wireless communication is unavoidable. The experiment measures the effect of the latency on the result. The experiment is used an IO output pin on MCU to indicate the moment from the first ADC trigger. The interval to the first sinusoidal signal of the DAC conversion means the latency. The oscilloscope shows the result which in the Fig. 3, the Ch4 (brown curve) is the input signal; Ch3 (light blue curve) is the result of the DAC circuit output; and the two dotted lines are used to mark the latency time by the rising edge (first trigger) in Ch1 and the first complete of DAC. It is observed that from the first signal sampling to the DAC conversion, the total latency takes about 9.56ms.

3.3 Timer occupy by Hardware switching

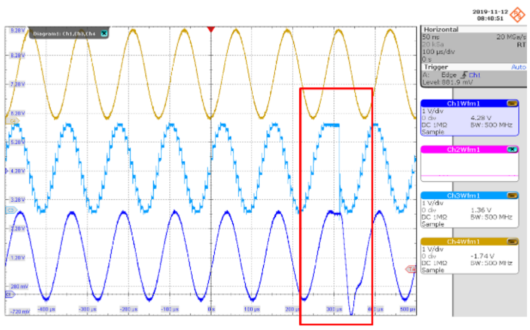


Fig. 4. The situation of hardware switching

Due to hardware limitation of the MCU, when the CPU is interrupted by RF receiver routine, the output pin can't be set. It causes the system can't update the DAC output in a fixed-length duration time. The reason is that the most of MCU never support hierarchy interrupt, when receiver interrupt, output timer interrupt for DAC update is disable. As shown in Fig. 4, the experiment is performed to measure this non-updating duration. In every 43.75us of transmission frame, there is about 3.2 ms of non-working interval.

3.4 *Mixed frequency*

The experiment is hoping that the system not just transmits the data with fixed or single frequency, but transmits the reality analog signal. Hence the experiment adds multi frequency to a signal in order to simulate it, and observe the result to know whether reconstruction is successful or not. This experiment also tests the filter performance of its exact bandwidth.

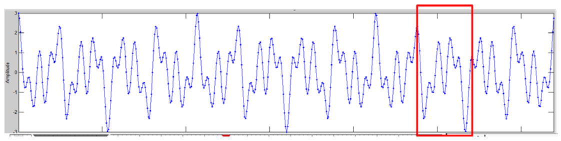


Fig. 5. The waveform data with 4kHz, 8kHz and 16kHz mixed frequency sinusoidal wave

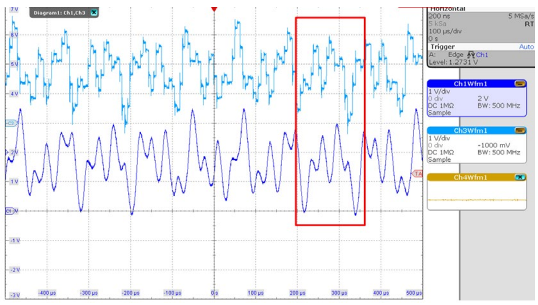


Fig. 6. The DAC output and active filter result of 4kHz, 8kHz and 16kHz mixed frequency sinusoidal wave

First of all, the waveform of the mixed frequency form 4 kHz, 8 kHz and 16 kHz sinusoidal waveform is produced as shown in Fig. 5. The Fig. 6 is measured from the output of DAC circuit. Ch3 (light blue curve) is the DAC output result and Ch1 (indigo curve) is the signal after filter.
The red square in the Fig. 5 and Fig. 6 marks the difference between the original signal and the signal after filter. From the result, it proves that within the 16 kHz bandwidth, all of the signal can reconstruct to the original signal successful.

3.5 *Signal noise ratio (SNR)*

SNR is defined as the ratio of signal power to the noise power. Since the analog continuous signal is transmitted wirelessly, the SNR is used to judge how much noise is in the output signal. In this paper, the measurement of

SNR is based on the frequency domain after the Fourier Transform. Fig. 7 shows the spectrum of the signal waveform. According to the frequency domain, calculate the energy of input signal and the noise except for the input signal. The last, the SNR can be calculated by (1), and it is effective evaluate the system.

$$SNR(dB) = 20\log_{10}\left(\frac{A_{signal}}{A_{noise}}\right) \quad (1)$$

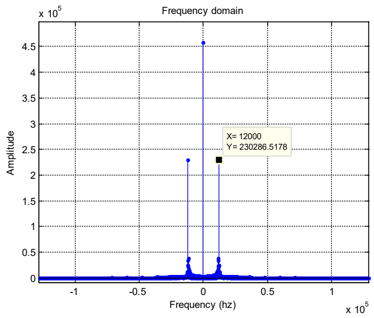


Fig. 7. The frequency domain figure of output result

Table 1. SNR of each signal input frequency

		Signal noise ratio
	1k	152.96dB
Input	2k	116.093dB
Signal	4k	98.0937dB
Freq.	8k	68.1dB
	10k	62.6037dB
	16k	41.3466dB

As mentioned above, due to the system only used MCUs to fulfill the RF reception and multiple IO to transmit the signal data, the reconstructed signal will have some vacant of signal in every transmission frame. The fixed duration reduces the SNR. The related data is measured as shown in Table 1, the lower effect to the lower frequency signal. It also shows the lower signal will have better SNR.

4. Conclusion

In this paper, the digital modulation is used to conduct the wireless communication. The system with digital modulation have the advantages of noise immunity and error checking with error-control codes. Moreover, the transmission power and transmission rate can be flexibly adjusted in the digital modulation.

As shown in experimental results, that the proposed system could effectively remove the equipment wiring in the industrial environment. Moreover, the system successfully simulate the situation of variety frequency waveform in the industrial environment, and reconstruct the output signal to the original continuous signal with a very low error.

The system which implement in this paper is proved that it can transmit the continuous analog signal by the wireless effective.

References

1. S. S. A. Abbas and K. L. Priya, Self Configurations, Optimization and Protection Scenarios with wireless sensor networks in IIoT, 2019 International Conference on Communication and Signal Processing (ICCSP), April 4-6, 2019
2. Z. Bi, L. D. Xu and C. Wang, Internet of Things for enterprise systems of modern manufacturing, IEEE Transactions on Industrial Informatics, January 16, 2014
3. Q. Chi, H. Yan, C, Zhang, Z, Pang and L. D. Xu, A reconfigurable smart sensor interface for industrial WSN in IoT environment, IEEE Transactions on Industrial Informatics, February 17, 2014
4. S. Nuratch, A universal microcontroller circuit and firmware design and implementation for IoT-based realtime measurement and control applications, 2017 International Electrical Engineering Congress (iEECON), October 23, 2017
5. Q. Gu, RF System Design of Transceivers for Wireless Communications, 1st edn. (Springer, New York, 2005), pp. 67–77.
6. A. A. Ignatius, D. C. Samjay, R. Prathish and A. I. Rexy, Automated Monitoring of Industrial Loads Using Iot and Control Using Wireless Technology, Middle-East Journal of Scientific Research, 2016
7. I. Müller, T. Companhoni, M. Götz, C. Protasio, Development of a Wireless Current Measurement Sensor Node for IoT Applications, 2018 VIII Brazilian Symposium on Computing Systems Engineering (SBESC), April 18, 2019

X-Y platform synchronous control with CANopen

Yu-Ming Guan[1], Chung-Wen HUNG*,[1], Shih-Ting Yu[1], Yu-Kai Chen[2]
[1]Department of Electrical Engineering, National Yunlin University of Science and Technology, No. 123 University Rd. Section 3, Douliu, Yunlin 640, Taiwan(R.O.C.),

[2]Department of Aeronautical Engineering, National Formosa University, No.64, Wunhua Rd., Huwei Township, Yunlin County 63201, Taiwan (R.O.C.)

*Email: M10712039@yuntech.edu.tw, * wenhung@yuntech.edu.tw ,M10612034@yuntech.edu.tw,ykchen@nfu.edu.tw*

Abstract

X-Y platform synchronous control with CANopen is proposed in this paper, the X-Y platform consists of three-phase hybrid stepper motor drivers and controllers which support the CANopen protocol. The communication network of the platform is based on CAN-bus and CANopen, and the proposed system supports the synchronization signal which meets the CiA301 communication protocol and CiA402 motion control protocol. The Interpolated Position mode of the CiA402 standard is implemented to improve the precision of position control. Compared with Interpolated Position mode, the other control method, cross-coupling control is also implemented to reduce the synchronization error of biaxial motion control. The experimental results show that both of the two method work well, and the comparison also is discussed.

Keywords: CAN (Controller Area Network), CANopen, X-Y Platform, synchronous control system, multi-motor, cross-coupling control.

1. Introduction

With the vigorous development of industry, production efficiency and precision machining are more and more important. And, the high precision position control has become the main development trend of motion control. The key point of the previous discussion is detailed in this paper, such as three-phase hybrid stepping motor, CAN, X-Y platform.

In the first part, the rotor and stator of the three-phase hybrid stepping motor are consisted of permanent magnet and toothed magnetic structure, and the structure can generate permanent magnet torque and reluctance torque [1]. And, the open loop control and the switching current triggers using pulse signals is the feature of the three-phase hybrid stepping motor. The features drives the motor with a stable angle without additional position feedback sensor. However, the resolution of position control is limited. To increase the accuracy, an encoder is installed for achieving higher precision.

About the control method, the constant current control is discussed in [2], and the resolution of step angle can be improved. And, the other method using SVPWM (Space Vector Pulse Width Modulation) algorithm to subdivision the current of the stator is proposed [3]. The algorithm could make the current approximate the sine wave, and then the low-frequency oscillation and noise suppression could be decreased. The above methods can improve the control accuracy to meet with the demand of modern location control.

The second part, the communication interface will be discussed. The CAN Bus is a kind of serial communication network, and is used in many fields, such as industrial communication, automotive electronics systems, etc. The CAN Bus can support the real-time internal distributed communication, but that traditional point-to-point communication could not satisfy [4]. CAN have features including high reliability, simple development, and low cost. However, the communication format of the application layer is not defined in CAN. To support the requirement for the automation application, the application layer protocol, CANopen, is proposed. Moreover, CAN-in-Automation (CiA) standard is proposed to define the basic CANopen communication protocol [5]. And, CANopen protocol is implemented in embedded devices in this paper.

Finally, in this paper, the X-Y platform is consisted of above two parts: motor control and CAN. There are many ways could control multi-axis systems, such as master-slave parallel control, master-slave series control [6] and cross-coupling control. The concept of the master-slave system architecture is to divide a system into two parts. One is the master system, and the other one is the slave system.

According to the difference of command transmission methods, the architecture could be divided into parallel control and series control. The parallel control is that the two-axis could be executed the same command action at the same time. The characteristic of parallel control is easy to control, but there isn't any communication between the two axes. When a serious location error occurred, it may cause damage to some mechanical structure, such as a car rear door system. And then, the series control method is that the command could be delivered from the master axis to the slave axis, and the slave axis will follow the master axis. Compared with parallel control, the series control structure can improve the location error effectively. However, the series control also has a shortcoming that is response is slower. The other control method, cross-coupling was proposed by Koren [7] in 1980, and the mutual motion between multi-axis systems could be improved by this method. And, a similar method, using tracking error of individual systems to compensate in time is proposed to improve the contour error [8]. The contour error refers to the shortest distance between the command position and the actual position. The cross-coupling architecture is shown in Fig. 1, and the effects of this architecture will be presented in tabular form in later chapters.

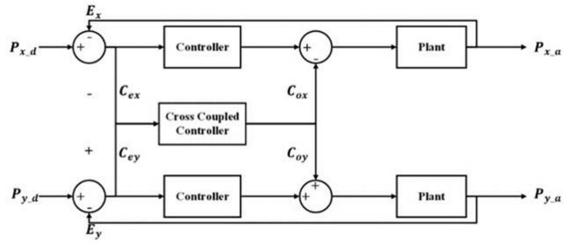


Fig. 1: Cross-scoupling architecture [7]

2. Hardware Architecture

The hardware part of the system will be introduced in this chapter. In order to meet the requirements of multi-axis control, CANopen is used as the main communication protocol. Two three-phase hybrid stepping motors and three microcontrollers are used in this paper. And, the one microcontroller is CANopen master, and the two of microcontrollers is CANopen slave which connected the motor driver. The system architecture achieving the decentralized control required for multi-axis control is shown in Fig. 2.

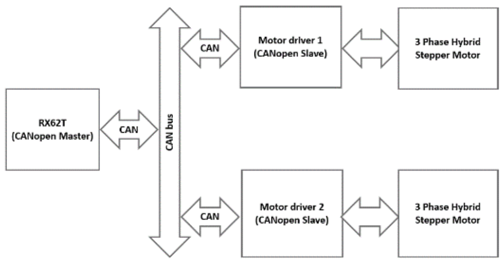


Fig. 2: System architecture diagram

The microcontroller, Renesas RX62T is used to control motor driver in this paper, and the pins and peripherals of RX62T are shown in Fig. 3. The three sets of PWM signal are used to control gate drive circuits, the required signal of three-phases six switched is generated. And, through the P11 IO pins, gate drive could be enabled or disabled. In addition, the feedbacks of the motor driver divided into three following parts:

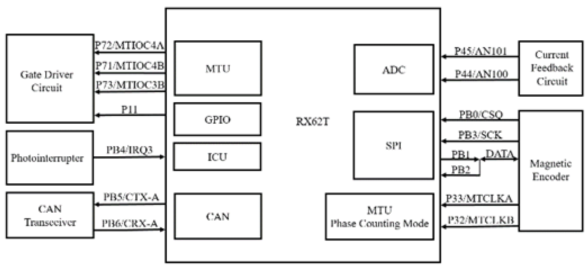


Fig. 3: Pins and peripherals of RX62T

2.1. *Current sensor*

The sensor is used to measure the current of the motor for PI constant current control, then the stability of the torque will be improved.

2.2. *Magnetic encoder*

The magnetic encoder is a high accuracy position sensor, and used to improve the position control accuracy.

2.3. *Photointerrupter*

The photointerrupter is used as the main detect way for homing mode. When the photointerrupter is triggered, an interrupt message will be sent to the microcontroller to reset home.

3. Software Design

CANopen is an application layer protocol built on the CAN's physical layer and data link layer. CANopen is often used in embedded systems, and it is a common protocol for industrial control. The specification of the CANopen device is defined in a standard, CiA. The basic CANopen device status and communication protocol are defined in CiA301, and specification of motion control is defined in CiA402. Moreover, CANopen devices could communicate through PDO (Process Data Object) or SDO (Service Data Object), and the data length of CANopen is not over 8 bits. And, the NMT (Network Management) function can be used to check the work status of nodes. The following control methods defined in CiA402 is implemented in this paper: IP (Interpolated Position) mode. And the Slave-CCC (cross-coupling control) mode is also implemented that will be analyzed performance compared with IP mode.

3.1. *Interpolated Position mode:*

IP mode is a location control method to coordinate multi-axis control, and the relevant driving elements can be coordinated through the time synchronization mechanism. Linear interpolation is applied in this mode. The interpolation period can be set by interpolation object of CANopen, and the interpolation data in a periodic is recorded as a reference for the next periodic interpolation. A FIFO (First In, First Out) structure is implemented as the Interpolated Position buffer in this paper, and is realized in the circular structure. The data of two-axis will be calculated by circular interpolation methods which are expressed as follows Eq. (1) and Eq. (2)[9].

$$X_{N+1} = Rcos(\theta_N + \Delta\theta) = X_N - \Delta\theta(Y_N + \tfrac{1}{2}X_N\Delta\theta) \quad (1)$$

$$Y_{N+1} = Rsin(\theta_N + \Delta\theta) = Y_N + \Delta\theta(X_N - \tfrac{1}{2}Y_N\Delta\theta) \quad (2)$$

Where X_{N+1} and Y_{N+1} are the target position of X-axis and Y-axis , and the X_N and Y_N are the present position of X-axis and Y-axis. θ_N is the present angle of the X-Y platform. The radius R and angle displacement $\Delta\theta$ determine the result of circular interpolation methods.

3.2. *Slave-CCC mode:*

In this mode, two slave devices will exchange their location information, and the new speed and location will be calculated by cross-coupling equations Eq. (3) and Eq. (4). And then, the new speed and location command will be executed.

$$X_{N+1} = E_{X_N} - \Delta\theta\left(E_{Y_N}\right) - \tfrac{1}{2}(E_{X_N})(\Delta\theta)^2 \quad (3)$$

$$Y_{N+1} = E_{Y_N} + \Delta\theta\left(E_{X_N}\right) - \tfrac{1}{2}(E_{Y_N})(\Delta\theta)^2 \quad (4)$$

Where E_{X_N} and E_{Y_N} are the error between present position and target position.

4. Experimental Result

The CANopen master, the RX62T board is shown in Fig. 5, and the three-phase hybrid stepper motor and driver circuit board are shown in Fig. 6.

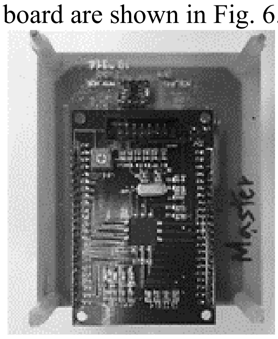


Fig. 5: CANopen master

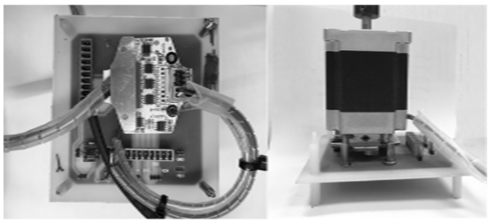


Fig. 6: Motor and driver circuit

An experiment is set for a circular motion control with a circular radius of 50 mm, in five different rotational speeds and two different control ways. The average RMS of error between command and actual position is shown in Fig. 7 and Table 1. The lower rotational speed, the smaller RMS error in the Interpolated Position mode and the Slave-CCC mode. However, the RMS error is the almost same between 188 and 376 sec./circular in the Interpolated Position mode, but the Slave-CCC's is drastically reduced.

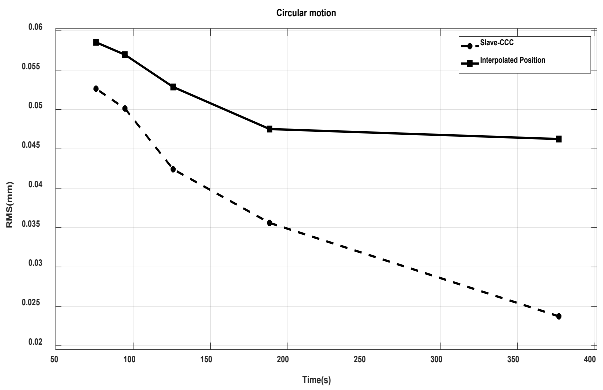


Fig. 7: Time for executing a circular motion in two different modes

Table 1: RMS(mm) of IP mode and Slave-CCC mode

Time / mode	79s	94s	125s	188s	376s
IP mode(mm)	0.0585	0.0569	0.0528	0.0475	0.0462
Slave-CCC(mm)	0.0526	0.0501	0.0424	0.0356	0.0237

5. Conclusion

A multi-axis synchronous control system based on CANopen is proposed in this paper. The standard CiA 301 and CiA 402 are supported by the proposed system. The experimental results of circular motion show the Interpolated Position mode and Slave-CCC mode have high accuracy. And the Slave-CCC mode has smaller error than Interpolated Position mode as the rotational speed is slower. The synchronous system could reduce the position error effectively.

Acknowledgements

This work is partially supported by the Ministry of Science and Technology, ROC, under contract No. MOST 107-2218-E-150-007, 107-2622-E-224-018-CC3 and 108-2221-E-224-045-.

References

1. Alberto Bellini, Carlo Concari, and Andrea Toscani. Mixed-mode PWM for High-Performance Stepping Motors. IEEE Transactions on Industrial Electronics, Vol.54, No.6, 2007, pp.3167-3177, last viewed in 2007.
2. Chawakorn Yodsanti and Mongkol Konghirun. Implementation of Two-Axis Highly Accurate Position Control of Space-Vector PWM Based Microstepping Bipolar Drive. 2008 International Conference on Electrical Machines and Systems, IEEE, 2008, pp. 1027-1031, last viewed in 2008.
3. Qi Guan and Jiaoyu Liu. The Vector Control Research and Implementation of The Three-phase Hybrid Stepping Motor. 2016 International Conference on Industrial Informatics-Computing Technology, Intelligent Technology, Industrial Information Integration (ICIICII), IEEE, 2016, pp. 384-387, last viewed in 2016.
4. Chung-Wen Hung and Bo-Kai Huang. Design and implementation of micro-stepping motor driver with CANopen. Taiwan master's thesis, Department of Electrical Engineering, National Yunlin University of Science and Technology, last viewed in 2017.
5. Dai Qiang Wang, Shi-You Gao, Yu-Qing Chen, Yi Wang, and Qiao Liu. Intelligent Control System Based on CAN-bus For Car Doors and Windows. 2009 3rd International Conference on Anti-counterfeiting, Security, and Identification in Communication. IEEE, 2009. pp. 242-245, last viewed in 2009.
6. Fan Yang, Ran Chen, and Zhi-hui Chen. Communication of the Wind Turbine Testing System Based On CANopen Protocol. 2011 IEEE Power Engineering and Automation Conference, IEEE, 2011. pp. 162-165, last viewed in 2011.
7. Jun-Xian Yang. The Identification and Control of a Twin Linear Servo System with Mechanical Coupling. NCKU Tainan, last viewed in 2003.
8. Yoram Koren. Cross-Coupled Biaxial Computer Control for Manufacturing Systems. Journal of Dynamic Systems, Measurement, and Control, 1980, 102.4: 265-272, last viewed in 1980.
9. Yan-Han Chen. XY Table Motion Control by Using Cross Coupling Law, Department of Electrical Engineering National Sun Yat-sen University Master Thesis, 2016

A PFC Converter with Voltages Double Characteristic for Universal Input Voltage Applications

Ching-Chun Chuang[1], Hung-Chi Lee[*,2], Chih-Chiang Hua[3], Chih-Wei Chuang[3], Chuan-Ming Niu[3]
[1]Bachelor Bachelor Program in Interdisciplinary Studies, National Yunlin University of Science and Technology, No. 123, University Rd., Section 3, Douliu, Yunlin 640, Taiwan (R.O.C),

[2]Department of Electronic and Computer Engineering, National Taiwan University of Science and Technology, No. 43, Section 3, Keelung Rd., Taipei 106, Taiwan (R.O.C),

[3]Department of Electrical Engineering, National Yunlin University of Science and Technology, No. 123, University Rd., Section 3, Douliu, Yunlin 640, Taiwan (R.O.C),

E-mail: austincc@yuntech.edu.tw, [*]D9602207@mail.ntust.edu.tw, huacc@yuntech.edu.tw, solarpower.david@gmail.com, a5712316@gmail.com

Abstract

In this paper, a current-fed bridgeless power factor correction rectifier with voltage-double is proposed for a hybrid electric vehicle charging system. The proposed PFC rectifier was simulated on a 3.4-kW prototype. The differences of the simulated THD and PF between the proposed converter and the conventional interleaved PFC converter are insignificant. The proposed rectifier shows an improved low-line efficiency compared to its conventional counterpart under 1.5-kW output power.

Keywords: current-fed, voltage doubler, power factor correction rectifiers, hybrid electric vehicle charging system

1. Introduction

The low transfer efficiency occurs in low line condition for general conventional power factor correction rectifier. This drawback can be improved by proposing PFC with voltage double characteristic. The topology can be seen on Ref. 1 & Ref. 2. However, PFC circuit should have multiple switches to realize function of double voltage in low line condition, it is not suitable for server power because issues of efficiency, reliability and cost. Basically, in design and develop battery charging system of electrical vehicle, optimal power circuit selection and power components dissipation evaluation should be well treated.

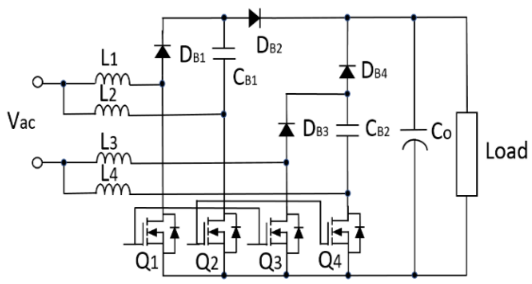


Fig. 1 The proposed current-fed bridgeless interleave PFC boost converter with voltage double characteristic

In this paper, improvement and design-in was conducted to solve low efficiency drawback of current-fed bridgeless PFC converter used in low line light load condition. The proposed PFC boost converter with voltage double characteristic improves efficiency of current interleave bridgeless PFC operated in low line and light load condition and meet input current harmonics specification as well. It can reach high power factor and high efficiency and suitable to apply for EV battery storage and charging system to have power reaching 3.4 kW.

2. Operating Analysis

Refer to Fig. 1, the operating and steady state circuit analysis of positive half cycle of the AC line input was implemented only. In the symmetrical control operating period of positive half cycle, test condition of duty cycle (0<D<0.5) was analyzed and deduced for the ideal waveform. Test condition for duty cycle (0.5<D<1) is similar to the other and was neglected.

Regard the semiconductor components as ideal, operating of one switching cycle period was discussed.

(a) For duty cycle (0<D<0.5),

Period 1$[t_0 - t_1]$: The ripple current of L_1, L_3,

$$\Delta i_{L1} = \frac{V_o - V_{ac}}{L_1 + L_3}\left(\frac{1}{2} - D\right)T_s \tag{1}$$

The ripple current of L_2, L_4

$$\Delta i_{L2} = \frac{[(1-D)V_o - V_{ac}]}{L_2 + L_4}\left(\frac{1}{2} - D\right)T_s \tag{2}$$

Period 2$[t_1 - t_2]$: The ripple current of Q_1, Q_3 and current of L_1, L_3 increase linearly as well

$$\Delta i_{L1} = \frac{V_{ac}}{L_1 + L_3}DT_s \tag{3}$$

The ripple current of L_2, L_4

$$\Delta i_{L2} = \frac{[V_{ac} - (1-D)V_o]}{L_2 + L_4}DT_s \tag{4}$$

Combine (3) and (4), the total input ripple current ΔI_{in} is the current of addition of L_1、 L_3 and L_2、 L_4

$$\Delta I_{in} = \frac{[2V_{ac} - (1-D)V_o]}{L_1 + L_3}DT_s \tag{5}$$

Period 3$[t_2 - t_3]$: The ripple current of L_1, L_3

$$\Delta i_{L1} = \frac{V_o - V_{ac}}{L_1 + L_3}\left(\frac{1}{2} - D\right)T_s \tag{6}$$

The ripple current of L_2, L_4

$$\Delta i_{L2} = \frac{[(1-D)V_o - V_{ac}]}{L_2 + L_4}\left(\frac{1}{2} - D\right)T_s \tag{7}$$

Period 4$[t_3 - t_4]$: The ripple current of Q_2, Q_4 and current of L_2, L_4 increase linearly as well.

$$\Delta i_{L2} = \frac{V_{ac}}{L_2 + L_4}DT_s \tag{8}$$

When Q_1、 Q_3 turn off, current of Q_1 is zero and ripple current of Q_3 is the same as that of L_3 current of L_1 turns to increase linearly

$$\Delta i_{L1} = \frac{(V_{ac} - DV_o)}{L_1 + L_3}DT_s \tag{9}$$

Advance to combine (8) and (9), the total input ripple current ΔI_{in} is the current of addition of L_2、 L_4 and L_1、 L_3.

$$\Delta I_{in} = \frac{2V_{ac} - DV_o}{L_1 + L_3}DT_s \tag{10}$$

(b) For duty cycle (0.5<D<1),

the derivation is similar, steady state waveform of proposed converter refer to Fig. 2 & 3.

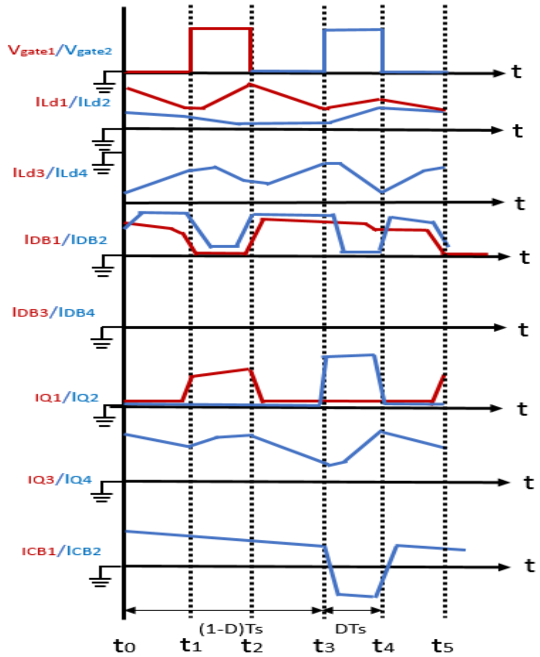


Fig. 2 Steady state waveform of proposed converter during duty cycle 0<D<0.5

3. System voltage conversion analysis

Refer to (11), when the current of L_2, L_4 increase, the equivalent equation is

$$V_{ac} = L_2\frac{di_{L2}}{dt} + L_4\frac{di_{L4}}{dt} \tag{11}$$

From equation (11), suppose $L = L_2 = L_4$ and

$$\frac{di_{L2}}{dt} = \frac{V_{ac}}{2L} \tag{12}$$

When current of L_2, L_4 decrease, the equivalent equation refer to (13) as

$$V_{ac} = \frac{V_o}{2} - L_2\frac{di_{L2}}{dt} - L_4\frac{di_{L4}}{dt} \tag{13}$$

From equation (13), suppose $L = L_2 = L_4$ and

$$\frac{di_{L2}}{dt} = \frac{1}{2L_2}\left(\frac{V_o}{2} - V_{ac}\right) \tag{14}$$

Observe Fig. 2 to analyze current of L_2, following volt-second balance theorem.

$$V_{ac}DT_s + \left(V_{ac} - \frac{V_o}{2}\right)(1-D)T_s = 0 \tag{15}$$

While the following voltage conversion ratio obtained.

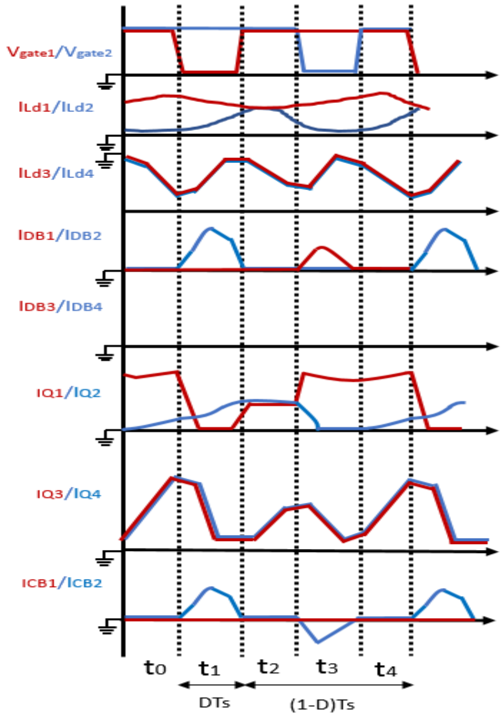


Fig. 3 Steady state waveform of proposed converter during duty cycle 0.5<D<1

$$\frac{V_o}{V_{ac}} = \frac{2}{1-D} \tag{16}$$

According to the above voltage conversion equation, higher voltage gain ratio can be obtained while duty ratio D < 0.5 and current of L_2 decrease relatively according to equation (14). The same derivation applies for duty cycle 0.5<D<1, while the following voltage conversion ratio obtained.

$$\frac{V_o}{V_{ac}} = \frac{1}{D-D^2} \tag{17}$$

According to the above voltage conversion equation, D <0.5 voltage gain ratio is similar to D > 0.5.

According to equation to constrain the lowest limit for duty ratio to avoid uncontrolled duty ratio. Therefore, when D is 0.1, V_o is 11 times of V_{ac}, is the minimum duty ratio of the proposed converter and the voltage gain is 11. Fig. 4 illustrates the voltage conversion gain ratio of the proposed converter.

4. Simulation Result

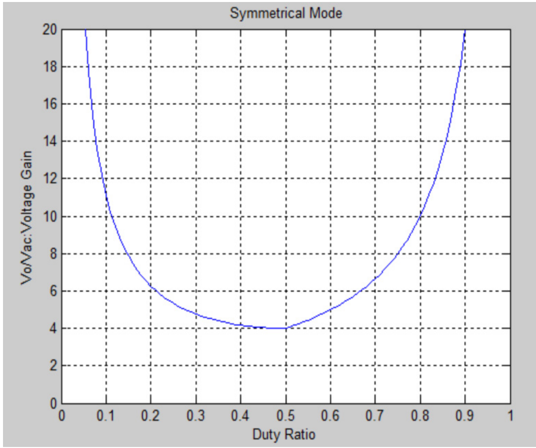


Fig. 4 Voltage conversion gain ratio of the proposed converter

The proposed PFC converter was compared and evaluated with conventional bridgeless interleave PFC converter. A set of 3.4 kW prototype design includes full range input (85-264Vrms) and 400V/8.5A output electrical characteristic. Fig. 5 was the efficiency comparison for ac input of 85Vrms, 264Vrms respectively.

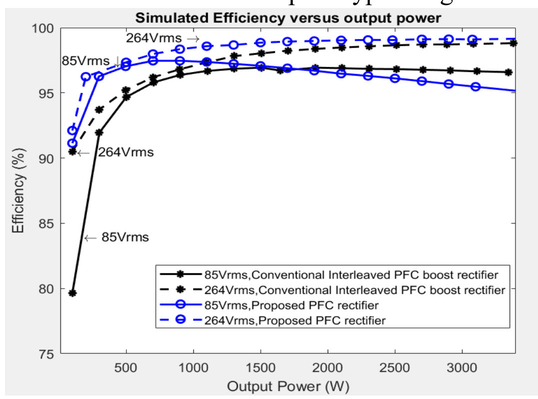


Fig.5 Simulation result of comparison of efficiency of the proposed converter and conventional interleave bridgeless PFC converter.

The simulation waveforms of output voltage V_o at $P_o = 3.4kW$ and $V_o = 400V$; input voltage V_{in} and input current I_{in} was stated in Fig. 6. Based on the result that the efficiency of the proposed converter can reach 98% during 264Vrms input voltage and out power of 0.5kW-3.4kW condition.

Fig. 7 and 8 simulate THD and PF during input voltage range 85-264Vrms. In Fig. 7, it can be observed that the

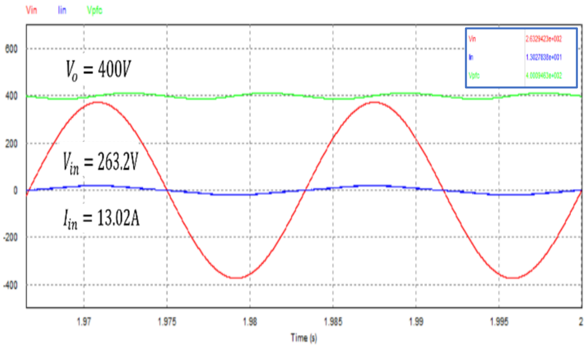


Fig.6 Simulation result of comparison of waveforms

low line operated THD of proposed converter is lower than the conventional one while it is opposite.

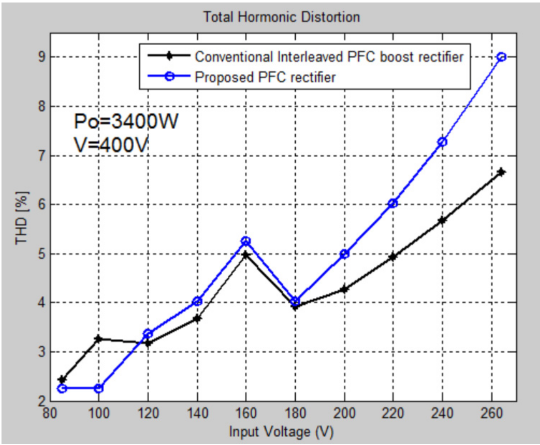


Fig. 7 The comparison of THD simulation result

when operated in high line. The PF of the proposed converter reveals in Fig. 11 is lower than the conventional one.

5. Conclusion

In this paper, the proposed PFC converter possesses high conversion efficiency when operates in low line condition. The proposed converter prototype was simulated and analyzed including efficiency, THD and PF. Comparison was made for full range input voltage, full load 3.4 kW and 400V output test conditions. The lower THD characteristic appears for the proposed converter when operates in low line condition and lower PF when operates for full input range. The efficiency reaches 97% when it was operated at 70kHz switching

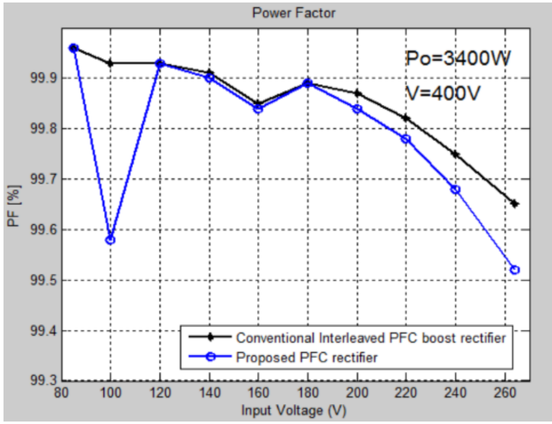


Fig. 8 The comparison of PFC simulation result

frequency, 264V input voltage and 0.5-3.4kW output power.

References

1. J. C. Salmon, Circuit topologies for single-phase voltage doubler boost rectifier, *IEEE Trans. Power Electron.*, vol. 8, no. 4, pp.521–529, Oct. 1993.
2. D. Maksimovic' and R. Erickson, Universal-input high-power-factor, boost doubler rectifiers, *in Proc. IEEE Appl. Power Eelctron. (APEC), Conf.*, 1995, pp. 459-465.vol. 1.
3. F. Musavi, W. Eberle, and W. G. Dunford, A High-Performance Single-Phase Bridgeless Interleaved PFC Converter for Plug-in Hybrid Electric Vehicle Battery Chargers, *IEEE Trans. on Industry Application*, vol. 47, no. 4, pp.1833–1843, July/August. 2011.
4. Y. S. Lai and K. M. Ho, FPGA-based digital-controlled power converter with universal input meeting 80 plus platinum efficiency code and standby power code for server power applications, *Proc. Int. Conf. ICRERA*, Nov. 2012.
5. S. H. Ryu, D. H Kim, and M. J. Kim, Adjustable Frequency-Duty-Cycle Hybrid Control Strategy for Full-Bridge Series Resonant Converters in Electric Vehicle Chargers, *IEEE Trans. on Industrial Electronics*, vol. 61, no. 10, pp.5354–5362, october. 2014.
6. D. Moon, J. Park, and S. Choi: New Interleaved Current-Fed Resonant Converter With Significantly Reduced High Current Side Output Filter for EV and HEV Application, *IEEE Transactions On Power Electronics.*, vol. 30, no. 8, pp. 4264-4271, august. 2015.

Design of a low pulse high current LLC resonant converter for EDM applications

Yu-Kai Chen[a] Min-Feng Lee[a] Yung-Chun[a] Wu Jui-Yang Chiu[b]
*[a]Department of Aeronautical Engineering,
National Formosa University, Huwei, Yunlin, Taiwan*
E-mail: ykchen@nfu.edu.tw and lord159951@gmail.com

[b]AcBel Polytech Inc.
No.159, Sec. 3, Danjin Rd., Danshuei Dis.,New Taipei City 251, Taiwan (R.O.C.)
Web-site: http://www.acbel.com

Abstract

The design of the LLC resonant converter is presented for use in Electric discharge machining (EDM) applications. The converter is designed to operate with a 3phase ac input voltage and will output controlled dc voltages during the striking and arc conditions of EDM process. The method of the LLC resonant converter can operate at zero voltage and zero current transitions, we change the output current and control the current waveform of the EDM.

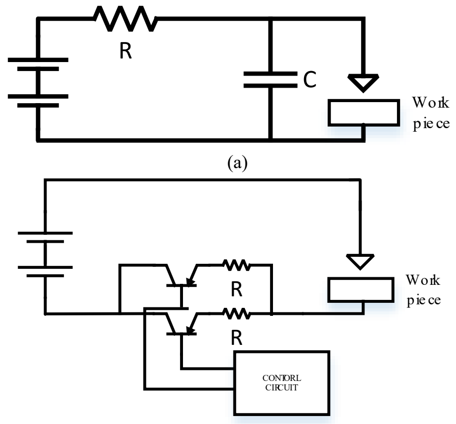
Keywords: Electric discharge machining (EDM), LLC, Resonant converter, Frequency control, Sic-MOSFET

1. Introduction

In this general, we have developed and improve the technology of performance equalization control of Electric discharge machining (EDM) applications. We change the output current and control the current waveform of the EDM. To control the current and the loss of the electrodes and increased the processing speed and the processed object will not be too Rough so we have designed a smoothly rising current to let the electrode lossless; and to processing the super-hard alloy we have designed a low pulse high output current.

2. Developments in Power Configurations

The shapes of voltage and current pulses in the discharge gap depend on the chosen power supply. There are two types of power supplies that have received most interest amongst the scientific community. They are Resistance Capacitance (RC) power supply, transistor switching circuit, Figure1 show the two type of power supply.



(b)

Fig 1(a) Resistance Capacitance (RC) power; (b) transistor switching circuit

Those two type of EDM power supply topology applies a square wave current as shown in Figure 2, the waveform of a fixed current pulse width method is used and control the pulse width of the discharge and keep current certain width to maintain the surface of the work

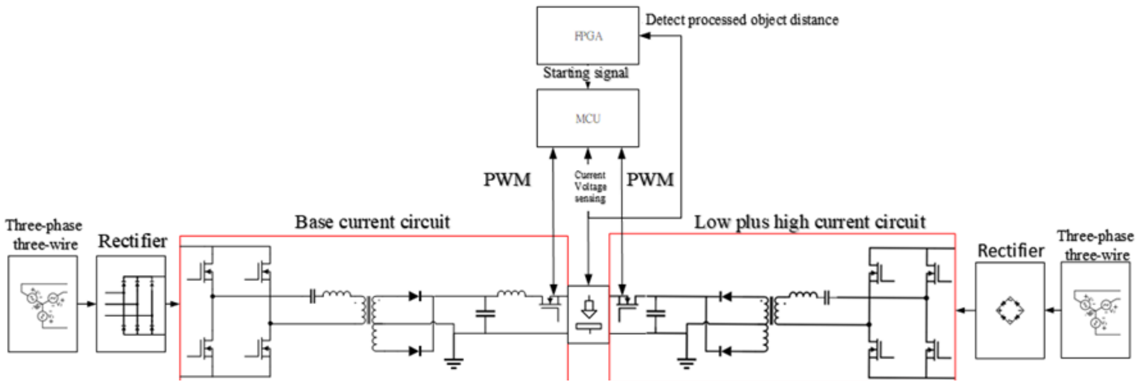


Fig 2 Circuit diagram of the EDM power supply using a full-bridge LLC resonant converter.

piece smooth. However, to increase the processing speed, the peak current Ip of the discharge current pulse will increase, but the processing surface will become rough, and the electrode consumption ratio will also increase. In order to reduce the electrode consumption, the discharge current pulse width T can be increased. At this time, as the discharge current pulse width T increases, the thickness of the work piece surface also becomes rough.

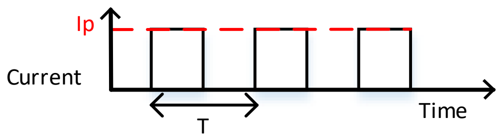


Fig 3 output current waveform

Therefore, a new circuit topology is proposed, and the output waveform is as shown in Figure 3. It consists of a slowly rising base current square waveform t1 to t2 and a low pulse high peak current waveform t2 to t3. These two waveforms are composed to find out whether the characteristics of the processed object can be maintained. The homogenization, acceleration of processing speed, reduction of electrode loss and improvement of the thickness of the machined surface. The relationship of these characteristics to the discharge current waveform and the realization of the current waveform circuit finally proposed.

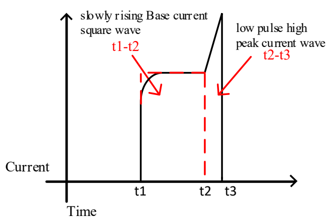


Fig 4 new design output waveform

3. Full-bridge LLC resonate converter experimental Prototype

The designed EDM impulse generator is a full-bridge resonant converter whose switching frequency is far higher than the machining frequency and we using Sic MOSFET as switch. LLC resonant converters are able to achieve the required voltage for the dielectric breakdown and, working above the resonant frequency current lags voltage so this topology achieves zero voltage switching, resulting in minimum switching losses.

3.1. *Full-bridge LLC resonate converter overview*

To achieved two different waveforms, so we set two parallel LLC circuit to create base current and low plus high peak current, these two converters are designed to operate with three-phase three-line ac input voltage and output will be controlled two different dc voltages during the EDM process, the method employed magnitude and frequency control to enable the converter to operate at zero voltage and zero current transitions. During the EDM process will have two different output current first is the base current, the base current is to reduce the electrode the current will be smoothly rising to the base current, when the base current has risen to the stable state, the second LLC will be started output the low pulse high current waveform shows in Figure 4.

3.2. *LLC resonant inverter analysis and design*

LLC resonant converter gain K is the product of the bridge switching gain, the resonance loop gain, and the transformer's primary-side turns (NP) and secondary-side turns (NS). The resonant tank circuit is composed of the element resonant inductor Lr, resonant capacitor Cr, and excitation inductance Lm. The resonant element is used to achieve zero voltage or zero current switching of the power switch, and the switching frequency is adjusted to change the output voltage. The resonance loop gain can be analyzed by analyzing the equivalent resonance The circuit is shown in Figure 5. The equivalent circuit shows that its resonance gain is shown in Equation (1).

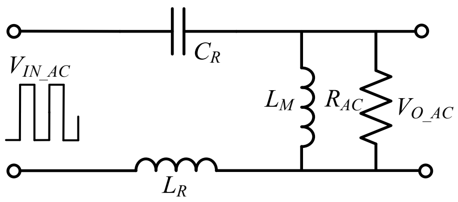


Fig 5 LLC series-parallel resonant tank equivalent circuit diagram

$$K(Q,m,F_x)=\left|\frac{V_{O\_AC}(s)}{V_{IN\_AC}(s)}\right|=\frac{F_x^{\,2}(m-1)}{\sqrt{\left(m\cdot F_x^{\,2}-1\right)^2+F_x^{\,2}\cdot\left(F_x^{\,2}-1\right)^2\cdot(m-1)^2\cdot Q^2}} \quad (1)$$

3.3. *Design the resonant tank*

The circuit parameter design The K value is obtained by the above method, and its verification parameters are already a better design of the circuit. Therefore, the Q value, Fx and fr can be used to solve the equation, and the resonance can be obtained by equations (2) and (3) The resonance inductance and resonance capacitance of the tank are obtained by formula (6) to complete the parameter design of the resonance tank element, the formula is as follows.

Quality factor Q:

$$Q=\frac{\sqrt{L_R/C_R}}{R_{AC}} \quad (2)$$

Resonant frequency fr:

$$f_r=\frac{1}{2\pi\sqrt{L_R\cdot C_R}} \quad (3)$$

The reflected load resistance at full load:

$$R_{ac}=\frac{8}{\pi^2}\times\frac{N_p^{\,2}}{N_s^{\,2}}\times R_o \quad (4)$$

Ratio of switching frequency fs and resonance frequency fr:

$$F_x=\frac{f_s}{f_r} \quad (5)$$

Ratio of primary inductance to resonant inductance m:

$$m=\frac{L_r+L_m}{L_r} \quad (6)$$

The influence of design parameters on voltage regulation and efficiency performance, and the design of resonant tank parameters according to system specifications. The ultimate design goal is to achieve the load to operate at the best performance under any conditions. The detailed design flowchart of the design method is shown in the Figure6

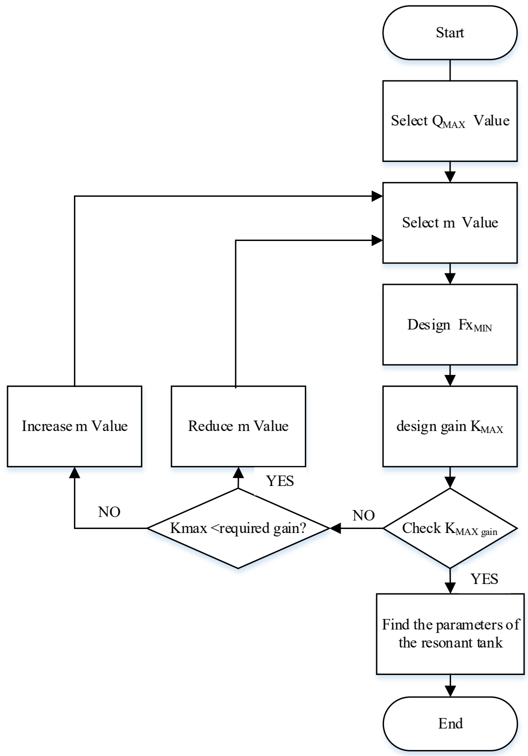


Fig 6 Resonant tank parameter design flow chart

4. Full-bridge LLC simulation results

Follow the step of the flow chart shown in Figure 6 to design the LLC resonant converter. The base current

Table 1 Base current wave form Specifications

Base current wave form	Specifications
Input voltage 3Φ3W	220V
Output voltage(V_{o1})	60V
Output power(P_{o1})	120W
Switching Frequency	100kHz
Waveform time(t1-t2)	300ms

output voltage has design 60v and the low pulse high current designs with 100v, the design data show in Tabel1 and Table2.

Table 2 Low plus high current peak waveform Specifications

Low plus high current peak waveform	Specifications
Input voltage3Φ3W	220V
Output voltage(V_{o2})	100V
Output power(P_{o2})	500W
Switching Frequency	100kHz
Waveform time(t2-t3)	100ms

4.1. *Full-bridge LLC simulation*

In this section, the diagram of the EDM system with ac to dc power supply and transistorized switching circuit as pulse generator designed in PSPICE and PSIM and schematic diagram is shown in Figure 7 there are two LLC converter which private 60v and 100v and the simulation output is shown in Figure 8.

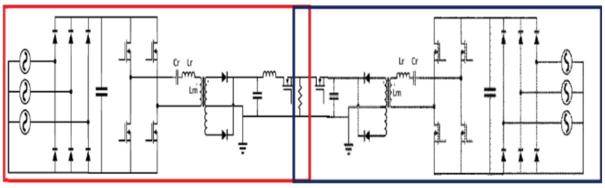


Fig 7 schematic diagram

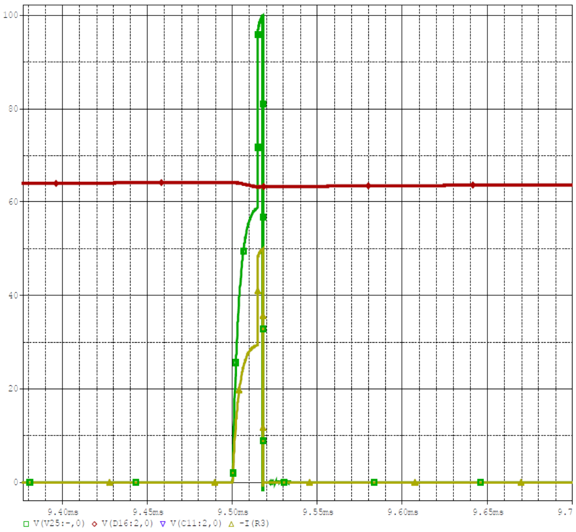


Fig 8 Results of parallel LLC simulation (green is output voltage; yellow is output current)

5. Experimental results

The simulation has approved that circuit can get the output that a smoothly rising current and a low pulse high peak current and next we design a prototype to test the waveform the prototype shown in Figure 9 and the output waveform shown in Figure 10.

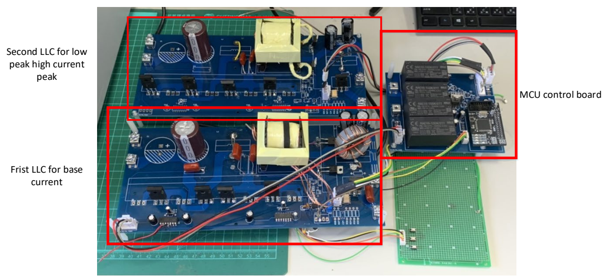
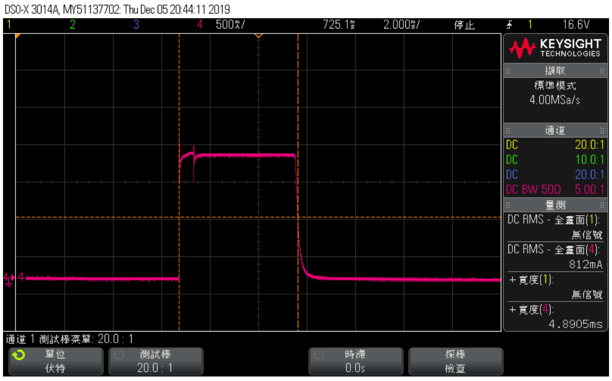
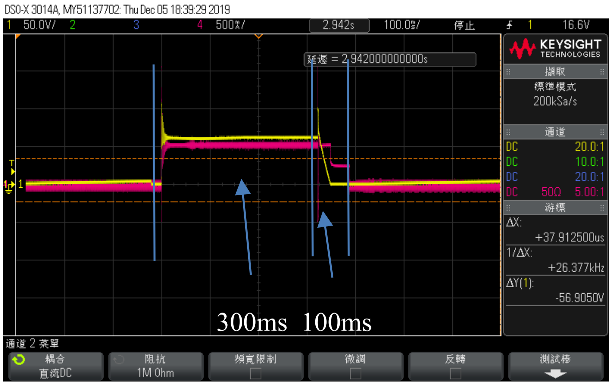


Fig 9 Prototype power system



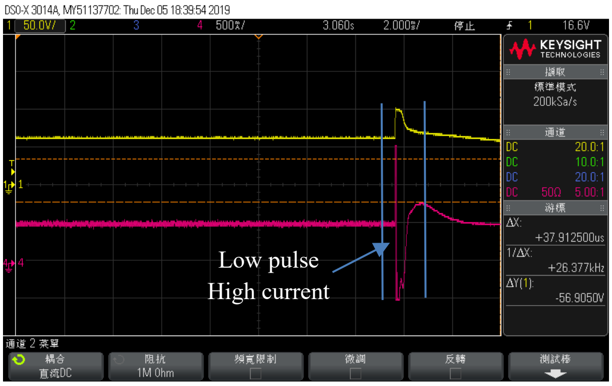
(a)

(I4:500mA/ div、Time：2ms/ div)



(b)

(VO1:50V/ div、I4:500mA/ div、Time：100ms/ div)



(c)

(VO1:50V/ div、I4:500mA/ div、Time：2m s/ div)

Fig 10 output waveform;(a)base current output 2A;(b)two

LLC output waveform(c) low pulse high peak current output

6. Conclusion

Design a parallel LLC circuit, and complete the basic wave circuit and low pulse high peak current output, with four Sic Mosfet. The output waveform still needs to be CLOSER to the simulation output and connected to The EDM machine to teste for electricity consumption and the processing speed.

7. References

1. Y. K. Wang and B. Y. Song, Research on Highly Efficient EDM Pulse Power Supply and Its Experiments (Trans Tech, Harbin, China, 2005).
2. R. L. Steigerwald, A comparison of half-bridge resonant converter topologies, (IEEE Trans. Power Electron., vol. 3, pp. 174–182, Apr. 1988)
3. W. G. Cole et al., Electron discharge machining apparatus and method,"(U.S. Patent 5 543 599, 1996).
4. Rosario Casanueva, High-Frequency Power Supply for Spark Erosion, (IEEE TRANSACTIONS ON POWER ELECTRONICS, 2005).
5. M. Yan, Y. Liu, Design, analysis and experimental study of a high frequency power supply for finish cut of wire-EDM (International Journal of Machine Tools and Manufacture, vol. 49, 2009) pp. 793–796.
6. Infineon, Improving Efficiency of Synchronous Rectification by Analysis of the MOSFET Power Loss Mechanism (Application Note, June 2009).
7. Infineon, Resonant LLC Converter: Operation and Design (Application Note, Step 2012).
8. Carl Odulio, ‚Design of an LLC resonant converter for EDM applications(IEEE, Auckland, New Zealand, 09 February 2017)
9. A. Looser, L. Linares, Novel power supply topology for large working gap dry EDM(IEEE, Sapporo, Japan, 09 August 2010)

Networking Integration and Monitoring System with CANopen Controller for Intelligent Production Line of Tool Machine

Chau-Chung Song[a,b*], Yu-Wei Ho[b,c], Chen-Pang Chen[b,c] and Yu-Kai Chen[a]

[a]*Department of Aeronautical Engineering,*
National Formosa University, Huwei, Yulin, Taiwan
[b]*Smart Machine and Intelligent Manufacturing Research Center,*
Nation Formosa University, Huwei, Yunlin, Taiwan
[c]*Graduate Institute of Aeronautical and Electronic Technology,*
National Formosa University, Taiwan, R.O.C.
E-mail: ccsong@nfu.edu.tw, 10678109@gm.nfu.edu.tw, ykchen@nfu.edu.tw, smileplayer0806@hotmail.com*

Abstract

In this paper, the development of CANopen node controller and the integration of intelligent production line is focused and studied by applying the CANopen protocol. While the basic network management and data transmission is followed by the dictionary objects in the CANopen CiA301 protocol, the controller needs to implement with the motor motion control specification subjected to the CiA402 protocol. The SCADA system communicates with each node through CAN Bus to complete the task of constructing and managing the entire network.

Keywords: Production Line, CANopen, Networking Integration, Tool Machine

1. Introduction

In the industrial system, the numbers and categories of electric circuit and controller contained are becoming more and more complicated. The control network becomes the key technology in the industrial system. Controller Area Network(CAN), this communication bus has the advantages of high noise immunity, fast transmission speed and easy installation. CANopen in Automatic(CiA) organization released CANopen communication protocol for CAN application layer. The basic CANopen device and communication profiles are given in the CiA301 specification released by CiA. The Monitor System is mainly responsible for the controlling the nodes states and function in the network and displaying the data posted back from the nodes for the convenience of our testing whole set of network system platform. The tool machine operation control of production line is tested and realized to evaluate and verify the feasibility and performance of CANopen networking system.

2. Production Line Communication System

2.1. *CANopen*

International Organization for Standardization(ISO) series specifies the data link layer and physical of the CAN in Open Systems Interconnection(OSI) model shown in Figure 1. Many manufacturers define CAN frame by themselves for their products. But, when different manufacturers need to integrate the system, it often waste time and costs more. Therefore, the biggest intention of CANopen is to unify the definition CAN application layer.

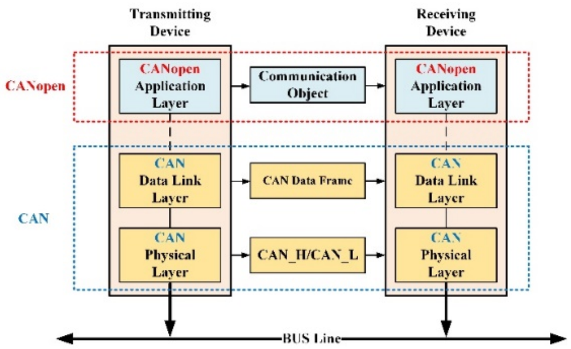


Fig. 1. CANopen Structure

2.1.1. *Communication Object(COB)*

The CANopen standard frame divides the 11-bit CAN-frame id into a 4-bit function code and 7-bit CANopen node ID. In CANopen the 11-bit id of a CAN-frame is known as communication object identifier, or COB-ID. COB-ID is shown in Table 1.

Table 1. Configurations of COB-ID

Object	Function Code	Resulting COB-ID
NMT	0000	0
SYNC	0001	80h
TIME STAMP	0010	100h
EMERGENCY	0001	81h-FFh
TPDO1	0011	181h-1FFh
RPDO1	0100	201h-27Fh
TPDO2	0101	281h-2FFh
RPDO2	0110	301h-37Fh
TPDO3	0111	381h-3FFh
RPDO3	1000	401h-47Fh
TPDO4	1001	481h-4FFh
RPDO4	1010	501h-57Fh
SDO(tx)	1011	581h-5FFh
SDO(rx)	1100	601h-67Fh
NMT Error Control	1110	701h-77Fh

CANopen according to communication feature divides communication feature into four several show below:

(i) Special Function Objects
  - Synchronization Objects
  - Emergency Objects
  - Time Stamp Objects

(ii) Network Management Objects(NMT)
(iii) Process Data Objects(PDO)
(iv) Service Data Objects(SDO)

2.1.2. *Object Dictionary(OD)*

CANopen devices must have an object dictionary, which is used for configure device parameter and communcaiton with the device.The complete OD consists of the six columns shown below:

(i) Index : 16-bits object address
(ii) Object name : Particular index within the OD.
(iii) Name: Simple textual description of the function of that particular object
(iv) Type: Information as to type of the object.
(v) Attribute: Defines the access rights for a particular object
(vi) M/O: Defines whether the object is Mandatory or optional.

2.2. *CANopen Communication Object*

2.2.1. *Network Management Object(NMT)*

In CANopen, the network management objects (NMT) mainly focuses the network states of monitoring nodes as well as the function like the control of node operation. In Figure 2 the network statement of a device is shown.

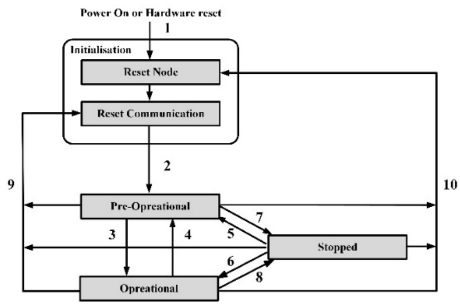


Fig. 2. NMT State Machine

Communication objects may only be executed if the devices involved in the communication are in the appropriate communication states. Table 3 shows the communication objects and states.

Table 3. States and Communication Objects

	Initialization	Pre-Operation	Operation	Stopped
PDO			√	
SDO		√	√	
SYNC		√	√	
Time Stamp		√	√	
EMCY		√	√	
Boot-up	√			
NMT		√	√	√

2.2.2. *Service Data Objects(SDO)*

According to data transfer direction divided into upload and download. Data frame of SDO consists of COB-ID, length and 2 to 8 Bytes data. Within 8 Bytes, Byte0 is Command Specifier(CS) which is data length of the object, Byte 1 and Byte 2 are index of the object within OD, Byte 3 is sub-index of the object, Byte 4 to byte 7 are parameter of the object. SDO frame is shown in Figure 3.

COB-ID	Length	Byte0	Byte1	Byte2	Byte3	Byte4~7
SDO	2~8	CS	Index		Sub-Index	Data

Fig. 3. CANopen Frame

2.2.3. *Process Data Objects(PDO)*

PDO transmission priority is higher than SDO, so PDO is often used to transfer data in real time. PDO divided into Transmit Process Data Object(TPDO) and Receive Process Data Object(RPDO). TPDO is transmitted from the node to the master. In contrast, RPDO is the node to receive messages from the master. Before using PDO, set PDO Communication Parameter and PDO Mapping Parameter through SDO to complete mapping.

2.3. *Device and Motion Control of CANopen*

CiA released CiA-402 protocol, it is the standardized CANopen device profile for digital controlled motion products. Establish device state machine for manage device. The state machine describes the device status and the possible control sequence of the drive shown in Figure 4. Within OD, Controlword(index:6040h) could switch device status. Statusword(index:6041h) could stored device current status.

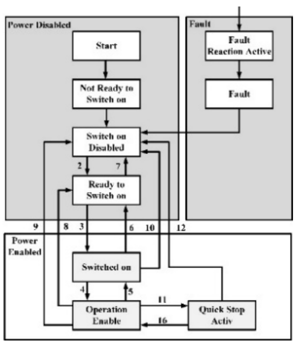


Fig. 4. Device State Machine

3. System Development and Design

3.1. *System Networking Structure*

CANopen network system structure designed is shown in Figure. The network main has one monitoring mode and two tool machine(six motor node). The main function of monitoring node is to monitor the whole CANopen network states. All equipment noodes possesses CANopne protocol to establish the whole set of CANopen network system.

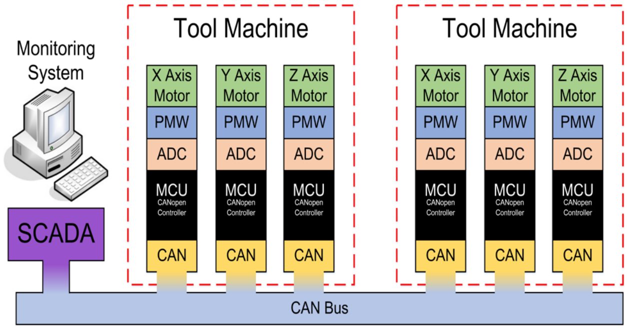


Fig. 5. Network System Structure

3.2. *CANopen Controller Design*

The design of the controller firmware to be written under the CANopen standard agreement, CANopen communication applications are mainly based on CiA-301 specification, which describes NMT, SDO and PDO. This proposed controller needs to implement with the motor motion control specification subjected to the CiA402 protocol. It is also additionally equipped with a motor-specific incremental encoder to receive the motor feedback signal, so as to realize the motor speed control and high-precision positioning control.

3.3. *Monitoring System*

The Supervisory Control and Data Acquisition (SCADA) system is the main monitoring of the Bus, which monitors the entire CANopen and the network status of each node, and displays the real-time CAN data below the SCADA System interface. Its function can be divided into four tabs, Initialize, NMT, Line, Figure 6 for the interface of SCADA System.

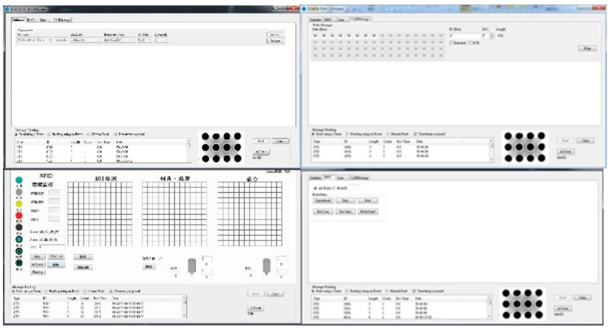


Fig. 6. Monitoring Interface: SCADA System

4. System Integration and Application

4.1. *Application field of production line*

Figure 7 is shown system structure of the production line. Figure 8 is shown the application architecture of this production line system consists of two tool machines and a set of handling robot. The controllers of the two tool machines are operated in accordance with the CiA protocol. The handling robot is controlled by I/O.

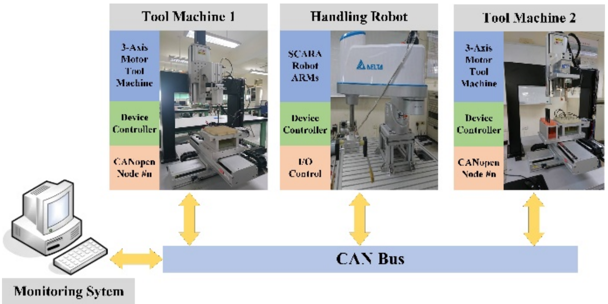


Fig. 7. System structure of production line

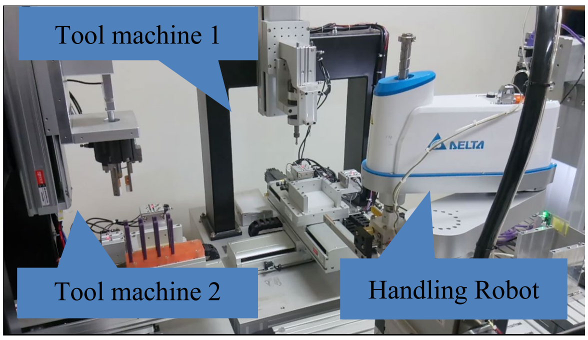


Fig. 8. Application field of production line

4.2. *Networking Integration of System Platform*

Turn on the power of each tool machines and use the SCADA System to initialize the CANopen network. After confirming that the CANopen network is connected, the NMT state will switch from Boot-up to Initialization and then to Pre-Operational. During the NMT state is Pre-Operational, then set the PDO parameter of each node. After setting, switch the NMT state to Operational. When the NMT state enters Operational, it can start to control motion device.

5. Conclusion

This paper is mainly to explore the industrial network. It is hoped that this set of network system platform could be applied in the industrial automation and make use of the CAN and CANopen features to make the industrial network control system to be more completed. The networking communication and motor control is developed for the motor controller node with CANopen. The motor controllers are provided with basic networking management and data transfer functions of CANopen. Additionally, equipped with the incremental encoder to receive the motor feedback, so as to realize the motor speed control and positioning control. Finally, with the actual application of the tool machine, the monitor system issues the commands that follow the protocol. Show the results of the combination of the CANopen local network and industrial automation.

6. References

1. Renesas Electronics Corporation, *RX62T Group, RX62G Group User's Manual: Hardware Rev.2.00*, Dec 27, 2013
2. CAN in Automatic, *CANopen Specification*, 1996
3. CiA-301, *CANopen Application layer and communication profile* 2002.
4. CiA-402, *CANopen Device Profile Drives and motion Control*, 2002.
5. AMP&MOONS' Automation, *CANopen User Manual,* 2013.
6. Yang Fan, Ran Chen, Zhi-hui Chen, *Model-based validation of CANopen systems,* France,5-7 May 2014
7. Mohammed Farsi and Manuel Bernardo Martins Barbosa, *CANopen Implementation: applications to industrial networks*, Research Studies PRESS LTD., 2000

Acknowledgements

This research was supported by the Ministry of Science and Technology, Taiwan, R.O.C. under Grant: MOST 107-2218-E-150-001 and 107-2218-E-150-007.

Illumination Manipulation and Specular Reflection Analysis of Still Image with Single Object

Hsuan T. Chang[1,*] Wei-Hsiu Chang

Photonics and Information Laboratory, Department of Electrical Engineering, National Yunlin University of Science and Technology, 123, Section 3,University Road, Douliu, Yunlin 64002, Taiwan

[1]E-mail: htchang@yuntech.edu.tw

Abstract

A lighting manipulation scheme for still images is proposed in this paper. By using the dichromatic-based model, the reflection component can be resolved through a single-image specular reflection removal method with the characteristics of color constancy. Finally virtual illumination can be generated through the recombination of the reflection component. The scenes illumination can be estimated using the proposed automatic method without knowing the illumination spectra, three-dimensional object modeling, or texture databases. Experimental results show that the method is useful for handling the single or multicolor objects in scenes.

Keywords: Chromaticity, Dichromatic-based model, specular removal, color consistence, virtual illumination.

1. Introduction

Establishing the relighting in the computer graphics needs the provision of the light sources from different directions. Using only one single input image is not enough to complete the relighting of the objects in the image. In this paper, we propose a method which can modify the illumination conditions in images. By separating the specular reflection and automatically evaluating the illumination in the image, the reconstruction of 3-D scene can be discarded and we can perform realistic control on illumination colors in images.

In recent research about specular highlight removal [1], most methods on separating the specular reflection components in a single image are based on the reflection model established by Shafer [2]. The reflection model assumes the complex light paths as a simple reflection component and analyzes the highlight information based on polarized photography. Thereafter, more related methods have been proposed. For example, the T-shape image color space proposed by Shafer and Klinker can be used to analyze the neighboring pixels of an object [3]. In this method, the reflected specular and diffused colors are viewed as orthogonal color vectors. However, in the bright image textures, the linear distribution model is hardly used to estimate the vectors in the color space.

The specular-free image (SFI) method determines the diffused components by estimating either the intensity or chromaticity of the image. Tan's method [4] generates the SFI by analyzing the maximum chromaticity space at first. In this method, the specular reflection is considered as the chromaticity deviation and the maximum chromaticity is set as the pixel common chromaticity. When the maximum chromaticity is extracted and discarded, the specular component can be removed. However, in a non-single color image, different maximum chromaticity in the chromaticity space for different textures could be observed. Therefore, the controversy exists on selecting the chromaticity. Yoon's specular-free two-band image (SFTBI) method [5], which uses the specular-invariant to determine the diffused reflection component, is thus proposed. A diffused reflection image with low chromaticity can be obtained. In addition, Shen proposed the modified specular-free image (MSFI) [6] method to improve Yoon's method.

The SFI method is usually used together with the neighboring region analysis. Although the SFI method usually cannot get precise component analysis during the separation process, SFI can easily obtain the result without specular components. The specular reflection pixels can be found by analyzing the difference between the SFI and the intrinsic image in the neighboring region. In Tan's method [4], the iteration process is used at low chromaticity range for both the SFI and normalized image. The high-quality or bright-color texture images can be successfully processed by setting regional chromaticity. Yoon and Shen proposed different mechanisms to improve the iteration efficiency. Other types of methods are based on the dichromatic reflection model. For example, Ping's inpainting method [7], which repairs the highlight points (also considered as region of interest, ROI). Rouf proposed the star filter to filter the highlight points based on the structural characteristics [8].

In this paper, the proposed method mainly deals with the color constancy based on the chromaticity and integrated the methods of high-light removal proposed by Tan and Shan [4], [6] to achieve the reflection removal purpose.

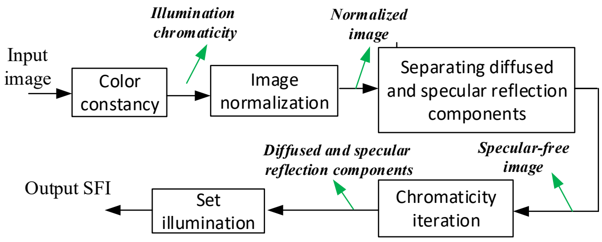


Fig. 1: The system diagram of the proposed method.

2. Proposed Method

Figure 1 shows the system diagram of the proposed method. The proposed illumination estimation method is based on the Ikeuchi's inverse-intensity chromaticity space [9]. In this space, the specular component will be labeled and then be transferred to the Hough space to statistically determine the corresponding chromaticity direction. Illumination estimation mainly aims to provide the correct intensity information during the separation of reflection components. The intensities of non-white light illuminations are inhomogeneous while changing the intensity. Therefore, we assume that the illumination transformation is correctly estimated using the same compensation mechanism.

The dichromatic-based model is used to determine the diffused reflection components. The SFI can be obtained using Tan's method [9]. Then the improved SFI, called the simplified mechanism of SFI can achieved. First, Tan's method mentioned that the image intensity can be divided into two parts according to the dichromatic-based model. That is,

$$\mathbf{I}(\underline{\mathbf{x}}) = w_d \int_{\Omega} S(\lambda)E(\lambda)\mathbf{q}(\lambda)d\lambda + w_S \int_{\Omega} E(\lambda)\mathbf{q}(\lambda)d\lambda, \quad (1)$$

where $\mathbf{I}(\underline{\mathbf{x}}) = \{I_r, I_g, I_b\}$ denotes the color vector of image intensity recorded by a camera. $\underline{\mathbf{x}}=\{x,y\}$ denotes the 2D coordinates, $\mathbf{q} = \{q_r, q_g, q_b\}$ denotes the 3D vector of sensor sensitivity. $S(\lambda)$ and $E(\lambda)$ denote the diffused and the illumination spectral distributions, respectively. w_d and w_S denote for the weighting factors of specular and diffuse reflection, respectively. The values depend on the geometric structure in the regional coordinate $\underline{\mathbf{x}}$. Image intensity is within the visible spectrum (Ω). As shown in Eq. (1), the dichromatic-based model describes the image pixel intensity by separating them into two reflection components: $w_d \int_{\Omega} S(\lambda)E(\lambda)\mathbf{q}(\lambda)d\lambda$, which denotes the diffused reflection components, and $w_S \int_{\Omega} E(\lambda)\mathbf{q}(\lambda)d\lambda$, which denotes the specular reflection components in an image. Diffused component includes the reflection spectrum of the object, while the specular one just depends on the illumination. The specular and diffused reflection components are two independent color vectors, which can be expressed as two bold-face characters, **B** and **G**, respectively. We also assume that the specular components are uniformly distributed in the scene so that the illumination colors are independent to the regional coordinates. The intensity equation of the original image can be replaced a simplified combination:

$$\mathbf{I}(\underline{\mathbf{x}}) = w_d\mathbf{B}(\underline{\mathbf{x}}) + w_S\mathbf{G}, \quad (2)$$

where
$$\mathbf{B}(\underline{\mathbf{x}}) = \int_{\Omega} S(\lambda)E(\lambda)\mathbf{q}(\lambda)d\lambda, \quad (3)$$

$$\mathbf{G} = \int_{\Omega} E(\lambda)\mathbf{q}(\lambda)d\lambda. \quad (4)$$

In order to analysis the relationship between the reflection components and scaling factor. The image chromaticity is defined by using the normalized RGB components. Here, the chromaticity (σ), diffuse chromaticity (Λ), and specular chromaticity (Γ) are defined as

$$\boldsymbol{\sigma}(\underline{\mathbf{x}}) = \frac{\mathbf{I}(\underline{\mathbf{x}})}{I_r(\underline{\mathbf{x}})+I_g(\underline{\mathbf{x}})+I_b(\underline{\mathbf{x}})}, \quad (5)$$

$$\boldsymbol{\Lambda}(\underline{\mathbf{x}}) = \frac{\mathbf{B}(\underline{\mathbf{x}})}{B_r(\underline{\mathbf{x}})+B_g(\underline{\mathbf{x}})+B_b(\underline{\mathbf{x}})}, \quad (6)$$

$$\boldsymbol{\Gamma} = \frac{\mathbf{G}}{(G_r+G_g+G_b)}, \quad (7)$$

Substituting Eqs. (5) and (6) into Eq. (2), the chromaticity equation becomes

$$\mathbf{I}(\underline{\mathbf{x}}) = m_d(\underline{\mathbf{x}})\boldsymbol{\Lambda}(\underline{\mathbf{x}}) + m_s(\underline{\mathbf{x}})\boldsymbol{\Gamma}, \quad (8)$$

where
$$m_d(\underline{\mathbf{x}}) = w_d\{B_r, B_g, B_b\}, \quad (9)$$

$$m_s(\underline{\mathbf{x}}) = w_s\{G_r, G_g, G_b\}. \quad (10)$$

In the reflection model shown in Eq. (1), in which only the specular reflection component exists (w_d=0), $\boldsymbol{\Gamma}$ will be independent of the specular geometrical parameter w_S. In the definition on chromaticity, the range of image chromaticity, diffuse chromaticity, and specular chromaticity are all within the range {0, 1}. That is, $\{\sigma_r+\sigma_g+\sigma_b\} = \{\Lambda_r+\Lambda_g+\Lambda_b\} = \{\Gamma_r+\Gamma_g+\Gamma_b\} = 1$. The illuminant chromaticity is estimated based on the color constancy method in Ref. [2] to evaluate the chromaticity $\boldsymbol{\Gamma}^{\text{est}} = \{\Gamma_r^{\text{est}}, \Gamma_g^{\text{est}}, \Gamma_b^{\text{est}}\}$. Assume that the evaluated chromaticity is correct. The image with corrected illumination is defined as a normalized image:

$$\mathbf{I}'(\underline{\mathbf{x}}) = m_d'(\underline{\mathbf{x}})\mathbf{\Lambda}'(x) + m_s'(\underline{\mathbf{x}})\tfrac{1}{3}, \quad (11)$$

where
$$m_d(\underline{\mathbf{x}})\mathbf{\Lambda}(\underline{\mathbf{x}}) = [m_d'(\underline{\mathbf{x}})\mathbf{\Lambda}'(\underline{\mathbf{x}})]\mathbf{\Gamma}^{\mathrm{est}}, \quad (12)$$
$$m_s(\underline{\mathbf{x}})\mathbf{\Gamma} = \left[m_s'(\underline{\mathbf{x}})\tfrac{1}{3}\right]\mathbf{\Gamma}^{\mathrm{est}}. \quad (13)$$

The normalized image can be denoted as $\mathbf{I}'(\underline{\mathbf{x}}) = \frac{\mathbf{I}(\mathbf{x})}{\mathbf{\Gamma}^{\mathrm{est}}}$. The normalized illumination color is $\frac{\mathbf{\Gamma}}{\mathbf{\Gamma}^{\mathrm{est}}} = \{1,1,1\}$. In $\mathbf{I}'(\underline{\mathbf{x}})$, $\mathbf{\Gamma}' = \{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\}$ and $3m_s = m_s'$. Normalized image will be considered as a diffused reflection image illuminated by a standard white light.

Since Tan's specular-to-diffuse mechanism does not generate a real diffuse image, the image can be seemed as a fake diffused reflection image. In this mechanism, the maximum chromaticity is used. The definition of maximum chromaticity of image pixels is as follow

$$\tilde{\sigma}'(\underline{\mathbf{x}}) = \frac{\max(I'_r(\underline{\mathbf{x}}), I'_g(\underline{\mathbf{x}}), I'_b(\underline{\mathbf{x}}))}{I'_r(\underline{\mathbf{x}})+I'_g(\underline{\mathbf{x}})+I'_b(\underline{\mathbf{x}})}, \quad (14)$$

where $(I'_r(\underline{\mathbf{x}}), I'_g(\underline{\mathbf{x}}), I'_b(\underline{\mathbf{x}}))$ is obtained in the normalized image, σ' denotes the chromaticity of the normalized image. Note that $\tilde{\sigma}'$ is different from σ' and the range of $\tilde{\sigma}'$ is not between 0 and 1. With Eqs. (11) to (13), the maximum chromaticity of image pixels in Eq. (14) can be rewritten as:

$$\tilde{\sigma}'(\underline{\mathbf{x}}) = \frac{m_d'(\underline{\mathbf{x}})\tilde{\Lambda}'(\underline{\mathbf{x}}) + m_s'(\underline{\mathbf{x}})\frac{1}{3}}{m_d'(\underline{\mathbf{x}})[\Lambda_r'(\underline{\mathbf{x}}), \Lambda_g'(\underline{\mathbf{x}}), \Lambda_b'(\underline{\mathbf{x}})] + m_s'(\underline{\mathbf{x}})}.$$

By setting $\tilde{\Lambda}' = \max(\Lambda'_r, \Lambda'_g, \Lambda'_b)$, both m_d' and m_s' can also be determined. Since the maximum chromaticity of diffused reflection is usually higher than that of specular reflection and, in generally, $\tilde{\Lambda}' > \frac{1}{3}$, and

$$\tilde{\sigma}'_{\mathrm{diff}} > \tilde{\sigma}'_{\mathrm{spec}}, \quad (15)$$

$$\frac{\tilde{\Lambda}'}{[\Lambda_r', \Lambda_g', \Lambda_b']} > \frac{m_d'\tilde{\Lambda}'(\underline{\mathbf{x}}) + m_s'\frac{1}{3}}{m_d'[\Lambda_r', \Lambda_g', \Lambda_b'] + m_s'}. \quad (16)$$

In the normalized image, $(\Lambda_r' + \Lambda_g' + \Lambda_b') = 1$. Removing the pixel $(\underline{\mathbf{x}})$ and substituting m_s' ($m_s' = m_d'(\frac{\tilde{\Lambda}' - \tilde{\sigma}'}{\tilde{\sigma}' - \frac{1}{3}})$) into Eq. (16), we can finally obtain the equation that can represent the relationship between the image chromaticity and illumination chromaticity:

$$\tilde{I}'(\underline{\mathbf{x}}) = m_d'\left(\tilde{\Lambda}' - \tfrac{1}{3}\right)\left(\frac{\tilde{\sigma}'}{\tilde{\sigma}' - \frac{1}{3}}\right). \quad (17)$$

In the above equation, m_d' is computed by assuming the same chromaticity. In the specular pixel $(\underline{\mathbf{x}}_1)$ and diffused pixel $(\underline{\mathbf{x}}_2)$, the same chromaticity is used such that $\tilde{\Lambda}'(\underline{\mathbf{x}}_1) = \tilde{\Lambda}'(\underline{\mathbf{x}}_2) = \tilde{\sigma}'(\underline{\mathbf{x}}_2)$ ($m_s' = m_d'$ here). We can determine m_d' by using:

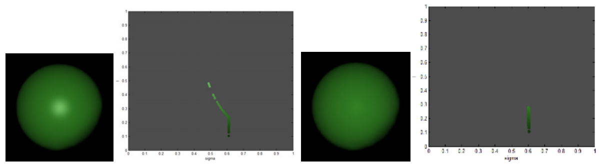
$$m_d'(\underline{\mathbf{x}}_1) = \frac{\tilde{I}'(\underline{\mathbf{x}}_1)[3\tilde{\sigma}'(\underline{\mathbf{x}}_1) - 1]}{\tilde{\sigma}'(\underline{\mathbf{x}}_1)[3\tilde{\Lambda}'(\underline{\mathbf{x}}_1) - 1]}. \quad (18)$$

The finally result is

$$m_s'(\underline{\mathbf{x}}_1) = [I'_r(\underline{\mathbf{x}}_1) + I'_g(\underline{\mathbf{x}}_1) + I'_b(\underline{\mathbf{x}}_1)]m_d'(\underline{\mathbf{x}}_1),$$

where the diffused reflection component is

$$m_d'(\underline{\mathbf{x}}_1)\tilde{\Lambda}'(\underline{\mathbf{x}}_1) = \tilde{I}'(\underline{\mathbf{x}}_1) - \frac{m_s'(\underline{\mathbf{x}}_1)}{3}. \quad (19)$$



(a) (b) (c) (d)

Fig. 2: (a) The original image; (b) Projection of pixel intensity in (a) into the chromatic space ($\tilde{\sigma}'$, $\tilde{I}'$); (c) The SFI with the maximum chromaticity; (d) The projection of pixel intensity in (c) into the chromatic space ($\tilde{\sigma}'$, $\tilde{I}'$).

Figures 2(a) and 2(b) show an input image and its 2D projection of the maximum chromatic intensity $\tilde{\sigma}'$ and intensity $\tilde{I}'$, respectively. In Fig. 2(b), the *x* axis denotes the maximum chromaticity intensity $\tilde{\sigma}'$, while the *y* axis denotes the intensity $\tilde{I}'$. Fig. 2(c) shows the SFI obtained by suing Tan's method with setting the $\tilde{\Lambda}' = \max(\Lambda'_r, \Lambda'_g, \Lambda'_b)$ in the maximum chromaticity for all image pixels. In Fig. 2(d), the SFI specular component has the same chromaticity with the diffused one.

Yoon proposed a simplified SFI method based on the property of invariant specular reflection, whose mechanism is shown as follows:

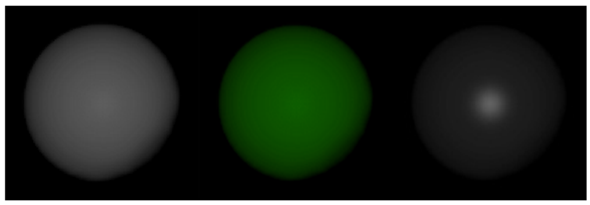
$$\mathbf{I}_{sf}(\underline{\mathbf{x}}) = \mathbf{I}'(\underline{\mathbf{x}}) - \mathbf{I}_{min}(\underline{\mathbf{x}}), \quad (20)$$
$$\mathbf{I}_{min}(\underline{\mathbf{x}}) = m_d\,\lambda_{min}(\underline{\mathbf{x}}) + \tfrac{1}{3}m_s'(\underline{\mathbf{x}}), \quad (21)$$

where $\mathbf{I}_{sf}$ presents SFI, $\mathbf{I}_{min}$ presents the minimum pixel value of the normalized image in RGB channels. Similarly, we assume that the specular reflection is also invariant in the HSV space:

$$\mathbf{I}_{min}(\underline{\mathbf{x}}) = \begin{cases} S = 0 \\ V = \min(I_r, I_g, I_b) \end{cases}. \quad (22)$$

Different from the Tan's method, we use the HSV color space to determine the SFI, which is shown in Fig. 3(b). The S and V components present the saturation and scalar intensity values, respectively. And the specular and diffused components of chromaticity are the same with that in the input image. By subtracting the change of specular component, the intensity of SFI shown in Fig. 3(a) will be a uniform intensity. Fig. 3(c) shows the invariant specular reflection part of the original image. The SFI component is determined as follows:

$$\mathbf{I}_{sf}(\underline{\mathbf{x}}) = m_d\left(\lambda(\underline{\mathbf{x}}) - \lambda_{min}(\underline{\mathbf{x}})\right). \quad (23)$$

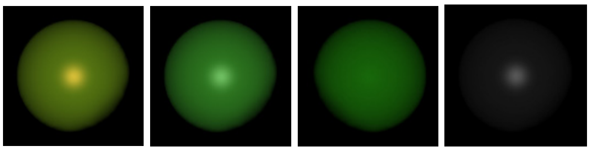


(a) (b) (c)

Figure 3: (a) intensity of SFI; (b) SFI; (c) invariant specular

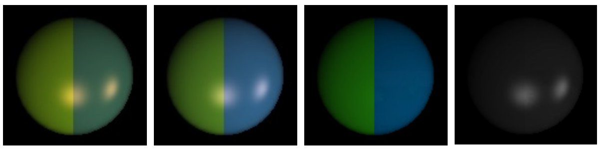
part $\mathbf{I}_{min}$.

The illumination setting for the final output SFI image is the last step. The user can define a new reflectivity by setting $\sum_{u\in\{r,g,b\}} \mathbf{\Gamma}^{\text{set}}$. That is, the separating specular and diffused reflections parts can be edited for the final combination. Finally, by replacing the original chromaticity with the new one $\mathbf{\Gamma}^{\text{set}} = \{\Gamma_r^{\text{set}}, \Gamma_g^{\text{set}}, \Gamma_b^{\text{set}}\}$, the final SFI image can be obtained.



(a) (b) (c) (d)

Fig. 5: (a) Image of a spherical object illuminated by a halogen lamp; (b) Normalized image with chromaticity RGB = (0.5324, 0.3077, 0.1594); (c) Separated diffused component; (d) Separated specular component.



(a) (b) (c) (d)

Fig. 6: (a) Dichromatic image of a spherical object illuminated by a halogen lamp; (b) Normalized image with chromaticity RGB =(0.2187, 0.2695, 0.7812); (c) Separated diffused component; (d) Separated specular component.

3. Experiment Results

Suppose that the reflection light is generated from a fixed illumination in a scene. The proposed system can automatically detect the light source and then perform image normalization so that the image can be separated into the normalized image, diffused image, and specular image. By setting the chromaticity, the user can obtain the image with a modified illumination color. Figures 5 and 6 provide two demonstrations of the proposed method on resetting the chromaticity. Figure 5(a) shows an input image, in which the single-color spherical object is illuminated by a halogen lamp with the color temperature 4700 °K. Figure 5(b) shows the normalized image using the designated illumination chromaticity RGB = (0.5324, 0.3077, 0.1594). Figures 5(c) and 5(d) show the extracted diffused and specular components, respectively. There are two colors in the spherical object shown in Fig. 6(a). The illumination is the same as that in Fig. 5(a). However, the designated illumination chromaticity RGB= (0.2187, 0.2695, 0.7812) in Fig. 6(b) is different from that in Fig. 5(b). By using the proposed method, Figs. 6(c) and 6(d) show the separated diffused and specular components, respectively.

4. Conclusion

We propose a method to simulate the object images under various illumination colors without constructing the 3-D model of the scene or requiring multiple input images. In addition, the reflectivity can also be assigned and be combined with the separated diffused components to obtain the designated image. The improved mechanism of SFI is used to accelerate the iteration process. Given an input image, the system will separate the specular and reflection components at first. Thus the high-light removal can be achieved. A graphical user interface has been implemented so that the users can easily manipulate the system to obtain the images with various illumination conditions such as the chromaticity and reflectivity.

Acknowledgements

This research is partially supported by Ministry of Science and Technology, Taiwan under the contract number MOST 108-2221-E-224-039.

References

1. Artusi, F. Banterle, and D. Chetverikov, “A survey of specularity removal methods,” Computer Graphics Forum, vol. 30, no. 8, pp. 2208-2230, 2011.
2. S.A. Shafer, “Using color to separate reflection components,” Color Research and Applications, vol. 10, pp. 210-218, 1985.
3. G.J. Klinker, S.A. Shafer, and T. Kannade, “A physical approach to color image understanding,” International Journal of Computer Vision, vol. 4, no. 1, 1990.
4. R.T. Tan and K. Ikeuchi, “Separating reflection components of textured surfaces using a single image,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 2, pp. 178-193, 2005.
5. K.J. Yoon, Y. Choi, and I.S. Kweon, “Fast separation of reflection components using a specularity-invariant image representation,” 2006 IEEE International Conference on Image Processing (ICIP), pp. 973-976, 2006.
6. H.L. Shen and Q.Y. Cai, “Simple and efficient method for specularity removal in an image,” Applied Optics, vol. 48, pp. 2711-2719, May 2009.
7. T. Ping, S. Liny, L. Quanz, and H.Y. Shum, “Highlight removal by illumination-constrained inpainting,” Ninth IEEE International Conference on Computer Vision, vol. 1, pp. 164-169, 2003.
8. M. Rouf, R. Mantiuk, W. Heidrich, M. Trentacoste, and C. Lau, “Glare encoding of high dynamic range images,” 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 289-296, 2011.
9. R. T. Tan, K. Ikeuchi, and K. Nishino, “Color constancy through inverse-intensity chromaticity space,” Journal of the Optical Society of America, A, vol. 21, no. 3, pp. 321-334, 2004.

Robotics Education for the 2019 MakeX Robotics Competition

Jia-Ming Hsiao*
Dep. Of Multimedia and Game Development Management,
Far East University, Taiwan, Republic of China

E-mail: hsiaojm@mail.feu.edu.tw
www.feu.edu.tw

Abstract

This paper is proposed to illustrate the robotics education experience for 2019 MakeX Robotics Competition in Taiwan. It is promoted in the class to make use of multidisciplinary learning within the fields of science and technology. Students are passionate about innovation by engaging them in exciting Robotics Competition. The spirit of creativity, teamwork, fun and sharing is appreciated in the competition. Alliance competition and cooperation strategy in competition is another topic in teaching and learning.

Keywords: robotics education, multidisciplinary learning, MakeX.

1. Introduction

MakeX robotics competition is held by Makeblock corporation to promote multidisciplinary learning within the fields of science and technology. The spirit of creativity, teamwork, fun and sharing is appreciated in the competition[1]. The theme of 2019 season is "My City" and there are four themed events including Spark, Starter, Challenge and Premier. MakeX Starter is a multitasking and alliance collaborative competition for students between 6 to 16 years old. MakeX Challenge is a confrontational competition between the Red and Blue camps for teenage between 11 to 18 years old. In 2019 MakeX Taiwan contest, the Starter and Challenge events are organized to pick the qualified teams who compete for a prize of MakeX world finals. This paper is proposed to illustrate and share the education experience subjected to the 2019 season of MakeX robotics competition. The advantage and disadvantage of MakeX robotics competition are also discussed.

2. The Motivation of Participating MakeX

There are many robotics competition held in Taiwan every year. The robot platform may be designated in some competitions like WRO (World Robot Olympiad), FLL (FIRST LEGO League) and VEX Robotics Competition. MakeX is a brand new robotics competition in Taiwan since 2018. MakeX is a themed robotics competition similar to WRO and FLL with the platform of Makeblock products. The themed robotics competition is better than the non-themed one in the viewpoint of education. The students and teachers are motivated by the theme of robotics competition to explore the possibility of applying robot to solve practical problems. The multidisciplinary learning concerned with STEAM (Science, Technology, Engineering, Art and Mathematics) is committed to inspire the young students to explore, design and solve real world problems. The competition missions can be transformed to courses to develop problem solving strategies through the

*Address: No.49, Zhonghua Rd., Xinshi Dist., Tainan City 74448, Taiwan (R.O.C.).

realization of electronic devices, mechanical structures and software programming, as well as alliance communication and collaboration. In the other hand, the total price of Makeblock robot platform is lower than LEGO products, and is more affordable for rural schools.

3. Teaching and Training

The competition missions are briefly described in this section with respect to MakeX Starter and MakeX Challenge separately.

3.1. *MakeX Starter*

The theme of 2019 MakeX Starter is "City Guardian". It is aimed at environment protection, green energy and home guardian. The designated age ranges of team members are correspond to elementary school students and junior high school students in Taiwan. There are two parts which are automatic stage and manual stage in MakeX Starter. In the automatic stage, six independent missions M01-M06 and three alliance missions M07-M09 are included. The robot should perform individually the following abilities:

(i) Line following
(ii) Detecting obstacle
(iii) Color identification
(iv) Moving props
(v) Manual remote control

The Red team and Blue team are alliance partners to earn points as high as possible.

The mechanical structure of robot is designed to be switchable to fit the need of independent and alliance missions. Examples of mechanical structure of independent and alliance missions are shown in Fig. 1 and Fig. 2 respectively.

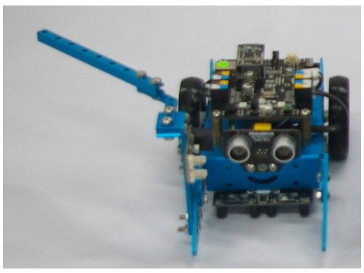


Fig. 1. One of mechanical structures of independent missions.

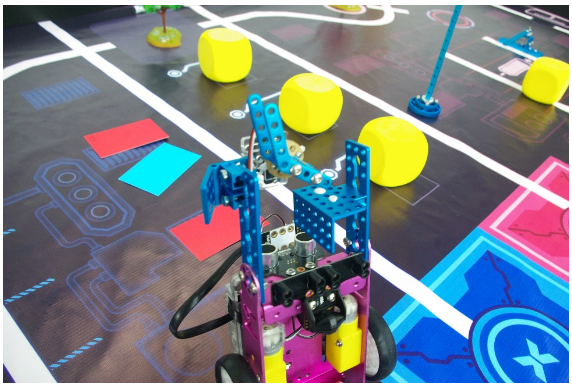


Fig. 2. One of mechanical structures of alliance missions.

Since the 2019 City Guardian mission is designed to avoid the human factor in automatic stage and the competition time is less than 2018 Blue Planet mission, the mechanical structure should be applied to as many missions as possible in automatic stage. The electronic devices include the mBot robot, color sensor, RGB line following sensor, ultrasonic sensor, LED matrix plate and a 9g servo motor. The control program consists of ten mission parts. Each mission program can be lunched by the times of pressing the on-board button and then covering the on-board light sensor of mCore control board.

Students are inspired and guided to show their innovative ideas about the mission completing strategies. The process of solving missions is shown as Fig. 3. Students are also required to communicate with alliance team to discuss how to complete the alliance missions as soon as possible. It is even necessary to assist alliance partners to modify mechanical structures and/or control program, i.e. sharing.

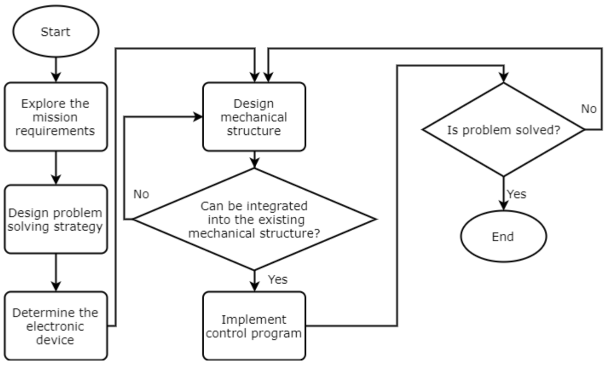


Fig. 3. Process of solving single mission

3.2. *MakeX Challenge*

The theme of 2019 MakeX Challenge is "Courageous Traveler". It is a confrontational competition corresponding to junior and senior high school students in Taiwan. Courageous Traveler is focusing on training and improving the students' ability of problem exploration, strategy analysis, structure design, software programming, precise remote control and communication with alliance. There are two camps called Red alliance and Blue alliance. Each camp is consisted of two teams. During the competition, the robot should perform

(i) Balls collection
(ii) Highland climbing
(iii) Bottle moving
(iv) Balls shooting (to attack bottles)
(v) Flag insertion

The mechanical structures are more complex than MakeX Starter "City Guardian". The robot chassis is designed and assembled by students to satisfy the competition missions. Students should need more knowledge and pay more patience, perseverance, and time to this competition. It is also required to construct the robot with finite parts and resource within the competition pack. The robot applied to 2019 MakeX Taiwan contest is shown in Fig. 4 and Fig. 5. The four omnidirectional wheels are attached to a smart encoder motor each, and arranged as "卍" shaped chassis structure since it provides simpler directional control[2]. The mechanism of collecting balls is realized by a DC motor and three gears with tracks tied by a number of nylon cables. The balls loading structure is similar to the balls collecting one. The function of shooting balls is implemented by two sets of two-stage gear train driven rubber wheel to rub balls by high rotation speed. The elevation angle of balls shooting mechanism is adjustable. The main control board, power expansion board, power management module, battery and Bluetooth module are integrated in a plate screwed on the vertical back side of robot. It is arranged to balance the weight of balls shooting structure in order to keep the center of gravity within the robot. If the center of gravity of robot is not proper, it will result in falling while climbing highland and non-smooth movement. The bottle moving function is implemented by attaching an angled beam to a servo motor. It is limited to clamp a standing bottle. Some teams attending the Courageous Traveler competition realize a claw mechanism to clamp both standing and falling bottles, therefore perform the defense function by picking and lifting falling bottles. The remote control function makes use of Bluetooth module and Bluetooth remoter. The function selections are programmed to press different remoter buttons.

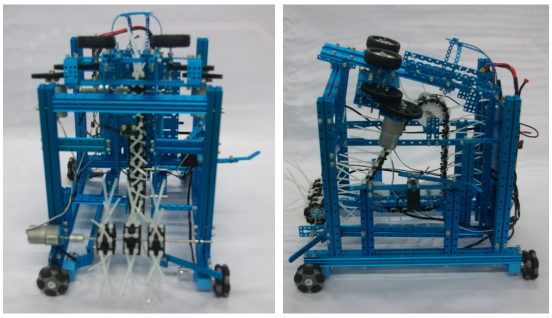


Fig. 4. Front and left side view of robot.

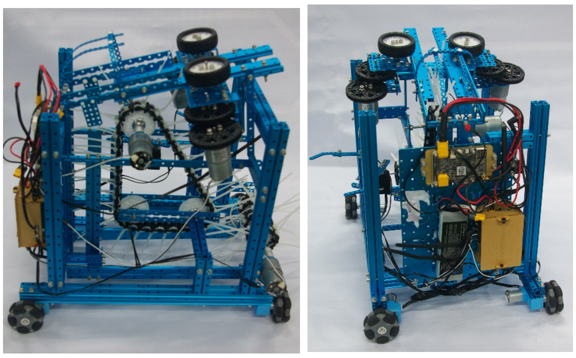


Fig. 5. Right side and back view of robot

Since the team members come from different senior high schools, it is more difficult to gather all team members to discuss the mechanism, strategy, software programming and alliance strategy. The meetings are taken on Saturday and Sunday. Limited by the parts and meeting time, the team implement simpler structure but approximately full function robot. Due to the improper structure of two-stage gear train and the distortion of plastic gears resulting from high torque, the distance of shooting balls is not long enough to attack the bottles of opposite camp. The competition strategy is then a defense type. The attacking attempt is performed by climbing highland and shooting the nearer bottles. The most attacking efforts are assigned to the alliance team. Our team passed the tournament and entered the finals. Before the finals, the alliance team attempt to improve the robot shooting function of ours, but it seemed to be not functional. Although the effort was not effective, the competition spirt of creativity, teamwork, fun and sharing is realized by these teenage students. The alliance teams wined the Third place of 2019 MakeX Taiwan contest in the end. Fig. 6 shows the competition situation of Courageous Traveler.

Fig. 6. The competition situation of Courageous Traveler.

4. The advantages and disadvantages of MakeX

Similar to other themed robotics competitions, there are advantages and disadvantages in the MakeX competition. These advantages are:

- A themed competition: encourage the team to explore the possibility and feasibility of applying robots to solve real world problems.
- Both competition and collaboration are valued: not only compete, but also collaborate.
- Alliance strategy: teams of some competitions in Taiwan just press the start button and then finished, there are more alliance strategies needed to think and communicate in MakeX.
- Team-working and sharing: the alliance missions are completed by two alliance teams, more team-working and sharing are required.
- Affordable price: the total price of participating is lower than FLL, FTC (FIRST Tech Challenge) and FRC (FIRST Robotics Competition) with the similar competition sprite.
- There are upgrade packs of robot expansion packs and competition arena packs.

The disadvantages are:

- The base-map can not be available separately. It comes with the whole competition arena pack.
- The quality of base-map could be more improved.
- The power of MakeX Starter robot is designated to the battery from 2019 Starter expansion pack. The battery from 2018 Starter expansion pack is not allowed. This results in the rise of price, and it is not reasonable and acceptable for some teams.

In summary, MakeX Starter is a suitable and meaningful themed robotics completion for the mBot robotics education. Each mission of MakeX Starter can be transformed into a teaching course. The training time and teacher's efforts can be reduced by the combination of these courses. MakeX Challenge is cheaper than FTC and FRC with the similar competition and collaboration sprite. MakeX robotics competition could be another choice of robotics education and competition for rural schools.

Acknowledgements

It is thankful for the assistance from JinYuan International Corporation, and IOP Robotic Technology Co.,Ltd. The training courses of MakeX Starter are supported by IOP Robotic Technology Co.,Ltd. The competition packs and arena of MakeX Challenge are supplied from JinYuan International Corporation.

References

1. MAKEX Web site, https://www.makex.io/en/about_us.
2. MakeX Robotics Competition Committee, *Competition Tutorial-Reference manual for teachers: MakeX Challenge Courageous Traveler*, Nanjing Baymax Education Technology Co., Ltd., 2019, pp. 74–76.

A Promoting Method of Role Differentiation Using a Learning Rate that Has a Periodically Negative Value in Multi-agent Reinforcement Learning

Masato Nagayoshi, Simon Elderton
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

There have been many studies on multi-agent reinforcement learning (MARL) in which each autonomous agent obtains its own control rule by RL. Here, we hypothesize that different agents having individuality is more effective than uniform agents in terms of role differentiation in MARL. In this paper, we propose a promoting method of role differentiation using a wave-form changing parameter in MARL. Then we confirm the effectiveness of role differentiation by the learning rate that has a periodically negative value through computational experiments.

Keywords: reinforcement learning, multi-agent, negative learning rate, role differentiation

1. Introduction

Engineers and researchers are paying more attention to reinforcement learning (RL) [1] as a key technique for realizing computational intelligence such as adaptive and autonomous decentralized systems. Recently, there have been many studies on multi-agent reinforcement learning (MARL) in which each autonomous agent obtains its own control rule by RL. Then, we hypothesize that different agents having individuality is more effective than uniform agents in terms of role differentiation in MARL. Here, we define “individuality” in this paper as being able to be externally observed, but not a difference that we are incapable observing, such as a difference of internal construction.

We consider that differences in interpretations of experiences in the early stages of learning have a great effect on the creation of individuality of autonomous agents. In order to produce differences in interpretations of the agents’ experiences, we utilized Beck’s “Cognitive distortions” [2], which is a cognitive therapy.

In this paper, we propose a “fluctuation parameter” which is a wave-form changing meta-parameter in order to realize “Disqualifying the positive”* which is one of the “Cognitive distortions”, and a promoting method of role differentiation using the fluctuation parameter in MARL.

We then confirm the effectiveness of role differentiation by introducing the fluctuation parameter into the learning rate, especially having a periodically negative value, through computational experiments using "Pursuit Game" as one of the multi-agent tasks.

2. Q-learning

In this section, we introduce Q-learning (QL) [3] 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:

$$Q\big(s(t-1), a(t-1)\big) \leftarrow Q\big(s(t-1), a(t-1)\big) + \alpha_{\mathrm{Q}}\{r(t-1) + \gamma \max_{b \in \boldsymbol{A}_{\mathrm{Q}}} Q(s(t), b) - Q(s(t-1), a(t-1))\} \quad (2)$$

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

$$\pi(a|s) = \frac{\exp(Q(s,a)/\tau)}{\sum_{b \in A_{\mathrm{Q}}} \exp(Q(s,b)/\tau)} \quad (3)$$

where τ is a positive parameter labeled temperature.

3. Fluctuation Parameter

RL has meta-parameters κ to determine how RL agents learn control rules. The meta-parameters κ include the learning rate α, the discount factor β, ε of ε-greedy which is one of the action selection methods, and the temperature τ of Boltzmann action selection method.

In this paper, the following fluctuation parameter using damped vibration function is introduced into this κ.

$$\kappa(t_{\mathrm{p}}) = \begin{cases} \kappa + A\cos(2\pi(t_{\mathrm{p}}/\lambda) + \phi) & (t_{\mathrm{pa}} < t_{\mathrm{ps}}) \\ \kappa + A\cos(2\pi(t_{\mathrm{p}}/\lambda) + \phi) \times t_{\mathrm{ps}}/t_{\mathrm{pa}} & (otherwise) \end{cases} \quad (4)$$

where $A, t_{\mathrm{p}}, t_{\mathrm{pa}}, t_{\mathrm{ps}}, \lambda$ and ϕ is the amplitude, the phase, the damped phase, the initial phase of damping, the wavelength, and the initial phase parameter of the fluctuation, respectively. The phase t_{p}, the damped phase t_{pa}, the initial phase of damping t_{ps}, and the wavelength λ are needed to set proper units.

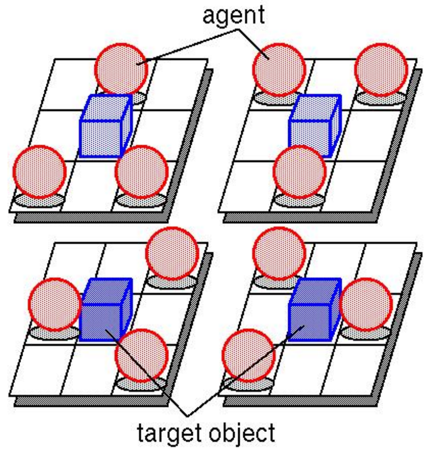


Fig. 1. Capture positions.

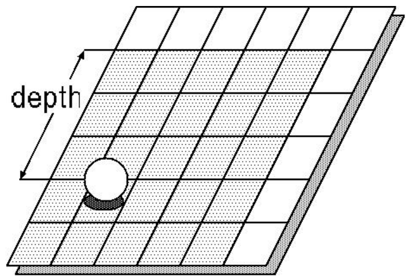


Fig. 2. Range of the agent view (indicated by grayed area).

4. Computational Examples

4.1. *Pursuit Game*

The effectiveness of the proposed approach is investigated in this section. It is applied to the so-called "Pursuit Game" where three RL agents move to capture a randomly moving target object in a discrete 10 ×10 globular grid space. Two or more agents or an agent and the target object cannot be located at the same cell. At each step, all agents simultaneously take one of the 5 possible actions: moving north, south, east, west or standing still. A target object is captured when all agents are located in cells adjacent to the target object and surrounding the target object in three directions as shown in Fig. 1.

The agent has a field of view, and the depth of view set at 3 as shown in Fig. 2. Therefore, the agent can observe the surrounding $(3 \times 2 + 1)^2 - 1$ cells. The agent determines the state by information within the field of view.

The positive reinforcement signal $r_t = 10$ (reward) is given to all agents only when the target object is captured, and the positive reinforcement signal $r_t = 1$ (sub reward) is given to the agent only when the agent is located in the cell adjacent to the target object and the reinforcement signal $r_t = 0$ at any other steps. The period from when all agents and the target object are randomly located at the start point to when the target object is captured and all agents are given a reward, or when 100, 000 steps have passed is labeled 1 episode. The period is then repeated.

4.2. *RL Agents*

All agents observe the only target object in order to confirm the effectiveness of role differentiation, e.g. moving east of the target object. Therefore, the state space is constructed with a 1 dimensional space.

Computational experiments have been done with parameters as shown in Table 1. In addition, all initial Q-values are set at 5.0 as the optimistic initial values.

Table 1. Parameters for Q-learning

Parameter	Value
α_{Q}	0.1
γ	0.9
τ	0.1

4.3. *Example (A): Same Amplitude*

The effectiveness of role differentiation by introducing 4 the fluctuation parameters, in which the initial phase $\phi = 0$, $\lambda = 500$ [step], and the amplitude $A = \{0.1, 0.12, 0.15, 0.17\}$, into the learning rate of QL (hereafter called "0.1", "0.12", "0.15", and "0.17",respectively) are investigated in comparison with an ordinary QL without fluctuation parameter (hereafter called "constant"). Here, the fluctuation parameters of all agents take the same value. The unit of the phase $t_{\mathrm{p}}$is set [step] which is the same as the wavelength λ, the unit of the damped phase $t_{\mathrm{pa}}$is set [episode], and the initial phase of damping is set at $t_{\mathrm{ps}} = 1000$ [episode]. The range of values which the fluctuation parameter for $\alpha_{\mathrm{Q}} = 0.9$ can take e.g. [0.0, 0.2], [-0.05, 0.25] on the condition of $A = 0.1, A = 0.15$, respectively. If the unit of the phase $t_{\mathrm{p}}$is set [episode] and the learning rate is

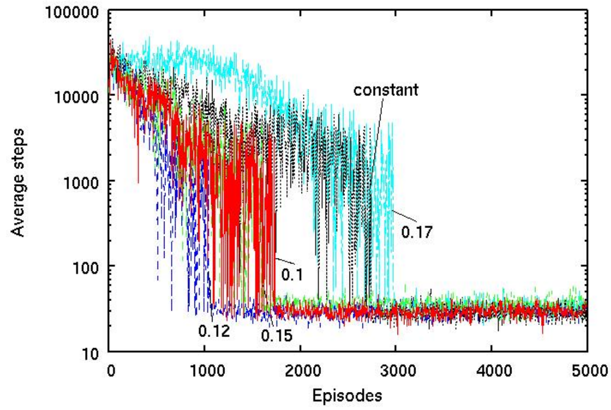


Fig. 3. Required steps of various amplitude parameters in the learning rate.

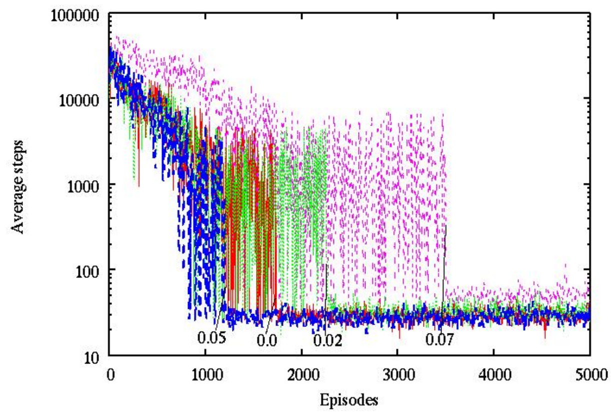


Fig. 4. Required steps of various shifting size of amplitude in the learning rate.

negative, then the control rule which the agent obtains is random. This is the result of negative learning at each step. Therefore, the unit of the phase $t_{\mathrm{p}}$is set [step].

The average numbers of steps required to capture the target object were observed during learning over 20 simulations with various amplitude parameters in the learning rate, as described in Fig. 3.

It can be seen from Fig. 3 that, (1) "0.12" shows a better performance than any other methods and with regard to promoting role differentiation. (2) "0.17" shows a worse performance than "constant" with regard to promoting role differentiation.

Thus, the learning rate that has a periodically a negative value only slightly shows a better performance than the learning rate that has any non-negative value. It could be considered that this is the result of preventing over-fitting by having periodically a negative value.

4.4. *Example (B): Various Amplitude*

In this section, the effectiveness of role differentiation by introducing 4 fluctuation parameters, in which the initial phase $\phi = 0, \lambda = 500$ [step], and the shifting size of the amplitude: $\pm 0.0, \pm 0.02, \pm 0.05, \pm 0.07$ around $A = 0.1$, into the learning rate of QL (hereafter called "0.0", "0.02", "0.05", and "0.07", respectively) are investigated. For example, the amplitudes of 3 agents are $0.05, 0.1, 0.15$ on the condition of the shifting size of the amplitude: ± 0.05 around $A = 0.1$. The same as Example (A), the unit of the phase $t_{\mathrm{p}}$is set [step] which is same as the wavelength λ, the unit of the damped phase $t_{\mathrm{pa}}$is set [episode], and the initial phase of damping is set at $t_{\mathrm{ps}} = 1000$ [episode].

The average numbers of steps required to capture the target object were observed during learning over 20 simulations with various shifting size of amplitude in the learning rate, as described in Fig. 4.

It can be seen from Fig. 4 that, (1) "0.05" shows a better performance than any other methods with regard to promoting role differentiation. (2) "0.02" shows a worse performance than "0.1". (3) "0.07" shows a worse performance than any other method.

Thus, the moderate shifting size of amplitude in the learning rate among the agents shows a better performance than the in the case of same amplitude in the learning rate among the agent.

5. Conclusion

In this paper, we proposed a "fluctuation parameter" which is a wave-form changing meta-parameter in order to realize "Disqualifying the positive" which is one of the "Cognitive distortions", and a promoting method of role differentiation using the fluctuation parameter in MARL.

Through computational experiment using the "Pursuit Game", we confirmed the effectiveness of role differentiation by introducing the fluctuation parameter into the learning rate, especially having a periodically negative value.

Our future projects include evaluating the effectiveness of promoting role differentiation using our proposed fluctuation parameter in order to realize "Jumping to conclusions"†, "Making "must" or "should" statements"‡, and "Overgeneralizing"§ with a state space filter [4].

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number JP19K04906.

References

1. R. S. Sutton and A. G. Barto, Reinforcement Learning (A Bradford Book, MIT Press, Cambridge, 1998).
2. A. T. Beck, Cognitive Therapy and the Emotional Disorders (International University Press, New York, 1976).
3. C. J. C. H. Watkins and P. Dayan, Technical note: Q-Learning, *Machine Learning* **8** (1992), 279-292.
4. M. Nagayoshi, H. Murao and H. Tamaki, A State Space Filter for Reinforcement Learning, in *Proc. of AROB 11th'06* (2006), 615-618.

† Reaching preliminary conclusions (usually negative) with little (if any) evidence.

‡ It can be seen as demanding particular achievements or behaviours regardless of the realistic circumstances of the situation.

§ Making hasty generalizations from insufficient evidence.

Verification of a Combination of Gestures Accurately Recognized by Myo Using Learning Curves

Kengo Kitakura
Graduate School of Education, Hiroshima University, Address
Kagamiyama 1-1-1, Higashi-hiroshima, Hiroshima 739-8524 Japan

Hideyuki Tanaka
Graduate School of Education, Hiroshima University, Address
Kagamiyama 1-1-1, Higashi-hiroshima, Hiroshima 739-8524 Japan

E-mail: tanakalpha@hiroshima-u.ac.jp[*]

Abstract

This paper studies verification of a combination of hand gestures recognized by using the Myo armband as an input device. To this end, relationship between data distribution and learning curves is investigated for binary classification and multi-class classification problems. A verification method is then proposed for finding a combination of gestures accurately classified. Experiments show effectiveness of the proposed method.

Keywords: Learning Curve, Data Distribution, Myo Armband, American Sign, Reliable Gesture Recognition

1. Introduction

A technique for hand gestures recognition from sEMG (surface ElectroMyoGraphy) is useful for extending means of human communication. To find the rule of sEMG that corresponds to the state of hand gestures, machine learning is often used (e.g. Refs. 1–2). Recently, the Myo armband[3] is one of the most popular sEMG acquisition systems, because it is relatively inexpensive and easy to remove. The Myo armband is being regarded as one of the tools for operating VR[4] and for communicating with hearing impairment,[1] and so on.

In developing various applications of Myo as an input device, it is important to verify a combination of gestures learned from the Myo data, because we cannot develop a reliable application without finding a combination of gestures accurately classified. To find such a combination, we need to confirm that the number of the data used for learning is sufficient, but we cannot confirm from a single matrix of classification accuracy.

There is a method for confirming that the number of the data used for learning is sufficient; that is using a learning curve.[5] The meaning of the learning curve in the context of machine learning is mainly divided into two.[6] One is a graph that is created by plotting performance measure against the training iteration on the condition that the number of training data is fixed. The other is a graph that is created by plotting performance measure against the number of the data used for training. In this study, we mean the learning curve by the second graph; it is created by plotting classification accuracy of training or test against the number of data for training.

In previous studies, the learning curve has been used for predicting the data required for DNA classification[7] and for comparing the methods of machine learning algorithms,[8] and so on. These studies have used the averaged learning curve, which is an average of the discrimination accuracy of all classes. In contrast, Ref. 9 proposed to use the learning curve for individual classes.

Learning curves are useful for verifying a combination of gestures accurately classified, because they possibly offer information on the boundary and data distributions. Learning curves indeed indicate accuracy of a classifier, and accuracy depends on the boundary and data distributions. However, the relationship between learning curves and data distributions has not been clear. If the relationship is clarified, it would be useful for inferring them in a high dimensional space.

There are a lot of researches for gesture recognition using the Myo armband.[1,2,10] But the researches have not been done for finding a combination of gestures accurately classified in the light of learning curves.

In this paper, we study relations of learning curves and data distribution. We then verify a combination of gestures accurately classified using a Myo armband.

2. Myo Armband and Gestures

Myo (Fig. 1) is an armband type gesture controller developed by Thalmic Labs. It measures sEMG of the arm (Fig. 2). Myo has sEMG sensors of 8 channels and its sampling frequency is 200Hz. The sEMG measured by each channel is converted analog to digital, and sent to the PC as an integer value between -128 and 128 by Bluetooth. The signal indicates the amount active muscle fibers and is a dimensionless value.

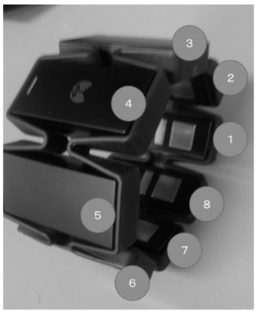


Fig. 1. Myo armband

Fig. 2. Wearing Myo armband

Let us represent the sEMG by $v_l(t)$, where l is the number of channel ($l = 1, \ldots, 8$) and t is the time index ($t = 1, \ldots, N_v$). We use the average of the absolute value $w_l(t)$ for each channel in 1 second as features for classification and define as follows:

$$w_l(j) = \frac{1}{200}\sum_{t=1}^{200} |v_l(200(j-1)+t)|, \tag{1}$$

where $j = 1, \ldots, N_w$ and $l = 1, \ldots, 8$. It should be noted that the features (1) were used by Ref.11 as MAV.

American sign language (ASL) is a sign language mainly used in North America. It is published[13] by NIDCD and shown in Fig. 3. We deal with the sEMG data of 24 gestures except "j" and "z" in ASL, because we classify the hand gestures by using only the sEMG data and the gestures of "j" and "z" include motion of fingers. To conduct experiments, we make training and test dataset. The size of each dataset is 24×100. We use *K*-nearest neighbor of the scikit-learn toolkit for classification.

Fig. 3. Amerincan Sign Language. [12]

3. Learning Curve and Data Distribution

In this section, given the training and test data sets, we study characteristics of learning curves. We first consider a binary classification problem given the uniformly distributed data, and then a multi-class one given the normally distributed data. Suppose that the training data $x_1, \ldots, x_N \in \mathbb{R}^d$ are given.

We construct the learning curve by increasing the training data one by one chosen from the given data. We compute the learning curve by taking the average, because the learning curve depends on how to increase the training data. We will show how to increase the data in Appendix A.

3.1. *Binary classification problem*

Suppose that the dimension f of the feature space is 1 and consider a binary classification problem ($c = 2$). Assume that the training data belonging to A and B are given and are uniformly distributed as follows:

$$0.1 \le \xi_{\mathrm{A}}(k) \le 0.5, \quad 0.6 \le \xi_{\mathrm{B}}(k) \le 1.0,$$

where $\xi_{\mathrm{A}}(k)$ and $\xi_{\mathrm{B}}(k) \in \mathbb{R}$ are respectively the k th training data ($k = 1, ..., N$). We hence have training data**:**

$$\boldsymbol{x_k} = [\xi_{\mathrm{A}}(k), \xi_{\mathrm{B}}(k)]^{\boldsymbol{T}} \in \mathbb{R}^d \quad (k = 1, \dots, N),$$

where $d = fc = 2$. We moreover suppose that the test data corresponding to A and B are given, and that they are respectively labeled as A' and B'. The test data are uniformly distributed as follows:

$$0.3 \le \xi_{\mathrm{A'}}(k) \le 0.7, \quad 0.4 \le \xi_{\mathrm{B'}}(k) \le 0.8,$$

where $\xi_{\mathrm{A'}}(k)$ and $\xi_{\mathrm{B'}}(k) \in \mathbb{R}$ are respectively the k th test data ($k = 1, ..., M$).

We show the distribution of training and test data in Fig. 4. The classification boundary is 0.55, if the number of the training data is sufficiently large, because the boundary is between 0.5 and 0.6 in the training data set and the boundary 0.55 maximizes the margin of classification. If the number of training and test data, N and M, are sufficiently large, the percentage of the blue area in Fig. 4 represents the misclassification rate:

$$e = \frac{(0.7-0.55)+(0.55-0.4)}{(0.7-0.3)+(0.8-0.4)} = 0.375.$$

The misclassification rate e is obtained by the volume of the misclassification areas of A' and B' (blue areas), supposing the data of A' and B' are uniformly distributed. The accuracy rate is moreover given by $1 - e$ for large M and N.

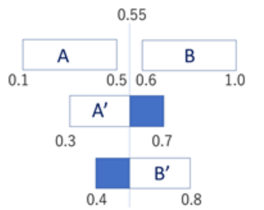


Fig. 4. The distribution of training and test data. The first row represents the training data, and the second and third row do the test data. The data in the blue area is misclassified by the classification boundary 0.55.

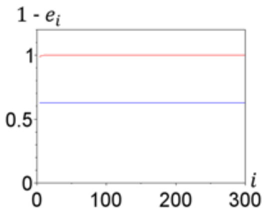


Fig. 5. Learning curves. Blue and red curves respectively represent the learning curves for the training and test data. The value of learning curve of the test data of i 300 is approximately 62.5%.

The learning curve is created by plotting the classification accuracy $1 - e_i$ against the number of training data i, where e_i is misclassification rate for i. The accuracy of the learning curve $1 - e_N$ approaches $1 - e$, as the number of the training data N becomes large. We therefore see that the accuracy rate $1 - e$ of the classification for the test data (Fig.4) is related with that for sufficiently large data in the learning curve (Fig. 5). This fact implies that the accuracy rate of K-nearest neighbor is $1 - 0.375$ (62.5%), if N and M are sufficiently large.

Let us consider the case where N and M are finite. We suppose that the classifier sets a classification boundary at z, given N training data. We express the total numbers of k satisfying the following inequalities respectively by $M_{\mathrm{A'}}$ and $\widetilde{M}_{\mathrm{A'}}$:

$$0.3 \le \xi_{\mathrm{A'}}(k) \le z, \quad z \le \xi_{\mathrm{A'}}(k) \le 0.7,$$

meaning that the number of data correctly classified is $M_{\mathrm{A'}}$ and that the one misclassified is $\widetilde{M}_{\mathrm{A'}}$; i.e. the following equation is established:

$$M = M_{\mathrm{A'}} + \widetilde{M}_{\mathrm{A'}}.$$

In the same way, we express the total numbers of k satisfying the following inequalities respectively by $M_{\mathrm{B'}}$ and $\widetilde{M}_{B'}$:

$$z \le \xi_{\mathrm{B'}}(k) \le 0.8, \quad 0.4 \le \xi_{\mathrm{B'}}(k) \le z,$$

meaning that the number of data correctly classified is $M_{\mathrm{B'}}$ and that the one misclassified is $\widetilde{M}_{\mathrm{B'}}$; i.e. the following equation is established:

$$M = M_{\mathrm{B'}} + \widetilde{M}_{B'}.$$

The accuracy rate $1 - e_N$ is then given by

$$1 - e_N = \frac{M_{\mathrm{A'}} + M_{\mathrm{B'}}}{2M}. \qquad (2)$$

The left hand side of (2) is related to the learning curve at N, and the right hand side depends on the test data and the boundary set by the N training data. In other words, the right hand side represents the percentage of data that does not violate the boundary. It should be noted that the test data are uniformly distributed, but the right hand side of (2) can be calculated regardless of the distribution, by just counting the number of not violating the boundary.

3.2. *Multi-class classification problem*

We study a multi-class classification problem. Let us consider data in feature space of $f = 2$ dimensions and classify them into $c = 3$ classes. Suppose that normally distributed data $\boldsymbol{w}_{\mathrm{A}}(n), \boldsymbol{w}_{\mathrm{B}}(n)$, and $\boldsymbol{w}_{\mathrm{C}}(n) \in \mathbb{R}^f$ ($n = 1, \dots, N$) are given as training data and labeled as A, B, and C, respectively. We then construct vector variables $\boldsymbol{x}_n \in \mathbb{R}^d$ for ($n = 1, \dots, N$):

$$\boldsymbol{x_n} = [\boldsymbol{w}_{\mathrm{A}}(n)^T, \boldsymbol{w}_{\mathrm{B}}(n)^T, \boldsymbol{w}_{\mathrm{C}}(n)^T]^T.$$

Assume that normally distributed data $\boldsymbol{w}_{\mathrm{A'}}(n), \boldsymbol{w}_{\mathrm{B'}}(n)$, and $\boldsymbol{w}_{\mathrm{C'}}(n) \in \mathbb{R}^f$ are given and corresponding to the labels of A, B, and C, respectively. Figs. 6 and 7 show the training data A, B, and C, and test data A', B', and C'.

The green lines in Figs. 6 and 7 are the classification boundaries trained by all the data A, B, and C. The results of drawing individual learning curves for the training and test data are shown in Figs. 8-10. Averaged learning curves for test and training data are shown in Fig. 11.

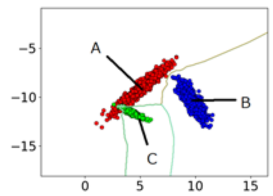


Fig. 6. Training data indicating three classes of A, B and C. The green line is the boundary of the classes.

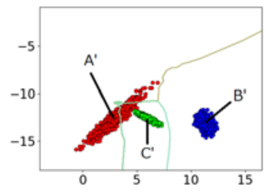


Fig. 7. Test data indicating three classes of A', B' and C'. The green line is the boundary of the classes.

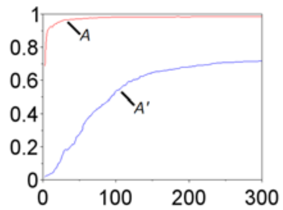


Fig. 8. Learning curves (A, A'). Red line is the learning curve of training, and blue line is that of test.

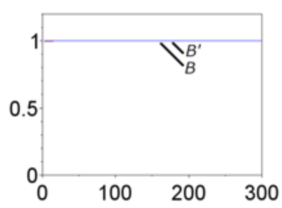


Fig. 9. Learning curves (B, B'). Red line is the learning curve of training, and blue line is that of test.

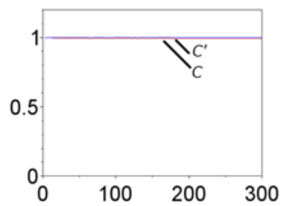


Fig. 10. Learning curves (C, C'). Red line is the learning curve of training, and blue line is that of test.

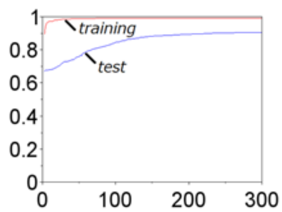


Fig. 11. Learning curves (averaged). Red line is the learning curve of training, and blue line is that of test.

From the individual learning curves of training and test data in Figs. 8, 9, and 10, we examine the relationship between the classification boundary and the distribution of the training and test data in Figs. 6 and 7.

In the same way as deriving (2), we see that the accuracy of the test data B' is equal to the ratio of the data B' that does not violate the boundary determined by the training data A, B and C. The label B' is indeed classified with 100% accuracy as shown in Fig. 9, indicating that all test data of B' are not outside of the trained area of B (Fig. 7). The label B is also classified with 100% accuracy (Fig. 9), showing that all training data of B can be trained correctly (Fig. 6). We observe that the same things for the label C' and C. On the other hand, we see from the learning curve in Fig. 8 that accuracy of classification of A' at 300 sets of data is about 70%. This fact implies that the test data of A' cross the classification boundary and that some of them are outside of the classification area of A as shown in Fig. 7.

We can obtain the ratio of misclassification caused by changes of distribution between training and test, from the gap between the learning curves of training and test data. For example, there is a gap between A and A' in Fig. 8. In this case, the ratio of test data correctly classified is lower than that of training data. This is because the test data A' violates the boundary defined by the training data A in Fig. 7. On the other hand, the data B' correctly classified by the boundary set by the training data. In this way, we can find test data seriously affected by distribution changes, by constructing individual learning curves.

4. Experiments

We investigate characteristics of the data of Myo in section 4.1 and propose a method for finding a combination of gestures accurately classified in section 4.2.

Let us consider data in a feature space of $f = 8$ dimensions and classify them into $c = 24$ classes; they corresponds to characters of the alphabet except "j" and "z". Using $w_l(j)$ in (1), we define the data of Myo in the feature space as: $\boldsymbol{w}(j) = [w_1(j), w_2(j), \dots, w_8(j)]^T \in \mathbb{R}^f$.

Suppose that the training data for the label "a" to "y" except "j" are respectively given by $\boldsymbol{w}_{\mathrm{a}}(n)$ to $\boldsymbol{w}_{\mathrm{y}}(n)$ except $\boldsymbol{w}_{\mathrm{j}}(n)$ $(n = 1, \dots, N)$. We moreover construct vector variables:

$$\boldsymbol{x_n} = [\boldsymbol{w}_{\mathrm{a}}(n)^T, \boldsymbol{w}_{\mathrm{b}}(n)^T, \dots, \boldsymbol{w}_{\mathrm{y}}(n)^T]^T \in \mathbb{R}^d,$$

where $d = fc$.

4.1. *Visualization of Myo data*

We visualize the features extracted from the data of Myo to see the characteristics of the data distribution. Figs. 12 and 13 indicate the results of principal component analysis showing 8-dimension feature values obtained from Myo by reducing the dimensions to 2. Each of the three classes corresponds to a, b, and f gestures. Even if Myo is not removed between training data acquisition and test data acquisition, the data distribution changes between training and test data acquisition, and it makes learning difficult. This is true for the case where Myo is removed as well. We will therefore study verification of a combination of gestures using learning curves based on the investigation of the boundary and the test data in Section 3.

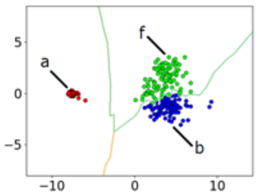


Fig. 12. Training data of Myo.

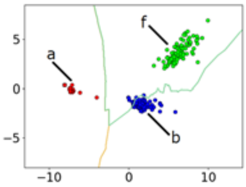


Fig. 13. Test data of Myo.

4.2. *A method for verification*

We verify a combination of gestures that can be classified with high accuracy among ASL. We first acquire 24 training and test data labeled "a" to "y" except "j". We then make verification, using individual learning curves and averaged ones and conducting experiments for acquiring data.

Suppose that the numbers of training and test data are the same ($N = M$). Let us draw individual learning curves of training and test data. If accuracy is low for a label (e.g. "k") in both training and test data at N and if the gap between them are very small, then there is a possibility that accuracy of the classifier for the label (e.g. "k") may be enhanced by excluding another label that makes conflict for classification. We should therefore keep such a label (e.g. "k"), if the gap between learning curves of training and test data is small. On the other hand, if there is a large gap between the training and test data in learning curves of a label (e.g. "f"), then the label (e.g. "f") should be excluded from the classification target.

We show an example of learning curves for a gesture to be excluded in Fig. 14. Even if the learning curve of training is seen that classification is possible, the gesture should be excluded from the classification target in case that the learning curve of test data at N indicates low accuracy because of the gap between the training and test data. We show another example of the gesture regarded as a classification target in Fig. 15. Accuracy for training and test data are both high, and there is almost no gap between them.

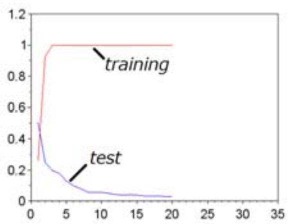


Fig. 14. Example of learning curves for a gesture ("f") to be excluded.

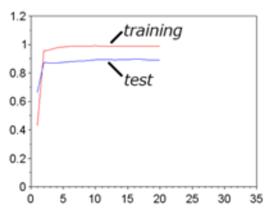


Fig. 15. Example of learning curves for a gesture ("k") not to be excluded.

Let us reduce the number of the combination of gestures from c_s to c_e for finding a reliable one. Based on the above consideration, the combination of gestures is verified by the following verification algorithm.

[Verification algorithm]:

Step 1: Draw individual learning curves and an averaged one for the number of classes c_s.

Step 2: Exclude the gesture that has a large gap between learning curves of the training and test data. If there are no more gestures to be excluded, then go to Step5.

Step 3: Re-acquire data for gestures that are not excluded in Step 2 and draw the individual and averaged learning curves.

Step 4: Repeat Step 2 and Step 3 until there are no more gestures that can be excluded in Step 2.

Step 5: Re-acquire data for not excluded gestures and draw the individual and averaged learning curves. Check if an averaged learning curve has a satisfactory discrimination accuracy. If it is unsatisfied, go to Step2.

In Step2, excluded gestures are determined by referring the accuracy indicated by the learning curve as shown in Fig14 and Fig15. Suppose that the number of the training data is α. In this experiment, α is 20. The value of α depends on the user who allows how much time for learning. The more α is increased, the more time is needed for learning. Of course, the user can reduce the value of α by seeing the learning curve.

We make experiments and apply the proposed method to the data. Since data distribution depends on removal of Myo between experiments, we investigate verification by taking the interval time between test and training data acquisition into account. We thus consider two cases: For the first case we do not remove Myo between data acquisition (**Case 1**), and for the second case we do it. We moreover consider two cases in the second case: There is little time in the interval (**Case 2**), and there are several days (**Case 3**). We apply the verification algorithm and conduct experiments in the order of Cases 1, 2, and 3, since the number of the reliable combination is decreased in the order of them.

4.3. *Results of experiments*

We determined the threshold γ for the gap of accuracy between training and test data as 20%. We obtained a combination of gestures with high accuracy for each Cases 1, 2, and 3, extracting a combination of gestures that was not in a trade-off relationship. As a result, we found the followings. In Case 1, 5 classes of "e", "k", "q", "r", "y" can be classified with 100% accuracy. Also, In Case 2, the classifier obtained by the verification algorithm can classify the 4 classes "e", "k", "q", and "y" with 99% accuracy. But in Case 3, it was possible to classify with 99% accuracy, if the gestures were classified into two classes of "k" and "q".

5. Conclusions

In this study, we investigated the relationship between data distribution and learning curves, and we then verified a combination of American sign language that is accurately classified using a Myo armband. In addition, as a result of investigating the characteristics of the data acquired from Myo, it was found that the distribution of data changes between the interval of training and test data acquisition, and verifications were hence carried out for the cases of different intervals. It remains a future topic how to determine the threshold γ for the gap to ensure accuracy.

Appendix A. Appendices

Suppose that the data $\boldsymbol{x}_1, \dots, \boldsymbol{x}_N \in \mathbb{R}^d$ are given. We pick up the data from them and increase the training data from 1 to $n (n \leq N)$. There are many combinations in increasing training data via picking up them one by one. We construct the learning curve of a classifier by averaging them for different combinations, because the learning curve depends on how to increase the data.

We explain how to construct the learning curve. Let us consider a permutation map $s_N: \{1, \dots, N\} \to \{1, \dots, N\}$. Since there exist $N!$ maps for s_N, we describe them by $s_{N,k}$ ($k = 1, 2, \dots, N!$),where $s_{N,i} \neq s_{N,j}$ $(i \neq j)$. For $s_{N,k}$, define $\widehat{\boldsymbol{x}}_{1,k}, \widehat{\boldsymbol{x}}_{2,k}, \dots, \widehat{\boldsymbol{x}}_{N,k} \in \mathbb{R}^d$ as:

$$\begin{bmatrix} \widehat{\boldsymbol{x}}_{1,k} \\ \vdots \\ \widehat{\boldsymbol{x}}_{N,k} \end{bmatrix} := S_{N,k} \begin{bmatrix} \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_N \end{bmatrix}, \quad S_{N,k} := s_{N,k} \otimes I_d,$$

where I_d is the Identity matrix of $d \times d$, and $\otimes$ is the symbol for the Kronecker product.[13] Since $\widehat{\boldsymbol{x}}_{1,k}, \dots, \widehat{\boldsymbol{x}}_{N,k}$ are given by changing the order of $\boldsymbol{x}_1, \dots, \boldsymbol{x}_N$, the following equation holds for $k = 1, \dots, N!$

$$\{\widehat{\boldsymbol{x}}_{1,k}, \dots, \widehat{\boldsymbol{x}}_{N,k}\} = \{\boldsymbol{x}_1, \dots, \boldsymbol{x}_N\},$$

meaning that the set of $\boldsymbol{x}_1, \dots, \boldsymbol{x}_N$ and that of $\widehat{\boldsymbol{x}}_{1,k}, \widehat{\boldsymbol{x}}_{2,k}, \dots, \widehat{\boldsymbol{x}}_{N,k}$ are same. Define a set $D_{i,k} = \{\widehat{\boldsymbol{x}}_{1,k}, \dots, \widehat{\boldsymbol{x}}_{i,k}\}$ and let $\eta_{i,k}$ be the classification accuracy of the classifier learning from the training data $D_{i,k} (i = i_0, \dots, N)$, where i_0 is the minimum number of data required for the classifier. Let us randomly choose m maps from the maps $s_{N,k}$ $(k = 1, 2, \dots, N!)$ and take an average for accuracy $\eta_{i,k}$ for the training data $D_{i,k}$:

$$\hat{\eta}_i = \frac{1}{m} \sum_{k=1}^{m} \eta_{i,k}.$$

The index i and $\hat{\eta}_i$ construct a learning curve for the training data. Given the test data, let $\theta_{i,k}$ be accuracy of the classifier for the training data $D_{i,k}$ and take an average for $\theta_{i,k}$ for the training data $D_{i,k}$:

$$\hat{\theta}_i = \frac{1}{m} \sum_{k=1}^{m} \theta_{i,k},$$

where i and $\hat{\theta}_i$ construct a learning curve for the test data.

References

1. C. Savur, F. Sahin, American sign language recognition system by using surface EMG signal, *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2016, pp. 002872-002877.
2. L. C. Galea and A. F. Smeaton, Recognising Irish sign language using electromyography, *IEEE Content Based Multimedia Indexing (CBMI) Conference*, 4-6 Sept 2019.
3. U. C.-Allard, G. G. Turcotte, F. Laviolette, and B. Gosselin, A Low-Cost, Wireless, 3-D-Printed Custom Armband for sEMG Hand Gesture Recognition, *Sensors*, **19**(12), 2019.
4. Y. S. Pai, T. Dingler, and K. Kunze, Assessing hands-free interactions for VR using eye gaze and electromyography, *Virtual Reality*, **23**(2), 2019, pp. 119-131.
5. P. Langley, Crafting papers on machine learning, *ICML*, 2000, pp. 1207-1216.
6. C. Perlich, Learning Curves in Machine Learning, *IBM Research Report*, RC24756(W0903-020) March 5, 2009
7. S. Mukherjfe *et al*., Estimating dataset size requirements for classifying DNA microarray data, *Journal of computational biology*, **10**(2), 2003, pp.119-142.
8. C. Perlich, F. Provost, and J. S. Simonoff, Tree induction vs. logistic regression: A learning-curve analysis, *Journal of Machine Learning Research*, **4**(Jun), 2003, pp.211-255.
9. Y. Wahba, E. ElSalamouny, and G. ElTaweel, Estimating the sample size for training intrusion detection systems, *International Journal of Computer Network and Information Security*, **10**(12), 2017, pp. 1-10.
10. J. G. Abreu, J. M. Teixeira, L. S. Figueiredo, and V. Teichrieb, Evaluating sign language recognition using the myo armband, *IEEE 2016 XVIII Symposium on Virtual and Augmented Reality (SVR)*, 2016, pp. 64-70.
11. I. M. Donovan, J. Puchin, K. Okada, and X. Zhang, Simple space-domain features for low-resolution sEMG pattern recognition, *IEEE 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2017, pp. 62-65.
12. National Institute on Deafness and Other Communication Disorders(NIDCD), American Sign Language, https:// www.nidcd.nih.gov/health/american-sign-language

13. J. W. Brewer, Kronecker products and matrix calculus in system theory, *IEEE Transactions on circuits and systems*, **25**(9), 1978, pp.772-781.

Augmentative and Alternative Communication Device Based on Head Movement to Aid Paralyzed Victims with Speech Disabilities

Vihanga Ashinsana Wijayasekara
Informatics Institute of Technology, No 56, Ramakrishna Road, Colombo 6, Sri Lanka

Torin Wirasingha
Informatics Institute of Technology, No 56, Ramakrishna Road, Colombo 6, Sri Lanka
E-mail: vihanga.2017581@iit.ac.lk, torin.w@iit.ac.lk
www.iit.ac.lk

Abstract

Augmentative and Alternative Communication (AAC) devices have been proven to be as an alternative to help people, who are having communication difficulties. In this paper, high tech AAC device focusing on paralyzed victims or other motor disabilities which allow them to control with their head gestures is proposed. The proposed system consists of a mobile app, a controller and a head tracker. Easy to familiar with and less material cost are some of the key advantages of this device.

Keywords: Augmentative and Alternative Communication, AAC Device, Head Gestures, Speech Disabled

1. Introduction

Technological breakthrough improves day-to-day life, and it gives new opportunities to people with special needs. Augmentative and Alternative Communication (AAC) device is the best example which enhances the quality of life of people who are unable to use verbal speech to communicate. Communication methods used in AAC devices vary from paper and pencil boards to digital speech-generating devices. Eye gaze technology which enables the use of eye movements to operate a device and head pointing/tracking technology which is used to operate a device and does a major role in high tech AAC devices. However, such eye gaze and head pointing communication devices which are available today are not convenient enough to afford by the majority. This paper suggests an AAC device, operated using patient's head gestures and further prioritizing its cost factor for increased affordability.

2. Literature Review

2.1 AAC devices related work

There are a significant amount of researches have been done on AAC devices considering different interaction techniques between patient and the device. Those include mainly, designs based on eye blink detection [1], lip contour detection [2], Brain-Computer Interface (BCI) and EEG based approaches [3] and research based on breathing signal patterns as the main communication medium between the user and the device [4]. The Eye blink detection method uses an IR sensor to detect blinks. However, directing IR to an opened eye for long run may lead to cause health issues [5] and further such systems have been identified as less reliable on different environmental conditions where the IR in outdoor [6] and image recognition for identifying lip movements in dark environments. Some mobility-related issues have been identified under the implementation of BCI and EEG signal based AAC devices as it is more often composed of a separate computer apart from simple Single Board Computers (SBC), in order to carry out complex computations. Since the breathing pattern changes, based on different factors such as emotions, body temperature, diseases suffering from [7], there is a possibility to have unpredictable outputs by adopting breathing signal-based approaches for patient-device interaction.

In order to use most of the above products, the user has to convert his/her message to Morse code using the interaction technique. Then the device captures the Morse code, decrypt and convert it to voice output. So in order to use those devices, knowledge on Morse code is essential.

2.2 Research and commercial products based on head trackers

Many head trackers in the market are allowed the user to move the cursor via head movements. Most of them are not cost effective and such systems more often incorporated with optical sensors or IR cameras to detect head movements. While many systems have extra buttons to get user inputs, most of them have Dwell selection software which in most cases sold separately. Difficulties on the initial adaptation of very young age children and people with cognitive impairment on such systems as well as immediate recognition of the cursor process generally tend to dissatisfactions [8]. Further, it may cause pain with extended use. Several types of research have been carried out on developing systems based on capturing head movements using an accelerometer sensor. The accelerometer sensor which is integrated with such systems detects the head movements of the user reliably and further it is used to control a wheelchair as per the user preferences [9]. Rushikesh T. Bankar has developed a Head Gesture Recognition System using Field-Programmable Gate Array (FPGA) based on a Smart Camera. It is a vision-based gesture recognition system [10].

3. Methodology

The proposed AAC system composed of a Bluetooth enabled controller, head tracker and mobile app.

3.1 Mobile app

The app is an AAC app which can be controlled using head gestures. In order to be controlled via head gestures, the phone needs to connect with the controller through Bluetooth. The app provides 3 screens,

About Screen – This is the screen where Bluetooth settings appear.

Command Screen – Set of core vocabularies, which contain some verbs, pronouns and adjectives. There are 24 words act as touchable buttons and the words are highlight one by one. User has to perform a particular head gesture when the word that user expected to express is highlighted. Then it generates a voice output for the selected items respectively.

Type Screen – This screen allows the user to type sentences/phrases. It contains numbers from 0-9 and the English alphabet, including Space and Delete button as touchable buttons. As same as the Command Screen, the buttons will highlight, one by one and user should perform a particular head gesture when the button that user expected to express is highlighted. Then it will display on the screen.

Initially, the phone's Bluetooth needs to be turned on, then the user has to open the app and click "Connect the Controller" button. After the controller connected successfully with the phone, the app informs the user via a voice message that the phone is connected to the controller successfully. Then the calibration process starts. In the calibration process the application finds the easiest and most suitable two head gestures for the user according to each user's motor ability.

One gesture for selecting the button which is highlighted and another gesture for moving within the screens (About/Command/Type). Then the variation in the angle of inclination in x, y and z in each defined head gesture is analyzed. When a user does a head gesture, the controller compares the head gesture's angles of inclination in x, y and z variation with predefined head gestures' variations and proceed according to matched predefined head gestures.

Fig 1. Selecting loop in Command Screen

3.1.1 Limiting head movements

In order to reduce head movements, the following features were introduced in the app.

Autocomplete in Type screen - When the user starts typing a word, word suggestions are appearing on the top of the screen. Maximum 3 words are set to appear on the screen. Suggested words are stored in the app.

Dynamic Reordering in Command Screen – All the commands in the command screen are dynamically reordered by their usage. (most clicked commands appear on the top left side)

Text Predicting in Type Screen – When the user starts typing, a suggested sentence is displaying on the screen. The sentence is suggested by a machine learning model which is developed using LSTM recurrent neural network. All the sentences typed by the user are stored in a database and they are used to train the machine learning model.

3.2 Bluetooth enabled controller

The controller composed of a microcontroller (Arduino UNO), a Bluetooth module (HC-05) and a 16*2 LCD Display (LCM1602 IIC V1). The LCD display provides instructions to the user about the calibration instructions.

3.3 Head tracker

The head tracker consists of an accelerometer and gyroscope sensor module (MPU6050) which contains both 3-Axis accelerometer and 3-Axis gyroscope and it acquires the inclination of the current position of the head tracker unit based on its reference X, Y and Z axes. In order to get more accurate values, the acquisition process followed by complementary filtering is used.

3.4 Process

User can give 4 types of commands using predefined head gestures.

Table 1. Commands sent by the controller according to user behaviour

User's Expectations	Command send by Controller to the phone
Go to about Screen	1
Go to Command Screen	2
Go to Type Screen	3
Select a Command/Letter	4

When the user needs to select a Command/Button he/she needs to do a particular head movement defined in the calibration process. Then the controller sends "4" to the phone via Bluetooth. After the app read the value "4", it outputs the respective text assigned selected button.

When the user does the head movement for navigating which is defined in the calibration process the controller send "1" to the phone via Bluetooth. When the user again does the above movement, the controller sends "2", "3" respectively.

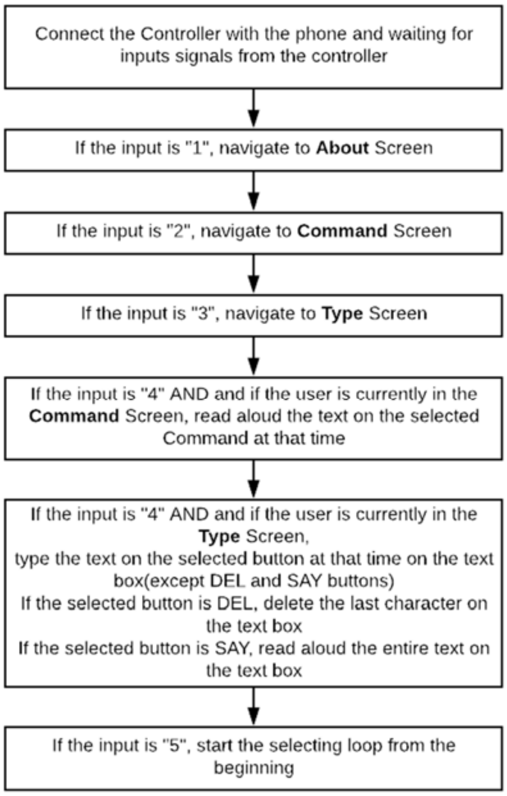


Fig 2. Entire flow in a Flow Diagram

4. Experiment and Results

The objective of the experiment is to determine the accuracy of the proposed system. Hence the test was made for healthy individuals. 8 healthy subjects were volunteered for the test including 6 male subjects and 2 female subjects.

Initially, test subjects were instructed to proceed with some random commands based on their preference and after that they were directed with 15 pre-selected commands using head gestures. Results based such commands mentioned earlier were evaluated based on its accuracy of achieving the exact output.

Table 2. Scores and Success rate for each individual who participated in the experiment

Subject	Score	Success Rate (%)
A	15	100
B	14	93.33
C	13	86.67
D	14	93.33
E	15	100
F	15	100
G	13	86.67
H	14	93.33
Avg. Success Rate		**94.16**

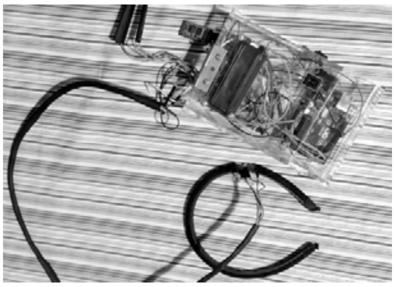


Fig. 3. Image of Controller and Head Tracker

5. Discussion

The proposed system was able to achieve average 94.14% success rate in the testing. Compare to other devices, this system gets a lower time for calibration and understanding the system. It took an average less than 15 minutes for each test user to familiarize with the system inclusive with the initial calibration process. With the high success rate of 94.16% the system can be applied to paralyzed individuals.

6. Conclusion

In this paper, the authors proposed an AAC system which cost less material cost and can be controlled using head gestures. During the testing process it was identified that, compared to many other devices, it takes considerably less time period for the user to become familiarized with the proposed head tracking system. All people selected bowing their heads as one of the gestures. In the experiment, the combination of multiple sensory outputs was suggested in order to get more accurate result. Future research will be focused on integrating other aid devices such as wheelchair controllers with the controller and working on the above-suggested idea.

7. References

[1] K. Mukherjee and D. Chatterjee, "Augmentative and Alternative Communication device based on eye-blink detection and conversion to Morse-code to aid paralyzed individuals," *2015 International Conference on Communication, Information & Computing Technology (ICCICT),* pp. 0-4, 2015.

[2] M. Anila and P. Radhika, "Lip contour detection based AAC device using morse code," *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET),* p. 1182–1187, 2017.

[3] M. Yeo, L. Jiang, E. Tham and W. Xiong, "Evaluation of a low-cost alternative communication device with brain control," *2015 IEEE 10th Conference on Industrial Electronics and Applications (ICIEA),* p. 229–232, 2015.

[4] R. G. Garcia et al., "Wearable augmentative and alternative communication device for paralysis victims using Brute Force Algorithm for pattern recognition," *2017IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM),* 2017.

[5] J. Cato, "Negative Effects of Infrared Waves," 2018. [Online]. Available: https://sciencing.com/negative-effects-infrared-waves-8592303.html.

[6] Y. Elsahar, S. Hu, K. Bouazza-Marouf, D. Kerr and A. Mansor, "Augmentative and Alternative Communication (AAC) Advances: A Review of Configurations for Speech Disabled Individuals," *Sensors,* vol. 19, no. 8, 2019.

[7] "Breathing Pattern," [Online]. Available: https://www.sciencedirect.com/topics/medicine-and-dentistry/breathing-pattern.

[8] S. Fager, L. Bardach, S. Russell and J. Higginbotham, "Access to augmentative and alternative communication: new technologies and clinical decision-making.," *Journal of pediatric rehabilitation medicine,* vol. 5, no. 1, pp. 53-61, 2012.

[9] R. T. Bankar and S. S. Salankar, "Head Gesture Recognition System Using Gesture Cam," *2015 Fifth International Conference on Communication Systems and Network Technologies,* vol. 8, no. 6, pp. 341-346, 2015.

[10] M. Davy and R. Deepa, "Hardware Implementation Based On Head Movement Using Accelerometer Sensor," *Int. Journal of Applied Sciences and Engineering Research,* vol. 3, no. 1, pp. 17-21, 2014.

The research about editing system of performance information for player piano. -Make inferences about whole musical composition by using DP matching system-

Mami Ezaki
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
Hayashi Lab, 680-4, Kawazu, Iizuka-City, Fukuoka, 820-8502, Japan†
E-mail: haya@mmcs.mse.kyutech.ac.jp

Abstract

We have developed a system that allows a piano to perform automatically. In order to play music in the manner of a live pianist, we must add expression to the piano's performance. In the case of music, there are often 1,000 or more notes in the score, requiring that an editor spend a huge amount of time to edit. Therefore, we have developed an interactive musical editing system that utilizes a database to edit music more efficiently.

Keywords: Automatic Piano, Knowledge Database, Computer Music, Music Interface

1. Introduction

We have developed a system that allows a piano to perform automatically. In this system, 90 actuators have been installed on the keys and pedals of a grand piano. These actuators execute key strokes and pedal movements to govern the piano's performance, e.g., "*Fig. 1.* (*Continued*)"

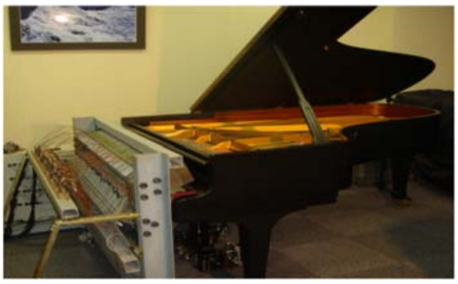


Fig.1 The automatic piano.

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. In this paper, we developed a system that automatically estimates the performance expression of unedited music using edited performance data and score data. We described a phrase search using DP matching and a method for selecting an optimal phrase for inference, and evaluated the entire song.

2. Editing support system

2.1. *Performance information*

The automatic piano that we have developed uses a music data structure that is similar to MIDI. We defined performance information, dividing it into two categories: the notes and the pedals. The note information is comprised of the six parameters involved in producing a tone: “Key” (note), “Velo” (velocity), “Gate”, “Step”, “Bar” and “Time”. “Velo” is the dynamics, given by the value of 1-127. “Gate” is the duration of the note in milliseconds. “Step” is the interval of time between notes, and it also exhibits tempo. “Bar” is the vertical line placed on the staff to divide the music into measures.
The pedal information is comprised of four parameters: “Key” (indicating the kind of pedal: “Damper” or “Shifting”), “Velo” (the pedaling quantity), “Time” (the duration for which the pedal is applied)”, and “Bar”, e.g., “*Table 1.* (*Continued*)”

Table.1 Note Information

Parameter	Key	Velo	Gate	Step	Time	Bar
Unit	—	—	ms	ms	ms	—
Reference	21 ～ 108	1～ 127	—	—	—	—

2.2. *How to make the data for player piano*

We show the Structure of the edit system how to make the data for player piano in “*Fig 2.(Continued)*” When we have the music data A that the pianist has played, and do not have the music data B that the pianist has played, we extract a characteristic of the pianist from music data A. Then we make the music data B’s musical performance by doing reasoning from characteristic of musical performance and information of musical grammar and score of music B.

2.3. *Search system*

As a result of the analysis, it was found that phrases of the same pattern existing in the same tune are performed in a similar expression. This time, we used DP matching to search for similar phrases.
DP matching is a technique used widely in the field of speech recognition, bioinformatics and so on. It has a feature that can calculate the similarity between two words that are different in a number of characters from each other. In “*Fig.4*”, the route of minimum cost in each point is taken, and the route with the lowest cost is assumed finally to be the optimal path. The cost at that time is defined as the distance between patterns. In this system, this distance is handled as a threshold to judge whether the phrases are similar to each other.
For example, if the cost moves up or to the right, then it is increased by 1. If it moves to the upper right, then it does not increase. Also, if the characters do not correspond in each point, then the cost is increased by 3.

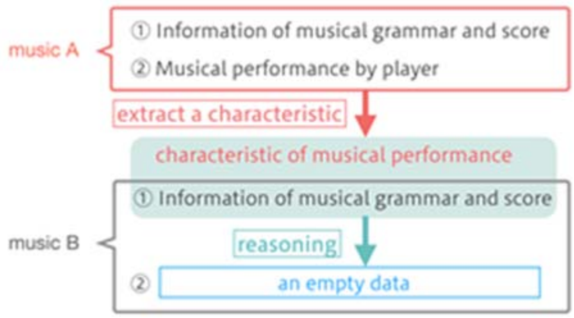



Fig.2. The structure of the edit system

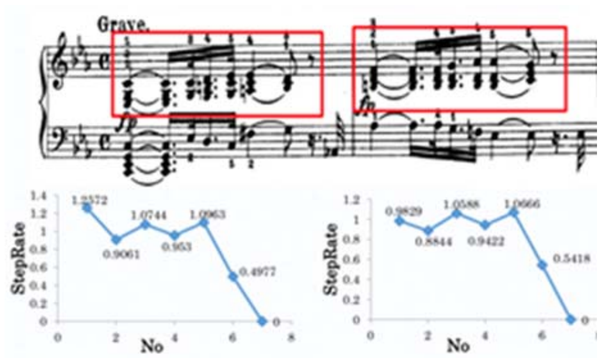


Fig.3. The discovery about the relation of same tunes and similar expression.

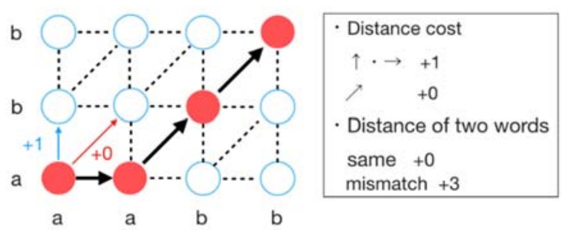


Fig.4. Result of dynamic programming matching method

2.4. *Selection system*

After done the Search system, there are a lot of similar phrases which has the same DP matching points. Similar phrases are phrases we found by doing DP matching from music data. Objective phrase is phrase we want to inference in this time. Therefore, we need to select the best phrases for objective phrases by doing the Selection system. In this year's system, three indicators were used from the viewpoint of music theory: Dynamic symbols, Beats and Steps, and Musical forms.

2.5. *Dynamic symbols*

If the Dynamic symbols are different between similar phrase and objective phrase, the performance is affected even if the phrase is the same, and the search phrase and the similar phrase that matches the dynamics on the score are preferentially selected.

Objective phrase	Similar phrases
pp	ppp
	pp
	p
	mp
	× mf
	× f
	× ff

Fig.5. Ocation about objective phrase's dynamics symbol is pp.

2.6. *Beat and Step*

Based on the musical grammar, it is known that strong beats are closely related to Step. In places considered to be strong beats, similar phrases are selected using the property that the Step value is larger than in other places. The position of the strong beat depends on the rhythms.

2.7. *Musical Forms*

If the Selection system in 2.5. and 2.6. does not narrow down the number of similar phrases, the selection is performed according to the Musical Forms. Musical Forms is a music format such as the Rond Forms or the Sonata Forms. For example, the Sonara Forms has a complex three-part form consisting of a presentation part, a development part, and a reproduction part. Even if different songs were in the sonata format, it was predicted that the presentation part had a presentation part feature, the development part had a development part feature, and each part had a unique feature. We decide same part has similar performance.

3. Inference system

3.1. *Sysmte overview*

The Velo value that the performer is executing for the dynamic symbol on the score is examined, and Velo suitable for the optimal phrase is inferred from the search phrase.

3.2. *Inference by the strength symbol*

In order to obtain the characteristic of the change of the Velo value due to the influence of the intensity sign, the tendency was examined from the performance data of the pianist. The pianist of the surveyed performance data is Gerhard Opitz, and its title is shown in "*Table 2*."

Table.2 The studied title of musical compositions.

	Title of musical compositions.
①	Sonate fur Klavier Nr.8 c-moll Op.13-1 "Grande Sonate Pathetique"
②	Sonate fur Klavier Nr.8 c-moll Op.13-2 "Grande Sonate Pathetique"
③	Sonate fur Klavier Nr.8 c-moll Op.13-3 "Grande Sonate Pathetique"
④	Sonate fur Klavier Nr.23 f-moll Op.57-1 "Appassionata"

As a result, the definition of Velo with detailed strength varied depending on the song, but the one with the high frequency of each Velo value was adopted as the strength symbol, and the results are summarized in "*Table 3*".

Table.3 The Velo value for a composite dynamic mark.

Dynamic mark	ppp	pp	p	mp	mf	f	ff
Velo	15～20	20～45	45～65	65～75	75～80	80～95	95～115

Inference using "*Table 3*" and Velo's inference formula for the xth sound are shown in "Eq. (1)" and "(2)". SVelo refers to the optimal phrase Velo.

$$\mathrm{Velo(x)} = \begin{cases} x = 1 \cap strong\ beat, chose\ maximum\ value\ of\ range \\ x = 1 \cap weak\ beat, chose\ median\ of\ range \end{cases} \quad (1)$$

$$\mathrm{Velo(x)} = Velo(x-1) - SVelo's\ value\ (x \geq 2) \quad (2)$$

4. Inference experiment

4.1. *Experimental method*

An experiment was conducted to compare the music reproduced using the editing support system with the performance of the pianist. The target song is Beethoven Piano Sonata No.14, Op.27-2("Moonlight Sonata"). "*Fig.6*" shows an example of a search phrase.

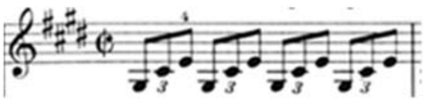



Fig.6. Search phrase

4.2. *Experimental result*

Using graphs and correlations, we compared inferred phrases with actual performance. The Search system found 18 similar phrases. "*Table 4*" shows the 10 data selected by Dynamics symbols. ○ indicates that the numerical value is 80% or more o

Table.4. Selecting method by using Dynamics symbols

1st number	182	214	226	819	831	857	869	1085	1097	1109
1st strong beat	△	△	○	△	○	○	○	△	×	○
2nd strong beat	△	△	○	△	○	○	○	△	×	○
3rd strong beat	△	△	○	×	○	○	○	△	×	○
4th strong beat	△	△	○	△	○	○	○	△	×	○
5th strong beat	△	○	△	△	△	△	×	△	△	×
6th strong beat	△	×	○	×	×	△	○	×	△	△

Table.5. The Velo value for a composite dynamic mark.

1st number	182	214	226	819	831	857	869	1085	1097	1109
Coefficient of correlation	-0.15	0.46	0.72	0.02	0.69	0.55	0.53	0.20	-0.16	0.52

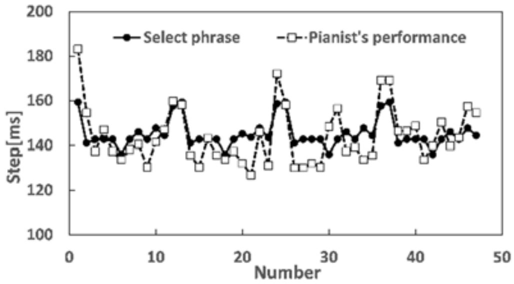


Fig.7. Comparing Select phrase to pianist's performance data.

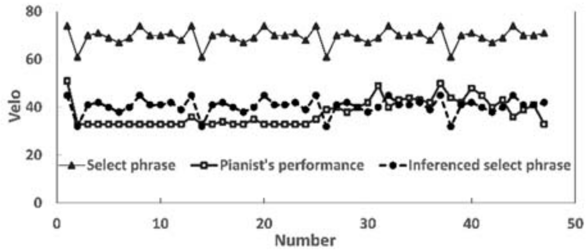


Fig.8. Comparing Select phrase to Infereced select phrase to pianist's performance data

4.3. *Consideration*

Table 5 shows that the phrase selected by the selection system is closer to the actual performance than other similar phrases. From "*Fig.8*", it was confirmed that the Velo was closer to the actual performance. Thus, the selection system and inference system newly added this year are considered to be effective.

5. Conclusion

This year, we introduced a selection system that selects the most suitable phrases and an inference system that infers the Velo values that are likely to be performed by the pianist, to the editing support system. The inference experiment showed that the performance expression of the unedited piece could be inferred. This selection system is based on concepts that are faithful to music theory. However, if the music style is very different, such as contemporary music, it is assumed that this system will not work. To solve this problem, we will test the system's behavior in inferring songs of different music styles and develop a new system that can select the best phrases for classical music of all era.

References

1. Y. Isomichi and N. Ogawa, Pattern Matching by Using Dynamic Progaming, 1975, Vol 16, No.1, pp.15-22.

A Development of a Model CubeSat with an Amateur Radio Transceiver for Education on Satellite Communication

Kentarou Konishi, Masahiro Tokumitsu
*Department of Electronic and Control Engineering, National Institute of Technology (KOSEN), Yonago College, 4448, Hikonacho, Yonago, Tottori 683-8502, Japan**

Taku Takada
Department of Social Design Engineering, National Institute of Technology (KOSEN), Kochi, 200-1, Monobe Otsu Nankoku, Kochi, 783-8508, Japan

Fumio Asai
Mmeber, The Radio Amateur Satellite Corporation, 10605 Concord St, #304, Kensington, MD 20895-2526, USA

Makoto Wakabayashi
Department of Electrical and Computer Engineering, National Institute of Technology (KOSEN), Niihama College, 7-1, Yakumocho, Niihama, Ehime, 792-8580, Japan

Kentarou Kitamura
Department of Mechanical and Electrical Engineering, National Institute of Technology (KOSEN), Tokuyama College, Gakuendai, Shunan, Yamaguchi, 745-8585, Japan
E-mail: tokumitsu@yonago.kosen-ac.jp
www.yonago-k.ac.jp/denshi/stafflab/tokumitsu

Abstract

We report the development of a model CubeSat for satellite communication education. The communication between satellites and ground stations is an essential part of satellite operations. As a critical part of the satellite systems, we focused on satellite communication to train the student so that they can develop and operate the CubeSats. The model CubeSat is equipped with fundamental components, such as an onboard computer, sensors, communication, and power supply, which are necessary for satellites. The software programs in the model CubeSat provide two functions: controlling and monitoring satellites, and transmission of telemetry and image data. The students can study the communication equipment and operations of satellite communication using our model CubeSat.

Keywords: Satellite commCubeSat, Raspberry Pi, Amateur radio, Space engineering.

1. Introduction

Nanosatellites are actively developed and launched all over the world by universities, the Japanese National Institute of Technology (KOSEN), senior high schools, and venture companies. The CubeSat is a nanosatellite, and its standard size is $10 \times 10 \times 10 cm^3$ (1U size). The development of the CubeSat aims to demonstrate

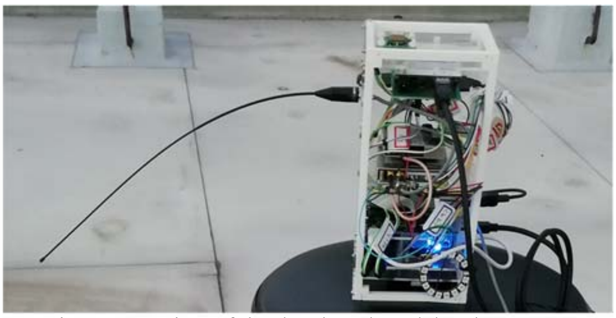


Fig. 1. Overview of the developed model CubeSat.

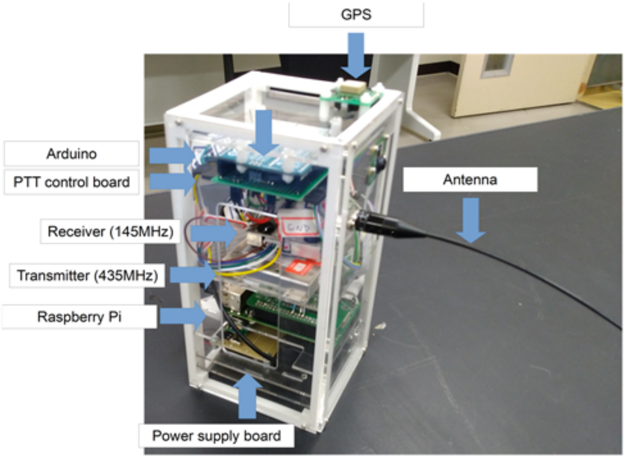


Fig. 2. Configuration of the developed model CubeSat.

experiments of new technologies or new scientific missions in space.

The communication between satellites and ground stations plays a vital role in satellite operation. The CubeSats transmit the mission and housekeeping data to the ground stations. The ground stations monitor the health status of the satellites and controls them by commands. Radio communication for satellite communication systems usually requires radio operator and radio station licenses in each country. For the universities, KOSENs, and senior high schools, education on the radio operators for satellite operation is an essential issue because satellite operation teams need to deploy licensed radio operators for satellite communication.

For the CubeSats, amateur radio bands are often used for communication between satellites and ground stations. The law of the amateur radio also requires radio operator licenses. The CubeSats mainly use the 144 MHz, 435 MHz, and 2.4 GHz bands for communication. Furthermore, the CubeSats use both analog mode and digital mode for satellite operation in amateur satellite communication.

In the universities, KOSENs, and other senior high schools, we consider the practical training material necessary for the education of radio operators. For this purpose, a CubeSat ground model has already been proposed and developed[1] for educating students. Similarly, there are other projects[2,3] that have developed the CubeSats for educational or practical training purposes.

We then extend the CubeSat ground model[1] to the model that is equipped with an amateur radio transceiver. The proposed model CubeSat is equipped with essential components, such as an onboard computer, sensors, a frame, a power supply, a Wi-Fi dongle, and a battery, as satellites.

In this paper, we focus on the education of satellite communication using the model CubeSat, in which the amateur radio transceiver for nanosatellites can be installed into the frame of the CubeSat. We report the prototype version of the developed model CubeSat and introduce the basic functions of the proposed CubeSat.

2. Overview of Model CubeSat

2.1. *Configuration of model CubeSat*

The proposed CubeSat is based on the basic model developed in a previous study[1]. The former model uses a Wi-Fi dongle for radio communication. Fig. 1 shows an overview of the developed model CubeSat. The dimension of the developed model CubeSat is 2 U. Fig. 2 shows the configuration of the developed 2U model CubeSat. The basic configuration of the proposed model CubeSat is the same as the former model. However, we installed additional hardware components as follows:

(1) The frequency control board of the transmitter and the receiver
(2) The transmission control board for CW / FM transceivers and the signal adjustment
(3) The Power supply board for hardware components

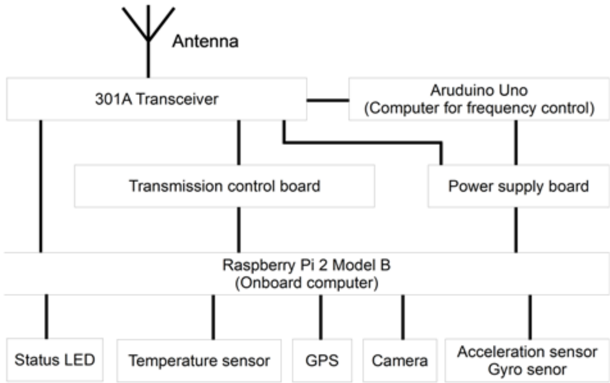


Fig. 3. System diagram of the developed model CubeSat.

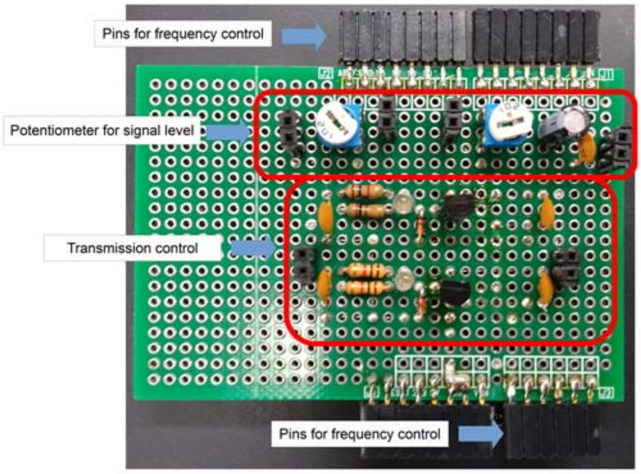


Fig. 4. Transmission control board

The developed model CubeSat is equipped with the Raspberry Pi 2 model B computer board. Fig. 3 shows the system diagram of the proposed CubeSat. The Raspberry Pi works as the main computer for monitoring and controlling the system of the model CubeSat. A temperature sensor, acceleration and gyro sensor, and a camera are installed into the model CubeSat. The proposed CubeSat transmits the sensor values as a beacon through the amateur radio.

For the model CubeSat, we developed both the hardware components and the essential programs. The CubeSat can be controlled by the commands through the amateur radio. The students can learn the communication system of the CubeSat and ground stations.

2.2. *Communication System of the model CubeSat*

The developed model CubeSat equips the amateur radio transceiver of type 301A[4]. The 301A implements FM analog mode / digital mode and CW transmission. The digital mode of the FM transceiver supports a speed rate of up to 1200 bps. Although this speed rate is slow, minimum operations are feasible. The amateur radio transceiver of the developed model CubeSat uses two bands for communication: 435 MHz and 145 MHz. Amateur satellite communications usually use two bands for downlink and uplink. For the developed model CubeSat, the 145 MHz band and the 435 MHz band are used for uplink and downlink, respectively.

We embedded the 301A transceiver into the CubeSat. We implemented the transmission control board for the transceiver. This board has three functions: transmission control, signal level adjustment, and frequency control. The 301A transceiver consists of a separate transmitter and receiver. The transmitter and the receiver require a frequency control by a microcomputer from outside of the transceiver. We implemented the transmission control board for Arduino Uno and its program. Furthermore, 301A requires an audio band signal for communicating digital mode. We also implemented a potentiometer on the transmission control board for adjusting the levels of input and output signals.

3. Communication Functions of the Model CubeSat

The ground stations can monitor and control the developed model CubeSat through digital packet communication as we also developed the software programs for the CubeSat. The developed model CubeSat uses AX.25 for packet communication. The modulation for AX.25 is by AFSK (Audio Frequency Shift-Keying) in which the digital data for transmission is modulated for the audio signals. We chose direwolf[5] for modulation of the transmission data as direwolf supports Linux (Raspbian for the Raspberry Pi is one of the Linux-based distributions).

Table 1 lists the commands implemented in our model. The operators of the ground stations send the command in AX.25 through the amateur radio transceiver. Fig. 5 shows an example of the image captured by the developed model CubeSat.

Table 1. Main commands of the developed model CubeSat.

Command No.	Description
000	Enable the CW beacon
001	Disable the CW beacon
010	Enable the FM packet beacon
011	Disable the FM packet beacon
020	Capture the image
021	Prepare transmission data of the image
022	Transmit the image data
030	Turn on the status LED
090	Reboot the onboard computer

Fig. 5. Example of the captured image.

4. Future work

There are still several future works for the development of the proposed model CubeSat. Although we have developed the model CubeSat, the frequency control of the transceiver uses another microcomputer of the Arduino Uno. We need to develop the frequency control function using the Raspberry Pi to reduce the number of components and simplify the system configuration.

We need to develop educational materials such as lecture materials and an operation manual about the model CubeSat, which will be helpful for students. We have already developed some materials for education on satellite communication[6]. We will develop new materials for the model CubeSat based on existing materials. Finally, we need to evaluate the developed model CubeSat and educational materials based on student experience. Accordingly, improvements must be made by implementing modifications based on the students' feedback.

5. Conclusions

We reported model the development of a model CubeSat prototype for satellite communication education. We focused on satellite communication to train the student so that they can develop the communication system and operate the CubeSats through amateur satellite communication. We implemented an amateur radio transceiver in the developed CubeSat. We also developed the software programs for the model CubeSat for controlling and monitoring satellites and transmitting telemetry and image data. In the future, we plan to improve the frequency control board and prepare documentation of educational materials.

Acknowledgements

This work was supported by Grants-in-Aid for Scientific Research (17K01169) by the Ministry of Education, Culture, Sports, Science and Technology, Japan. The authors would also like to thank members of the KOSEN space collaboration group for valuable comments and encouragement on this work.

References

1. J. Nakaya, T. Takada, Y. Kajimura, H. Tsuchiya, N. Uezono, Y. Sasaoka, S. Ueta, M. Wakabayashi, K. Kitamura, Development of CubeSat Ground Model Extended from CanSat: Application to Space Education at KOSEN, in *the Proceedings of the International Symposium on Space Technology and Science* (ISTS), pp. 2019-u-16, pp.1-6, 2019.
2. The AMSAT, AMSAT CubeSat Simulator, https://github.com/alanbjohnston/CubeSatSim (Accessed on 20/12/2019).
3. The OPUTSAT-KIT Project, OPUSAT-KIT, http://www.sssrc.aero.osakafu-u.ac.jp/activity/opusat-kit-project/ (Accessed on 20/12/2019).
4. Nishimusen corporation, 301A Transceiver for nanosatellite, http://www.nishimusen.co.jp/eisei2008/eisei2008.html (Accessed on 20/12/2019).
5. WB2OSZ, direwolf, https://github.com/wb2osz/direwolf (Accessed on 20/12/2019).
6. M. Tokumitsu, F. Asai, T. Takada, M. Wakabayashi, K. Kitamura, and K. Imai, A Report on Satellite Communication Lectures in the KOSEN Space Group, in the Proceedings in *the 32nd International Symposium on Space Technology and Science* (ISTS), 2019-u-08, pp.1-9, 2019.

Performance Evaluations on Data Estimation Technique with Statistical Properties of Telemetry Data for Corrupted Data in Amateur Satellite Communication

Yusuke Teranishi~~, and~~ Masahiro Tokumitsu
*Department of Electronic and Control Engineering, National Institute of Technology (KOSEN), Yonago College, 4448, Hikonacho, Yonago, Tottori 683-8502, Japan**

Taku Takada
Department of Social Design Engineering, National Institute of Technology (KOSEN), Kochi, 200-1, Monobe Otsu Nankoku, Kochi, 783-8508, Japan

Fumio Asai
Mmeber, The Radio Amateur Satellite Corporation,
10605 Concord St, #304, Kensington, MD 20895-2526, USA

Makoto Wakabayashi
Department of Electrical and Computer Engineering, National Institute of Technology (KOSEN), Niihama College, 7-1, Yakumocho, Niihama, Ehime, 792-8580, Japan
E-mail: tokumitsu@yonago.kosen-ac.jp
www.yonago-k.ac.jp/denshi/stafflab/tokumitsu

Abstract

Herein, we propose a novel data estimation technique for satellite communication and evaluate its performance. The proposed technique estimates the original data from the incorrect data, error detection codes, and supplemental information. First, we investigated the possibility of data estimation for long length data by a brute-force technique, and the results showed that the brute-force technique succeeded in estimating the original 1024-bit-long data from the incorrect data. Next, we applied the proposed data estimation technique to the incorrect data, and the results showed that the proposed technique succeeded in estimating the original 1024-bit-long data from the incorrect data.

Keywords: data estimation, frame check sequence, satellite communication, statistical analysis, telemetry data.

1. Introduction

Nanosatellites are actively developed and launched all over the world by universities, the Japanese National Institute of Technology (KOSEN), and venture companies. CubeSats are nanosatellites whose standard size is $10 \times 10 \times 10\ cm^3$ (1U size), and their development is aimed at experimenting with new technologies or new scientific missions in space.

Amateur radio bands of the 144 MHz band and the 435 MHz band are primarily used for communication between CubeSats and ground stations. However, both of these bands have some limitations with regard to data transmission. The characteristic limitations of the two bands are as follows (Fig. 1):

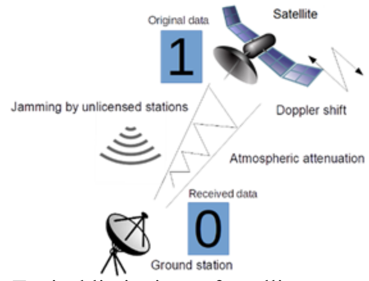


Fig. 1. Typical limitations of satellite communication

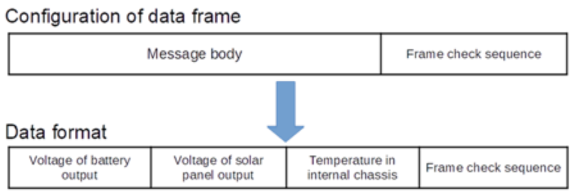


Fig. 2. Typical example of a data frame in satellite communication

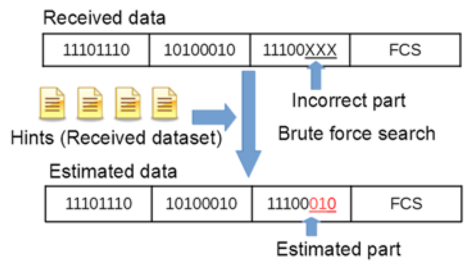


Fig. 3. Overview of the data estimation

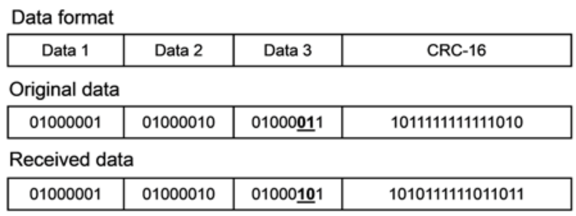


Fig. 4. Typical example of the comparison of correct and incorrect data in the data frame check sequence

(1) The maximum data rate is 9600 bps.
(2) The visible time of the satellites is approximately 15 minutes.
(3) Jamming can occur because of unlicensed radio stations.

The CubeSats transmit the data frames with error correction codes to enable the ground stations to verify the reliability of the data frames. The frame check sequence (FCS) comprises extra data to detect errors in the transmission data. The ground stations verify the received data using the FCS and discard it if it contains errors.

For the CubeSats, the data frame for satellite communication primarily consists of the message part and its FCS (Fig. 2). Furthermore, the CRC-16 of the FCS, which is the primary method used for amateur satellite communication, only allows the detection of the errors of the data frame. Several studies have attempted to correct the errors of the data frame by using the FCS[1]. However, these methods have some limitations in that they only allow the correction of consecutive errors for up to two bits and shorten the data[2]. We propose a novel technique to the original data frames from corrupted data frames that overcomes the limitations of the previous methods. The proposed method attempts to estimate the original data to correct the errors by using the statistical properties of the received data set and the FCS of the received data. Furthermore, the proposed method can be applied to data that contain randomized errors.

In this study, we evaluate the performance of the proposed method by applying it to estimate randomized test data and real data obtained by FUNcube-1[3,4]. In addition, we discuss the feasibility of the proposed approach for improving the reliability of data for satellite communication.

2. Data Estimation Technique

The proposed technique estimates the original data by using the statistical properties of the received data set and the FCS of the received data. Fig. 1 shows an overview of the proposed data estimation technique.

The basic principle of the proposed technique involves the use of the brute-force approach for estimating the original data transmitted by the CubeSats. In Fig. 1, the received data contains two consecutive bits. The proposed technique, which estimates the original data by the brute-force search, consists of two steps. (1) The brute-force search assigns either a one or a zero for each bit. (2) The proposed technique then calculates the FCS of the estimated data and compares it with the FCS of the received data. At the end of the data estimation process, the proposed technique accepts the estimated data for which both frame check sequences match.

Although the naïve brute-force technique is quite simple, collisions of the estimated data occur during its calculation process. We introduce the statistical properties of the received data set to the proposed algorithm to reduce collisions among the estimated data.

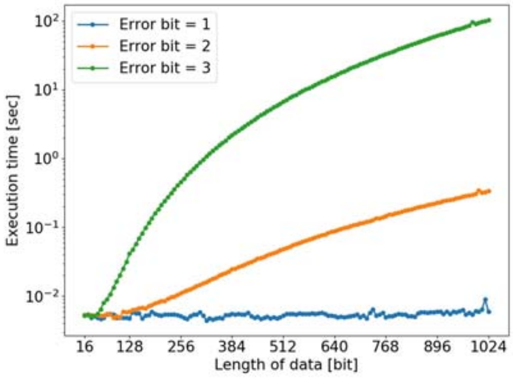


Fig.5. Execution time of the proposed technique.

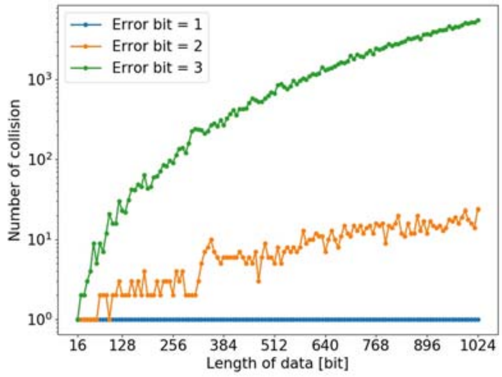


Fig. 6. Collisions of the estimated data

The brute-force technique attempts to generate and verify all combinations of the bit arrangement, and the proposed technique will generate data that have the same FCS.

We need to avoid generating unrelated data in the calculation process so that the proposed technique can exclude the data whenever some parts of the estimated data contain values outside their standard deviations. Although the brute-force technique is simple, some assumptions with regard to the received data can contribute to reducing the collisions of the estimated data and the computation time of the proposed algorithm.

3. Simulations

3.1. *Feasibility of data estimation*

We examine the proposed technique to investigate its feasibility with regard to data estimation. We generate the original data, where each bit is randomly assigned a value of either one or zero. For the tests, we flip some bits (less than four bits) of the original data to create the test data. We also calculate the FCSs for both the original and the test data. Then, we perform data estimation by checking whether or not both the FCSs match. For the analysis, we compare the length of the test and execution time; in addition, we also consider the length of the test and collisions of the estimated data.

Fig. 5 shows the execution time of data estimation performed using the proposed technique. The brute-force technique was successful in estimating the corrupted data with errors within three bits. For 1024-bit-long data with two bit errors, the proposed technique estimated the original data in less than 1 s.

Fig. 6 shows the number of data collisions in the data estimation process with respect to the length of the data. For errors within three bits, some estimated data collided with other data; that is, their FCSs matched each other.

3.2. *Performance evaluation*

We apply the proposed technique to the real telemetry data obtained from FUNcube. The test data consists of 14 fields, each with a length of 16 bits. Thus, the total length of the test data frame is 224 bits. We repeat the data estimation 100 times to measure the average performance.

Table 1 shows the number of collisions with respect to the number of erroneous bits. For the single bit errors, the proposed technique is consistently successful in estimating the original data with no collisions among the data. However, as the number of erroneous bits increases, the data estimated by the proposed technique comprises some errors. With regard to the four bit errors, although the number of collisions increases sharply to 3214.14 when statistical properties are not considered, the proposed technique narrowed the data points that could potentially be the original data to 52.03 by considering the statistical properties.

Table 2 shows typical examples of the estimated data with the same FCSs. However, each estimated data has different fields of values. For example, the FCS of estimated data 1 is the same as that of the original data. However, it can be excluded from the candidate data because data 2 deviates its standard deviation.

4. Discussion

According to the results of the simulation, the proposed technique can estimate the original data by using the statistical properties of the received data set.

Table 1. Average number of collisions obtained via the proposed technique with and without the statistical properties

Number of erroneous bits	1	2	3	4
Without considering statistical properties	1.00	3.50	69.53	3214.14
Considering statistical properties	1.00	1.27	5.06	52.03

Original data

	Data 1	Data 2	Data 3	Data 4	CRC-16
Value	11.6875	11.8438	13.9375	17.5469	EF26
Bit data	0100100111011000	0100100111101100	0100101011111000	0100110001100011	1110111100100110

Received data (The FCS is E1FA.)

	Data 1	Data 2	Data 3	Data 4	CRC-16
Value	11.6875	11.8438	12.4375	17.5469	EF26
Bit data	0100100111011000	0100100111101100	0100101000111000	0100110001100011	1110111100100110

Estimated data 1 (Data 1, data 2, and data 3 contain errors, and they deviate from their standard deviations.)

	Data 1	Data 2	Data 3	Data 4	CRC-16
Value	23.375	15.8438	12.4375	17.5469	EF26
Bit data	0100110111011000	0100101111101100	0100101000111000	0100110001100011	1110111100100110

Estimated data 2 (Estimated data 3 and data 4 are within their standard deviations.)

	Data 1	Data 2	Data 3	Data 4	CRC-16
Value	11.6875	11.8438	12.4375	16.7969	EF26
Bit data	0100100111011000	0100100111101100	0100101000111000	0100110000110011	1110111100100110

Fig. 7. Typical example of the estimated data; Gray-colored cells indicate that the data contain errors.

Furthermore, the proposed technique substantially decreases the number of collisions by the use of the statistical properties when compared with that obtained by the naïve brute-force technique. However, some estimated data still collided with other data. We need to improve the proposed technique to reduce the number of collisions among the estimated data.

In the simulation results, some estimated data with the same FCSs consisted of different field values. The estimated data with different field values compared to those of the original data can be used as supplemental data because the errors between the estimated and original values are small. Similar estimated data would be available for interpolated data instead of missing data. For the data estimation process, however, we need to improve the search performance of the proposed technique with regard to finding the original data.

5. Conclusions

We proposed a novel data estimation technique for satellite communication and evaluated its performance. The proposed technique estimates the original data from the incorrect data, error detection codes, and statistical properties of the received data set. The results of the first simulation showed that the brute-force technique succeeded in estimating the original data, with the length of 1024 bits, from the incorrect data. Based on the results of the second simulation, we showed that the proposed technique succeeded in estimating the original data, with the length of 224 bits, from the incorrect data.

Acknowledgements

This work was supported by the Grants-in-Aid for Strategic Information and Communications R&D Promotion Programme in 2018 by the Ministry of Internal Affairs and Communications, Japan.

References

1. P. Oechslin, “Making a Faster Cryptanalytical Time-Memory Trade-Off,” *in Advances in Cryptology - CRYPTO 2003. CRYPTO 2003, Lecture Notes in Computer Science 2729* (Springer, Berlin, Heidelberg, 2003).
2. M. Tokumitsu, K. Mikami, F. Asai, T. Takada, M. Wakabayashi, and Y. Ishida, "A Data Estimation Technique for Incomplete Telemetry Data based on a Genetic Algorithm with Data’ Statistical Properties," *in Proceedings of the 2019 International Conference on Artificial Life and Robotics*, ICAROB'18 (2018).
3. The FUNCube project, “Welcome to the FUNcube Web Site,” https://funcube.org.uk/. (2019) (Accessed on 20/12/2019)
4. D. Johnson, “G4DPZ, The Data Warehouse,” http://warehouse.funcube.org.uk/realtime.html?satelliteId=2, (2013). (Accessed on 20/12/2019)

Heritage Building Design Properties: Development of As Built Drawing by UAV Application via 3D Laser Scanner

D. Hazry, M. A. Azizan, N.Z Noriman, N. Romeli

Centre of Excellence for Unmanned Aerial Systems, Universiti Malaysia Perlis

E-mail: hazry@unimap.edu.my, azizi@unimap.edu.my, niknoriman@unimap.edu.my, norsyakilah@unimap.edu.my

Abstract

Inadequate information that represent the existing figure of the building after transmission of years and the eagerness of the Malaysian's architect to develop frantically modern development design that gradually neglecting the essence of cultural perseverance of the heritage building has been brawled excessively. Taking a human – factor constraint to conducted the physical survey, the technologies such as computers and digital media have given beneficial value that are time saving and less risky. Hence, the research aim for the development of the as-built drawing for Masjid Alwi, Kangar, Perlis has been completed via UAV application and 3D laser scanner, and results were recorded on the structural and architectural properties of Masjid Alwi that has been analyzed accordingly. Therefore, the needs of having an adequate as-built drawing due to the depletion of the ancient documentation and as a baseline for the designer maintaining the cultural identity in a way to design new development is imperative and possible with the aid of UAV application and 3D scanning.

Keywords: Heritage Building Design, As Built Drawing, UAV Application, 3D Laser Scanner

1. Heritage Building as a Baseline for Architect to preserve the Cultural Representation

The heritage building embraced the essence of cultural representation in each nation, which influenced by cultural inclusivity of the designer (Baker, 2014; Basir et. al, 2014). The evolution of the design ignited from race, religion, and country and state in the era of circulation. Current practice on Malaysian construction industry, a lackadaisical of adequate information that represent the existing figure of the building after transmission of years (Shukor et.al, 2015). In addition, the eagerness of the Malaysian's architect to develop frantically modern development design that gradually neglecting the essence of cultural perseverance of the heritage building has been brawled excessively (Newland, 2015). Taking a human – factor constraint (Samad et. Al, 2013; Harun, 2011) to conducted the physical survey, the technologies such as computers and digital media have given new opportunities in the obtaining the information on the heritage buildings that are time saving and less risky. The aim of this research is to reconstruct and develop a new as-built drawing of the heritage building by using the data analyzed from 3D laser scanner, UAV photogrammetry, and Building Information Modelling (BIM) software. The objective of the research is to generates as-built drawing of heritage buildings using BIM software (He et. al, 2017) and produce a guideline on design properties of heritage building in Malaysia.

2. UAV Application via 3D Laser Scanner

The Unmanned Aerial Vehicle (UAV) system is the low-cost system and flexibility for photogrammetry and remote sensing. Thus, this system becomes a new revolution in photogrammetry and suitable solution compared to other mobile mapping systems (Samad, 2013). The UAV is basic principle based on the photogrammetry. It is used camera on the UAV, flight planning and maps and analysis same as in photogrammetry but in low altitude. This low cost and low altitude aerial vehicle can give rapidly data and can update the data by real-time without waiting for long term like satellite images (Everaerts, 2008). Micro Aerial cars or as known as 'Flying Drones'make themselves useful in a number of architecture, engineering, and creation (AEC) packages. It has been widely used for developing 3D digital models of exterior facades, plan renovations thru virtual 3D simulation and even to mechanically generate 3-D printable prefabricated structures (Daftry, Hoppe, & Bischof, 2015). The Terrestrial Laser scanning is like their airborne counterparts; these are active sensors that emit laser signals to calculate distances based on the time delay of the returned laser pulses. They are able to record a dense array of distance return values (e.g., in the order of several laser returns per square centimeter over ranges of 50 m), which are assembled into highly detailed, digital 3-D landscape models that can be transformed into GIS DEMs (Sutthivanich, et.al, 2012). TLS provide dense point clouds measured along a predefined frame and representing the 3D geometry of a scene surface. The 3D coordinates of points of interest are then indirectly obtained through point cloud processing. Due to the development of ever more powerful stereo-matching algorithms, photogrammetry methods also lead to the creation of dense point clouds (Lachat et al., 2017), (Evgenikou & Georgopoulos, 2015). BIM is thus prone to being 'blind and deaf' to ongoing AECO processes, whilst an 'as-built' version can extra reliably and usefully assist information exchange and decision-making in the course of the project life-cycle (Zhong et al., 2017),(He et al., 2017) that in a way to actively tracking the changes in physical of the building, as built drawing helps in clarify a clear deviation and explicit content are obtained to act as a design properties guideline to the Malaysian architect in designing urban – heritage mosque.

3. Research Methodology

A preliminary study has been conducted as an initial exploration and important step before starting a project which include screening of literature review, pre site investigation to Masjid Alwi surrounding to predict the contingencies factor may rise from the surrounding especially wind factor and animal factor during the investigation. The data acquisition for this research is divided into two types of data which is from aerial photograph and land survey (3D laser scanning). For the aerial photo, the image is captured using DJI Inspire 2, it can capture with the IMU data to provide the geometrical position to the centimeter level with the help of PIX4D Mapper software that has been recommended compared with other software. IMU data are for the orientation of the photo. So, the point cloud has a position and true orientation that is Omega, Phi, Kappa, the roughness and evenness of the surface can easily distinguish. While for the land survey, 3D laser scanning are done using Terrestrial Laser Scanning (TLS) or Trimble 3D laser scanner. The second phase of this research is data processing. It divides into two data types. First is 3D modeling. To produce 3D modeling by using PIX4D Mapper software, it has 3 main step which is "Initial Processing", "Point Cloud and Mesh", and "DSM, Orthomosaic". Meanwhile, the data 3D laser scanning also process by using Building Information Software (BIM) to verify the as-built drawing. The last phase of this method included results and analysis of the metadata that will produce the final output of research.

1. Results and Discussion

4.1 Data Collection

UAV data collection using DJI Inspire 2 drone with DJI Ground Station App. Side lap and overlap are decisions based on the altitude fly and condition of the place. For this project, overlap and side lap are 80% and 70% for the area. Altitude is 100 m above MSL. There are several number of flight has been taken to ensure that all the acquired images are overlap with each other and can be process using Pix4D. Terrestrial Laser Scanner (TLS). Upon using laser scanner to collect data for internal and external of the mosque, marking area are need to locate the certain points of laser scanner in order to do the scanning process. For this procedure, traversing method have been used to mark the coordinates for inside and outside of the mosque. There are 10 points of coordinate that have been marked in order to do the scanning process for the external area of the mosque. The scanning process need to be done in bright and clear weather in order to get better scanning images. For the internal side, the scanning process have been done during night time after the Isya' prayer time in order to avoid noise and disturb people to do the prayers. For the internal data collection, the laser scanner has been moved into four separate place inside the mosque: - left-wing, right-wing, center area of mosque and level 1 prayer zone. This is because each of these place has been separated by walls in the mosque.

4.2 *Data Processing*

Data processing for this study are involving laser and the conventional method. For the UAV processing used Pix4D Mapper, both use AutoCAD as data analysis processing. Using Pix4D Mapper, the initial processing involved the key point extraction and matching that leads to camera model optimization and geolocation of the GPD/GCP.

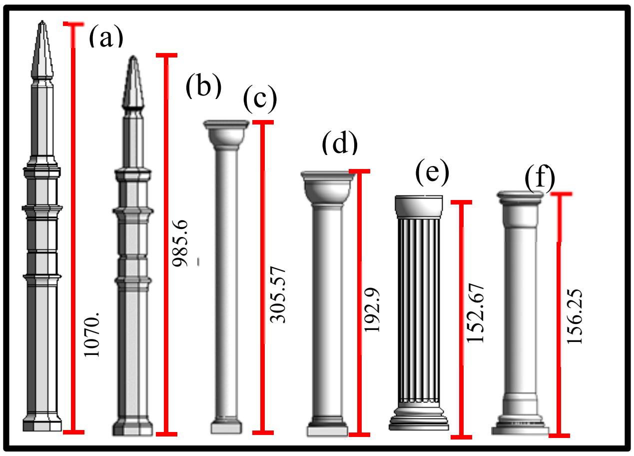


Fig 4.1: Type Roman Pillars That Use for Alwi Mosque and Height for Each Pillar.

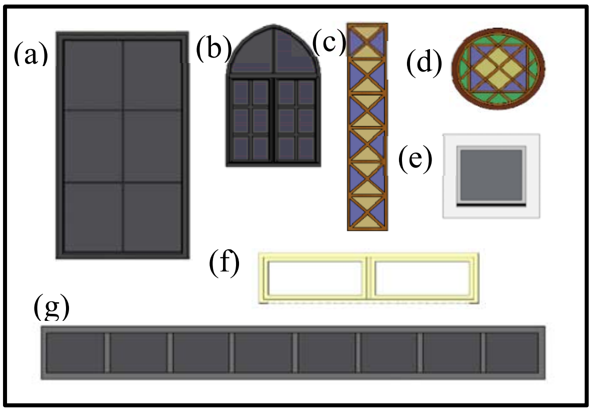


Fig 4.2: The Types of Windows in The Alwi Mosque

Upon the point cloud mesh phase, the point densification has been implemented and the 3D Textured Mesh will be produced. Therefore, the digital surface model (DSM) and orthomosaic information will be produced. Meanwhile processing the data obtain from terrestrial laser scanning involved initial site work traversing on external site of the mosque that apply Laser scanning inside and outside of the mosque that leads to pint cloud processing using magnet collage and 3D point cloud densification. Finally, exporting the data into Autodesk Revit using the selected BIM modification for pillar, inside and outside structures of the mosque. Figure 4.1, Figure 4.2 and Table 1 represent the findings obtain from the UAV application and 3D terrestrial laser scanning that can be used as a guideline for the designer to determine the conceptual design in orders to preserve the heritage building in Malaysia. The types of Roman Pillars as presented in Figure 4.1 completed with dimension and measurement has been rarely used in Malaysia's new blue print, meanwhile the window doors types that has influences of ancient Islamic design has been determined the size and dimension. The findings are significant towards urban – heritage building design in Malaysia that can be obtain from the as built drawing produced from the UAV Investigation and 3D scanning.

Table 1. Design Properties Available at Alwi Mosque

Design Properties	Qty	Size (cm)
Wall (thickness)		
▪ 37.50 cm		
▪ 25.00 cm		
▪ 30.00 cm		
Columns (refer Fig. 4.1)		
• Concrete		
(a) Round		
▪ 32.00 cm diameter	16	
(b) Rectangular		
▪ (a)	8	30 X 30
▪ (b)	3	50 X 50
▪ (c)	3	50 X 37
▪ (d)	2	15 X 25
▪ (e)	1	15 X 15
• Roman		
▪ (a)	2	1070.04
▪ (b)	4	985.67
▪ (c)	15	365.57
▪ (d)	88	192.97
▪ (e)	6	152.67
▪ (f)	2	156.25
Door		
• Glass		
▪ Main Door	1	248 X 536
▪ Sliding	5	209 X 177.5
▪ Double	6	200 X 250
• Wooden		
▪ Single		
o Door tower	1	203 X 76
o Door room	1	200 X 93
▪ Double	1	215 X 110
• Aluminum	1	200 X 80
Window (refer Fig. 4.2)		
• Colonial window		
➢ (a)	7	176.22 X 324.27
➢ (b)	8	209.61 X 120
• Classic		
➢ (c)	20	265.89 X 44.31
	5	135.50 X 44.31
➢ (d)	51	75 (diameter)
	8	130 (diameter)
• Fixed		
➢ (e) Single glass	4	50 X 50
➢ (f) Double glasses	1	205 X 62
	3	290 X 283
	6	200 X 250
	1	160 X 283
➢ (g) Eight glasses	1	65 X 535

5. Conclusion

The design properties view processed using Revit software. The plan shows fundamental of the as built drawing and depicted the 3D shape of the mosque including the dimension of the whole building. Thus, a baseline on new heritage design properties criterion will be created.

6. Acknowledgement

Special thanks to Centre of Excellence for Unmanned Aerial Systems, Universiti Malaysia Perlis for the bountiful facilities and technical support towards the data inspection process.

7. References

1. Aylin Orbasli 2008. Architectural Conservation. Blackwell Publishing, Oxford
2. Abdul Shukor, S. A., Wong, R., Rushforth, E., Basah, S. N., & Zakaria, A. (2015). 3D terrestrial laser scanner for managing existing building. Jurnal Teknologi, 76(12), 133–139. https://doi.org/10.11113/jt.v76.5895
3. Baker, W. (2014). The Role of Cultural Awareness. https://doi.org/10.1177/0033688208091144
4. Basir, W. N. F. W. A., Setan, H., Majid, Z., & Chong, A. (2014). Geospatial database for heritage building conservation. In IOP Conference Series: Earth and Environmental Science (Vol. 18). Institute of Physics Publishing. https://doi.org/10.1088/1755-1315/18/1/012087
5. Bortoluzzi, B., Efremov, I., Medina, C., Sobieraj, D., & McArthur, J. J. (2019). Automating the creation of building information models for existing buildings. Automation in Construction, 105. https://doi.org/10.1016/j.autcon.2019.102838
6. Daftry, S., Hoppe, C., & Bischof, H. (2015). Building with drones: Accurate 3D facade reconstruction using MAVs. Proceedings - IEEE International Conference on Robotics and Automation, 2015–June (June), 3487–3494. https://doi.org/10.1109/ICRA.2015.7139681
7. Eisenbeiss, H., & Zhang, L. (2006). Comparison of DSM
8. Everaerts, J. (2008). The use of unmanned aerial vehicles (UAVs) for remote sensing and mapping. 1-2.
9. Evgenikou, V., & Georgopoulos, A. (2015). Investigating 3D reconstruction methods for small artifacts. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 40(5W4), 101–108. https://doi.org/10.5194/isprsarchives-XL-5-W4-101-2015
10. Harun, S. N. (2011). Heritage building conservation in Malaysia: Experience and challenges. Procedia Engineering, 20, 41–53. https://doi.org/10.1016/j.proeng.2011.11.137
11. He, Q., Wang, G., Luo, L., Shi, Q., Xie, J., & Meng, X. (2017). Mapping the managerial areas of Building Information Modeling (BIM) using scientometric analysis. International Journal of Project Management, 35(4), 670–685. https://doi.org/10.1016/j.ijproman.2016.08.001
12. Lachat, E., Landes, T., & Grussenmeyer, P. (2017). Investigation of a combined surveying and scanning device: The trimble SX10 scanning total station. Sensors (Switzerland), 17(4). https://doi.org/10.3390/s17040730
13. Newland, J. (2015). Generating as-built drawings as a project gets built.
14. Samad, A. M., Kamarulzaman, N., Hamdani, M. A., Mastor, T. A., & Hashim, K. A. (2013). The potential of Unmanned Aerial Vehicle (UAV) for civilian and mapping application. 2013 IEEE 3rd International Conference on System Engineering and Technology. doi:10.1109/icsengt.2013.6650191

A Virtual System for Measurements and Analysis of the Respiratory Sounds for Diagnosis of Respiratory System

Ali AlMejrad

Biomedical Technology Department, Applied Medical Sciences College,
King Saud University, P.O.Box 10219, Riyadh 11433, Kingdom of Saudi Arabia
E-Mail: amejrad@ksu.edu.sa

Abstract

Respiratory problem is one of the most common health problems occurring in Saudi Arabia due to the continuous changes of the weather in addition to surface winds that cause dust during all seasons in the year. This problem affects most people and especially children and elderly in addition to the adults who have respiratory problems such as asthma. This is one of the problem that need emergency care as soon as it occurred. The goal of our research is to develop a compact respiratory diagnostic system using advanced signal processing that can be used remotely via the virtual instrumentation technology to help accurately diagnosis at early stages of respiratory diseases. It can also overcome the lack of expert physicians in rural regions and some urban clinics or health centers. The proposed system will be implemented using virtual instrumentation (VI) that consists of computer, microphone with simple analog circuit, digitizer and LabVIEW software (National Instruments Inc.). VI has been designed for easy measurement and analysis. In addition to that, it has features and ability to control the whole system of acquisition, play, display, processing and advanced analysis of the different acquired respiratory signals. Respiratory signals obtained from our system are analyzed for diagnosis purposes using advanced signal processing techniques in time and frequency domains in addition to the analysis of respiratory parameters. The obtained results can be printed in a report format including acquired respiratory signal, filtered signal, extracted segment for selection of phase, power spectral density, and analyzed respiratory parameters as a reference for diagnosis purpose. The developed system has been tested with real respiratory signals of normal and abnormal cases that proved to be efficient system while dealing with many respiratory problems conditions.

Keywords- Respiratory disease; Virtual Instrument; Signal Processing; power spectrum

1. Introduction

Respiratory problem is one of the most common health problems occurring in Saudi Arabia due to the continuous changes of the weather in addition to surface winds that cause dust during all seasons in the year. This problem affects most people and especially children and elderly in addition to the adults who have respiratory problems such as asthma [1-2]. This is one of the problem that need emergency care as soon as it occurred.

To diagnose these problems affecting the respiratory system, different methods and devices have been used. One of the widely used methods is auscultation of lung sound using stethoscope invented in 1816. It is cheap, easy to use and noninvasive but it is less accurate due to limitation of human ear which cannot distinguish the contents of mixed respiratory signal in addition to low signal to noise ratio due to deficiency in analyzing such signal [3-4]. Other methods and devices using many technologies based on computer technology have been developed for improving the monitoring and quantitative analysis of respiratory sounds [5].

Analysis of one cycle of respiratory sounds consisting of three phases, inspiration, respiration pause, and expiration quantitatively depends on the recording site of the respiratory system. The common three recording sites are the mouth, trachea and chest wall. The respiratory sounds are classified as normal or abnormal called adventitious. The respiratory sound heard on the chest wall of a healthy person is called normal respiratory sounds. The adventitious sounds are classified as two classes continuous and discontinuous. Continuous sounds are wheezes and rhonchi. Wheezes sound is as a whistle due to obstructions located in the lungs. Wheezes in the frequency domain have many harmonics with fundamental frequency above 400 Hz, and in the time domain has a time period which is longer than 250 ms. while rhonchi sound is low pitched and has a time duration of longer than 250 ms, with frequency of around 200 Hz or less. In the other hand, discontinuous adventitious sounds are called crackle sounds which occur due to either the sudden opening of bronchial alveoli called fine crackles, or the air bubbling through excessive secretions in the lung called coarse crackles [6,7].

One of the important techniques is the spectral analysis of this signal using spectrum analysis [8-11]. This method will help in displaying quantitative Respiratory information that is of great importance as being will help in the clinical assessment of respiratory system by distinguishing between healthy and diseased respiratory system [12-13].

Today, most of physicians already have computers which connected to the network and Internet that allow remote work. For most physicians, a system dedicated to one task and local use is expensive. The goal of our research is to develop a compact respiratory diagnostic system using advanced signal processing that can be used remotely via the virtual instrumentation technology to help accurately diagnosis at early stages of respiratory diseases. It can also overcome the lack of expert physicians in rural regions and some urban clinics or health centers. The proposed system will be implemented using virtual instrumentation (VI) that consists of computer, microphone with simple analog circuit, digitizer and LabVIEW software (National Instruments Inc.) [14]. VI has been designed for easy measurement and analysis. In addition to that, it has features and ability to control the whole system of acquisition, play, display, processing and advanced analysis of the different acquired respiratory signals. Respiratory signals obtained from our system are analyzed for diagnosis purposes using advanced signal processing techniques in time and frequency domains in addition to the analysis of respiratory parameters. The obtained results can be printed in a report format including acquired respiratory signal, filtered signal, extracted segment for selection of phase, power spectral density, and analyzed respiratory parameters as a reference for diagnosis purpose. To validate the developed system, it will be tested with real respiratory signals of normal and abnormal cases.

2. Materials and Methods

2.1 Hardware

The system hardware is composed of dedicated microphone, analog to digital converter interface card (ADC) and a personal computer (PC) containing a virtual instrument which is implemented based on LabVIEW software (Fig. 1).

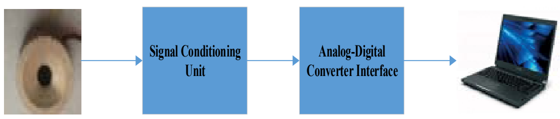


Fig. 1 Block Diagram of computer-Based Respiratory System

The Respiratory signals were recorded from subjects using dedicated microphone. To minimize the background noise, the amplification gain was set to obtain a clean spectrum on the spectral display. The audio output of analog circuit unit is sampled at 11,025 Hz with 12-bit resolution and then sent to a PC via the interface card [15].

2.2 Software

A compact program, Virtual Respiratory Measurements and Analysis System (VRMAS) is designed and developed based on the hardware described above using the graphical programming language (LabVIEW). This graphical programming language (GPL) is a powerful graphical development used for signal acquisition, acquired signal analysis and display of processed data [14]. GPL has a feature of its ability to build our instrument as a virtual instrument (VI). VI represent a PC-based control system that is used to acquire data from microphone and then manipulate the acquired data using a very high level graphical environment. Also using some special icons, specific methods and a combination of different measuring hardware in VI, it is possible to simulate our instrument in the computer. A VI is a system comprising of software especially designed software and a hardware to convert the measured inputs to digital to be processed or analyzed by the computer and then displayed on the computer screen.

VRMAS software is developed based on state diagram which consists of self-produced 4 modules without the need for any external add-ons. The main modules of VRMAS are acquisition of the audio Respiratory signal, advanced signal processing of acquired or simulated Respiratory data, analysis of Respiratory data and finally the printing and help tools as shown in Fig. 2. Using the main module of VRMAS, the user will be able to control the whole system of acquisition, play, display, processing and advanced analysis of the different acquired in-vivo clinical real data.

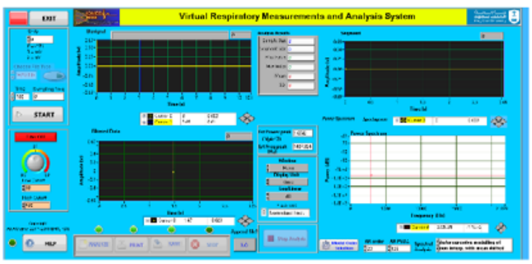


Fig. 2 GUI and Front Panel of Developed Virtual Respiratory Measurements and Analysis System

2.3 Respiratory Signal Processing Techniques

Many spectral methods have been used for the analysis Respiratory. These methods are classified as parametric and nonparametric [16]. The simplest classical nonparametric method used in most cases is Fast Fourier transform (FFT) due to simplicity and high processing speed. While parametric methods such as autoregressive (AR) method has some advantages such as smoother spectral components that enable to distinguish independently the different frequency bands. Therefore, we will apply both methods in our system to investigate the influence of spectral analyses using FFT method as compared with AR method on the spectral characteristics of Respiratory signals in normal and abnormal cases.

FFT spectrum is calculated using usually either the Blackman Tukey or periodogram method. In periodogram method, the appropriate segment length selected depends on the data. If shorter window is selected, greater stationarity will be obtained but the resolution of spectral estimates will be lower. Therefore, a trade-off between the resolution of spectral estimates and stationarity has to be considered. Hence the data are first divided into overlapped segments of 128 intervals followed by windowing each segment using selected window. Then FFT spectrum is calculated for each windowed segment followed finally by averaging the segment spectra. While AR spectrum is calculated by fitting the selected order of AR model into our Doppler signal data. The AR spectrum is obtained from the estimated AR model parameters which are solved using a forward-backward least squares method. In this method if the order of the model is increased, this will lead to decrease of estimated variance. To choose the appropriate model order, a compensation for the over-parameterizing is required. One of the well-known criterion for model order estimation used is the Akaike information criterion (AIC) [17]. The model order p that gives the minimum value of AIC (p) is selected and based on the literature and studies of Respiratory signals an order of 5 is selected [18].

3. Results of Software Application and Discussion

VRMAS Software is running as shown in the flowchart in Fig. 3. Acquired Respiratory data can be recorded and saved in files in different format in hard disk which can be loaded for data analysis. All modules are called from VRMAS main module and will be described below in the following sections.

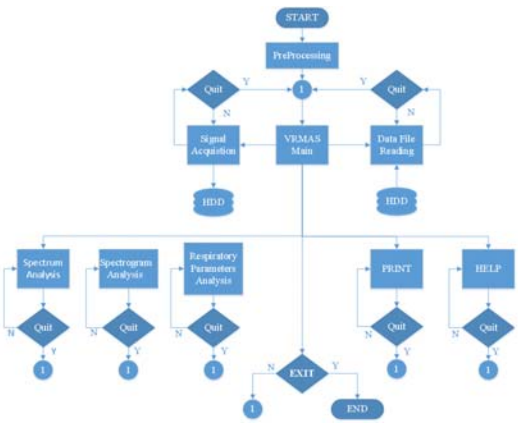


Fig. 3 VRMAS Software Flow Chart

3.1. Acquisition and recording of respiratory signals

Acquisition and recording of Respiratory signals from normal and abnormal cases, as digitized by DAQ card are displayed and also can be saved if required. The sampling frequency and recording interval can be adjusted using acquisition setup. Sampling frequency is selected to digitize the Respiratory signal sufficiently based on Nyquest theory. In the system control module, the Respiratory signal can be recorded both in txt and wav file formats. In Fig. 4, shows graphical user interface of VRMAS showing a case study of one cycle of Inspiration and expiration of Normal Vesicular Respiratory signal.

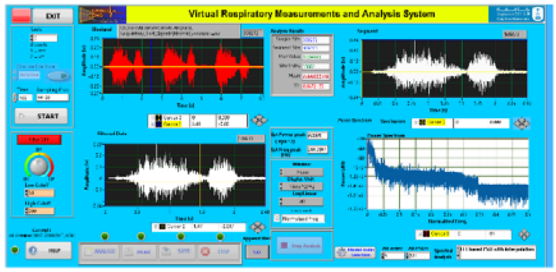


Fig. 4 GUI and Front Panel of Developed Virtual Respiratory Analysis System showing one cycle of Inspiration and expiration of Normal Vesicular Respiratory signal.

3.2. Reading of respiratory data from a file

Recorded Respiratory signal data is loaded as array from either wave format or text format file for time or frequency analysis. The reading of Respiratory data from file VRMAS module is shown in Fig. 5. As the Respiratory signal can be filtered, also it is possible to extract any portion of the Respiratory signal in the time domain to be available for required analyses of preprocessed Respiratory signal.

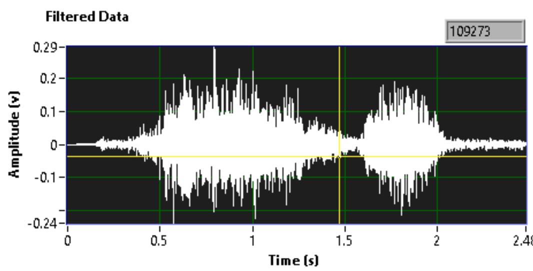


Fig. 5 Filtered Normal Vesicular Respiratory signal acquired from Respiratory System

3.3 Displaying and playing of respiratory signals

As Respiratory audio date is read from the file, Respiratory signal is playing using computer sound card

3.4. Analysis of respiratory signals

VRMAS has three analysis modules, spectrum analysis, spectrogram analysis and Respiratory parameters analysis that can be set up as shown in Fig. 6. They will be discussed in the following subsections below.

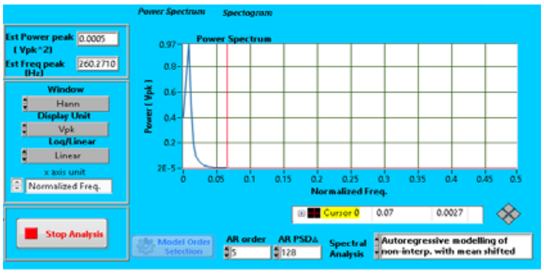


Fig. 6 Analysis of Respiratory Signals Window's Setup

3.5. Spectrum analysis

VRMAS spectrum analysis module is able to analyze Respiratory signal in frequency domain using Fast Fourier Transform (FFT) and Autoregressive modelling (AR). In this module linear or logarithmic input specifies linear or log spectrum output. Also display unit can be selected to one of the values listed in the display unit section display the output unit for the spectrum either as amplitude spectral density or power spectral density (PSD). In window section, to eliminate the spectral leakage during computing power spectrum, if required any window type can be selected from the list of many well- known window types such as Hamming, Henning, Blackman, Blackman- Harris, or Flat top. In this case shown a 1024 point FFT is applied with window function.

In the other hand autoregressive modelling is designed to be used as another method of spectral analysis for comparison to search the best method for diagnosis of diseases. AR modelling requires some parameters to select based on the model chosen. For the respiratory signals analyzed, the mean of the orders, which minimize the order criterions, is found to be 5 but highly varying in different segments. So it's better to select different orders for different segments to avoid overestimation or inadequate estimation. In some segments lower orders can be sufficient for estimation. Selecting a higher order for such a data sequence causes high frequency peaks, which are not desired. Also the order changes with the sampling frequency such that for higher sampling rates the adequate orders are higher than the ones for lower sampling rates. To select the best order, one tool of our system for model order selection has been built using AIC technique as shown in Fig. 7 for example of acquired crackle respiratory signal.

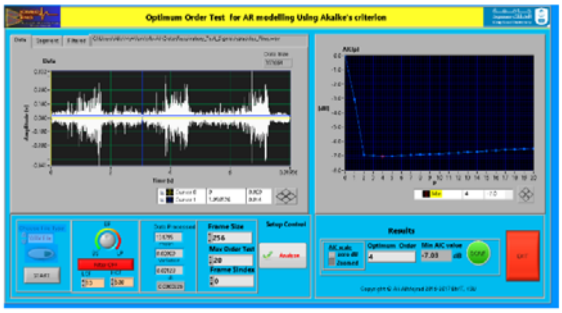
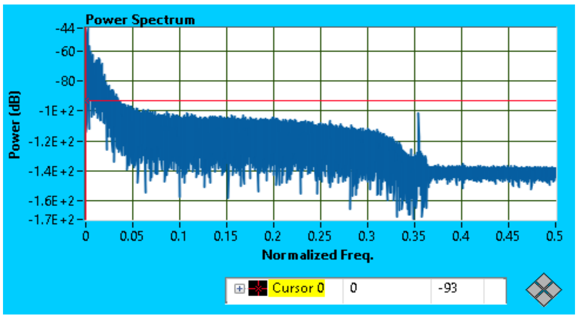
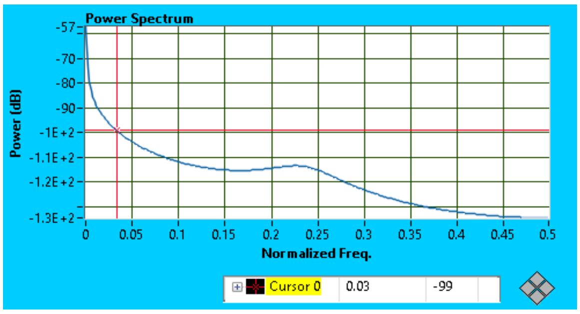


Fig. 7 Model Order Selection using AIC of Crackle Fine Respiratory Signal

From our search, it is found that better results are obtained as the data length increases. The AR order of Respiratory signal found to be 4 as shown from our experiment which is nearly the same as suggested from previous research, which was selected to be around 5 for AR PSD of 256. The results of spectral analysis using FFT and AR obtained using the selected order are shown in Fig. 8. It is clear the best smooth performance using AR method of Fig 8(b) compared to FFT method of Fig 8(a).



(a)



(b)

Fig. 8 Spectral Analysis of a Crackle Fine Respiratory signal using FFT(a) AR using model order of 4 (b)

3.6 Spectrogram analysis

A spectrogram is a visual representation of sound as shown in Fig. 9 which represents Respiratory spectrogram of a crackle fine respiratory signal. The horizontal axis represents time and the vertical axis represents frequency. When it is shown higher on the display means higher sound. The colors of the spectrogram at any point corresponds to relative intensity of the sound at particular time and frequency at that point. Spectrogram provides more information because it is based on actual measurements of the changing frequency content of a sound over time. As shown in the Fig.8 below the color table on the right side that can be changed from many options if desired

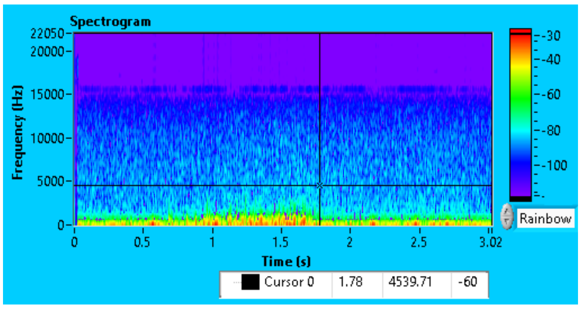


Fig 9. Spectrogram of a Crackle Fine Respiratory signal

3.7 Report Printing

The obtained results can be printed in a report format including Respiratory signal, PSD using the preferred method, spectrogram and respiratory parameters to be in one form to be used by clinician as a reference for diagnosis purpose

4. Conclusions

In this development work, a simple Respiratory measurement device consists of dedicated microphone with signal conditioning and ADC interfacing in addition to advanced digital signal processing software tools are used to develop our virtual Respiratory system on computer which can work remotely. The hardware used was as standard as possible, while the required specific signal processing methods were implemented in software tools and adopted for system objectives. Such composition of components reduced the cost and minimized specialized and expensive hardware. The digital signal processing algorithms were developed to build virtual implementation of instrument that will enable us easily to modernize and upgrade our system by changing simply the software algorithms. Hence, clinician will get the up-to-date system to work with, taking in consideration capital and time resources.

Our main features of our developed system discussed in this paper are flexibility, low cost and future improvement as being capable to be upgraded if required. In addition to that, our developed VRMAS has the following features and capabilities such as ability to work in both modes in real-time acquisition and storing data in different format and displaying mode and off- line analysis mode. In addition to that, it has features and ability to control the whole system of acquisition, play, display, processing and advanced analysis of the different acquired Respiratory signals. Respiratory signals obtained from our system can be analyzed using advanced signal processing techniques in time and frequency domains using spectral estimation methods such as FFT and AR in addition to display of the spectrogram and extraction of respiratory parameters for diagnosis purposes. The obtained results can be printed in a report format including Respiratory signal, power spectral density, and spectrogram for respiratory parameters indices as a reference for diagnosis purpose. The developed system has been tested with real normal and abnormal Respiratory signals and proved to be efficient system while dealing with many different conditions.

Acknowledgment

This work is supported with my great appreciation to the Deanship of Scientific Research and the Research center of College of Applied Medical Sciences at King Saud University**.**

References

[1]https://www.moh.gov.sa/en/HealthAwareness/EducationalContent/Diseases/Chest/Pages/001.aspx.
[2]https://www.pme.gov.sa/ar/Pages/default.aspx
[3] N. Gavriely, et al.," Spectral characteristics of chest wall breath sounds in normal subjects". *Thorax* 50:1292–1300, 1995.
[4] F. Dalmay et al.,"Acoustic properties of the normal chest", Eur. Respir. J., 8, pp. 1781-1789, 1995.
[5] A. R. A. Sovijarvi et al., "Standardization of computerized respiratory sound analysis". *Eur. Respir. Rev.* 10(77):585, 2000.
[6] H. Pasterkamp et al.," Respiratory Sounds. Advances beyond the stethoscope", *Am. J. Respir. Crit. Care Med.* 156:974–987, 1997.
[7] A.R.A. Sovij ¨arvi, et al.," Characteristics of breath sounds and adventitious respiratory sounds". *Eur.Respir. Rev.* 10(77):591–596, 2000.
[8] G. Charbonneau et al.," Basic techniques for respiratory sound analysis", *Eur. Respir. Rev.* 10(77):625–635, 2000.
[9] N. Gavriely et al.," Spectral characteristics of normal breath sounds", *J. Appl. Physiol.Respir. Environ. Exerc. Physiol.* 50(2):307–314, 1981.
[10] S.L. Jae and A. Oppenheim, "Advanced Topics in Signal Processing", Prentice- Hall, New Jersey, pp. 298-300, 1988.
[11] S.M. Kay, "Modern Spectral Estimation: Theory and Application", Prentice Hall, 1988.
[12] L.P. Malmberg et al.," Changes in frequency spectra of breath sound during histamine challenge test in adult asthmatics and healthy controls", Chest, 105, pp. 122-131, 1994.
[13] J.L. Semmlow, et. al, "Noninvasive detection of coronary artery disease using parametric spectral analysis methods", IEEE Eng. Med. Biol. Mag. 9 (1), 33–36, 1990.
[14] G.W. Johnson, "LabVIEW graphical programming: practical applications in instrumentation and control", 2nd edition, McGraw-Hill, New York, 1997.
[15] L.Vannuccini, et al., "Capturing and preprocessing of respiratory sounds", Eur. *Respir. Rev.* 10(77):616–620, 2000.
[16] F. Castanié, "Spectral Analysis: Parametric and Non-Parametric Digital Methods", 1st ed., Wiley, 2006.
[17] H. Akaike, "A new look at statistical model identification." IEEE Trans.AC., vol 19, pp. 716-722, 1974.
[18] J. Vanderschoot et al.,"AR model orders of lung sounds", Proc. Ann Int. Conf. IEEE EMBS, 14, pp. 2531-2532, 1992.

Real-Time Self Localization for Autonomous Robot of RoboCup MSL

Kaori Watanabe
Graduate School of Engineering, Tokyo Polytechnic University, 1583 Iiyama, Atsugi, Kanagawa 243-0297, Japan

Yuehang Ma
Graduate School of Engineering, Tokyo Polytechnic University, 1583 Iiyama, Atsugi, Kanagawa 243-0297, Japan

Tetsuya Yoshida
Graduate School of Engineering, Tokyo Polytechnic University, 1583 Iiyama, Atsugi, Kanagawa 243-0297, Japan

Hidekazu Suzuki
Faculty of Engineering, Tokyo Polytechnic University, 1583 Iiyama, Atsugi, Kanagawa 243-0297, Japan
E-mail: watanabe.kougei.karakuri@gmail.com, xmaancn@gmail.com, t.yoshida1993@gmail.com, hsuzuki@em.t-kougei.ac.jp
www.t-kougei.ac.jp

Abstract

This paper presents a self localization technique using an omni-directional camera for an autonomous soccer robot. The position information of the robot is important for strategic behavior and cooperative operation. Therefore, we have proposed the self localization method which generates the searching space based on a model based matching with white line information of soccer field, and which recognizes the robot position by optimizing the fitness function using Genetic Algorithm.

Keywords: Self-Localization, RoboCup Middle Size League, Soccer Robot, Genetic Algorithm

1. Introduction

The main focus of the RoboCup competitions is the game of football/soccer, where the research goals concern cooperative multi-robot and multi-agent systems in dynamic adversarial environments[1]. In the field of RoboCup, self-localization technique is important to estimate own position including goal and other robot positions and to decide strategy. Basically, we estimate the self-position with the image information, the environment information and the field information. In

Fig. 1. RoboCup MSL

this paper, we describe a real-time self localization method that applies a genetic algorithm (GA) for the RoboCup middle size league (MSL, Fig. 1), which has the widest field size (12x18 m). In Section 2 is hardware structure, which is overview of our past robot, driving module, ball handling module and kicking module. In section 3, vision systems, which are omni-directional camera module, self localization and ball recognition, are described.

2. Hardware structure

We have restructed the most part of hardware structure every year since our team was founded in 2008. We developed the robot based on recent MSL concepts that are high torque driving module, ball handling module, electrical kicking module using solenoid and USB3.0 camera system. We have called the platform as "Mugen"series[2], and "Mugen III(M-III)" shown in Fig. 3, which is improved and based on "Mugen" model, participated in RoboCup Japan Open 2017.

This high torque driving module equips 4-wheels and each wheel is Omni directional wheel. We use ball handling module for rotating a ball with natural direction. This kicking module can shoot a ball by solenoid.

3. Vision system

3.1. *Hardware of vision system*

The omni-directional vision system of our robot is consisted of the camera (FLIR, Flea3[3]), a varifocal lens (Vstone) and a hyperboloidal mirror (Vstone). We developed vision system shown in Fig. 2 for RoboCup MSL robot by combining with above elements. The image captured by this vision system is shown in Fig. 3(a), and the image size and frame rate are 512 × 512 [pixels] and 30 [fps] respectively.

3.2. *Self localization*

We employ a white line of MSL field for self localization. We have proposed the self localization method which generates the searching space based on a model based matching using white line information, and which recognizes the robot position by optimizing the fitness function which has the correct robot position as the maximum value of the function. And this proposed

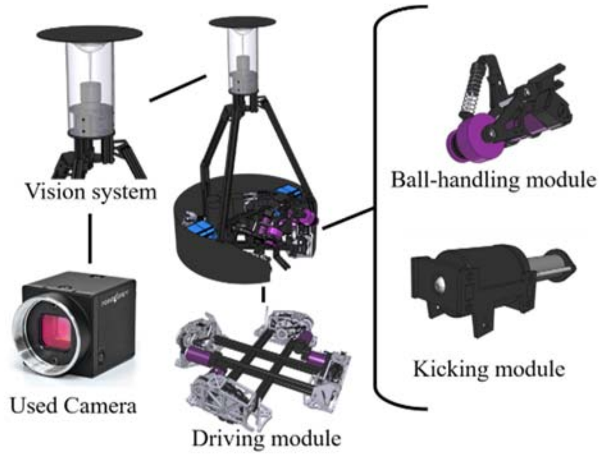


Fig. 2. "Mugen III" Mechanical layout

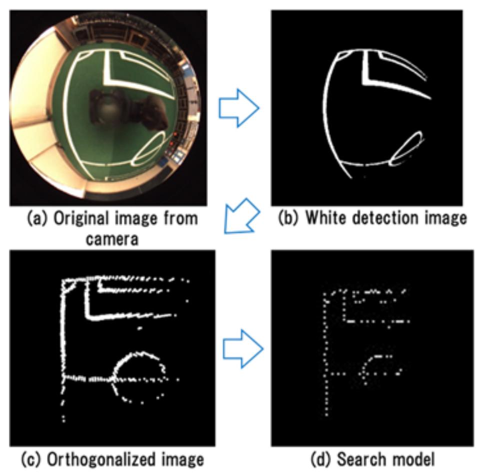


Fig. 3. Process of making search model

self localization method employs Genetic Algorithm (GA)[5] for optimization of the fitness function.

3.1.1. *Searching model*

Figure 3 shows the process of making the searching model of the proposed method. At first, we need the detection image of the white line for making the searching model based on the white line. We obtain the white detection image by employing the converting method of color space from RGB to HSV and to YUV like Fig. 3(b). And we generate the field information by orthogonalizing the white line information like Fig. 3(c). Moreover, we determine the searching model by thinning down the field information based on white line like Fig. 3(d). Therefor we use thinned model as the self localization as model.

3.2.2. *Model-based matching*

The proposed self localization method generates the searching space by model based matching between geometric information of the white line in the MSL field and above-mentioned searching model. Let us denote the set of pixels, which compose of the searching model shown in Fig. 4, as $S_f(\tilde{\phi})$. The position $\tilde{r} = (\tilde{x}, \tilde{y})$ and orientation $\tilde{\theta}$ of searching model in the image is represented as $\tilde{\phi} = [\tilde{x}, \tilde{y}, \tilde{\theta}]^T$. Then S_f movement in the matching area is expressed as $S_f(\tilde{\phi})$. And, if the pixel value of field image corresponding to the area of the moving model is expressed as $p(\tilde{r}), \tilde{r} \in S_f(\tilde{\phi})$, then the evaluation function $F(\tilde{\phi})$ of the moving model is given as follows.

$$\boldsymbol{F}(\tilde{\boldsymbol{\phi}}) = \sum_{\tilde{r} \in \boldsymbol{S}_f(\tilde{\boldsymbol{\phi}})} \boldsymbol{p}(\tilde{\boldsymbol{r}}) \tag{1}$$

The fitness function $F(\tilde{\phi})$ obtains the maximum value when the position of the searching model corresponds to the correct position that robot exist in the MSL field. Then, the problem of detection of robot position and orientation is converted to a searching problem of $\tilde{\phi}$ such that $F(\tilde{\phi})$ is maximized[6]. The calculation result of whole matching area shown in Fig. 4 shows Fig. 5. In this Fig. 5, the vertical axis represents the fitness value, and the horizontal axes represent the field plane. Here, we select only one depending on the value of an electric compass, because two maximum value exist in the function value caused by revolution symmetry of geometric characters of the MSL field.

3.2.3. *Genetic Algorithm*

In the proposed self localization method, we employ Genetic Algorithm (GA) for searching the maximum value of the fitness function $F(\tilde{\phi})$. A GA is an example of an artificial intelligence program and is well known as a parallel search and optimization process that mimics natural selection and evolution(Fig. 6). In the proposed method, an elitist model of a GA that preserves the best individual in the population at every generation is utilized and genetic coding using gray code, roulette selection and one-point crossover are employed. And, the

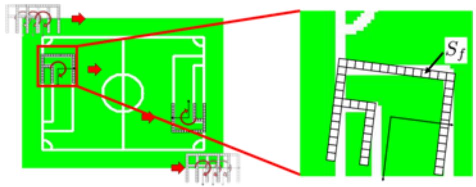


Fig. 4. Model matching

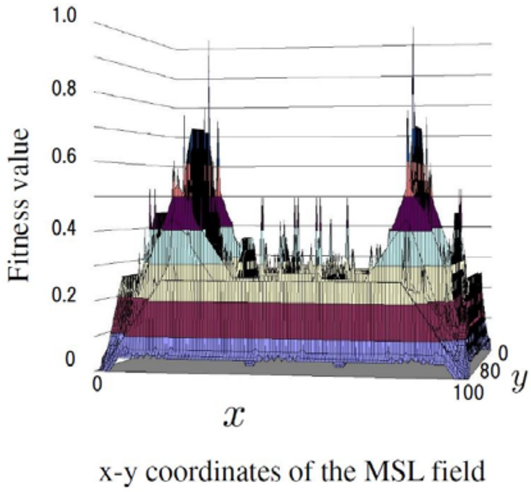


Fig. 5. Calculation result of the fitness function

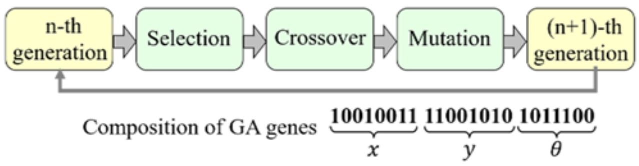


Fig. 6. GS's process

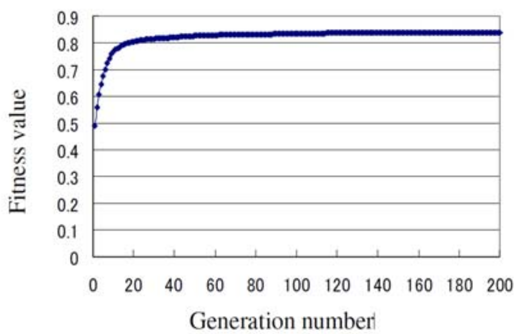


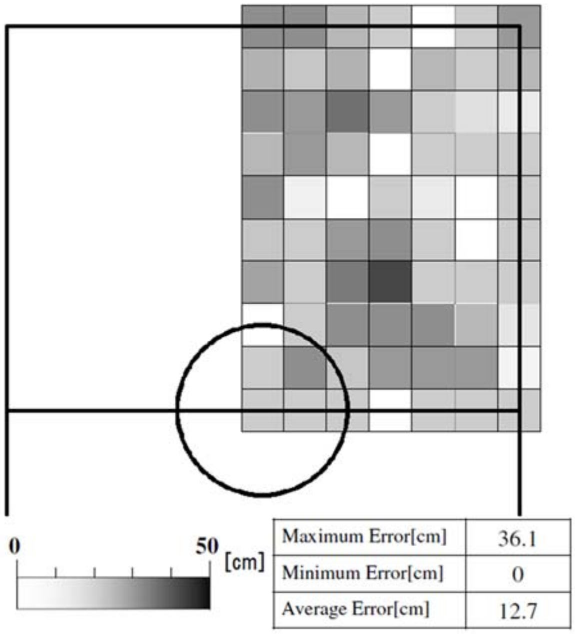
Fig. 7. Result of the convergence of GA

parameters of the GA process are determined by previous experiments. Figure 7 shows the result of the convergence of GA in case of self localization experiment using actual image that the robot captured at voluntary position. In this figure, the vertical axis represents fitness value of fitness function, and the horizontal axis represents the generation number of GA. The GA converged the maximum value, which means current position of robot in the MSL field, at about 60th generation in real time.

3.2.4. *Verification experiment*

We performed the self localization experiment to verify the effectiveness of the proposed method. Figure 8 shows the result of the verification experiment that checked the self localization error between correct position and detected position at the quarter area of the MSL field at interval of one meter. In this figure, each box represents the error as the brightness of gray scale. Average error of this experiment was 12.7[cm], and the accuracy of the self localization by the proposed method is enough to play soccer.

Moreover, we performed the experiment using five player shown in Fig. 9 on the assumption of MSL game. Laptop PC displays the position sended by each robot in real time, and the positions of five player described in the result image corresponded with the actual position of the robots.



Maximum Error[cm]	36.1
Minimum Error[cm]	0
Average Error[cm]	12.7

Fig. 8. Error of the self localization

Fig. 9. Verification experiment in the real world

4. Conclusion

In this paper, we have proposed the self localization method which generates the searching space based on a model based matching with white line information of RoboCup MSL soccer field, and which recognizes the robot position by optimizing the fitness function using Genetic Algorithm. Moreover, we verified the effectiveness of the proposed self localization method using GA. Furthermore, we confirmed that the accuracy of the self localization by the proposed method is enough to play soccer.

References

1. “RoboCup Federation official website,”, https://www.robocup.org/
2. 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)
3. FLIR, “Flea3(USB3 Vision Camera)”, https://www.ptgrey.com/flea3-usb3-vision-cameras
4. T. Yoshida, H. Suzuki, Real-time self-localization for RoboCup Middle-Size-League, Proceedings 32nd Fuzzy System Symposium, pp. 397-398, 2016. (In Japanese)
5. D.E.Goldberg, “Genetic Algorithms in Search, Optimization and Machine Learning”, Addison-Wesley, 1989.
6. 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.

Behavior Selection System for Robot Using Neural Network

Moeko Tominaga
Graduate School of Life Science Systems Engineering, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka, 808-0196, Japan

Yasunori Takemura
Faculty of Engineering, NishiNippon Institute of Technology
1-11, Aratsu, Kanda-town, Miyako-gun, Fukuoka, 800-0397, Japan

Kazuo Ishii
Dept. of Human Intelligent Systems, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka, Japan
E-mail: tominaga.moeko382@mail.kyutech.jp, takemura@nishitech.ac.jp, ishii@brain.kyutech.ac.jp
www.kyutech.ac.jp

Abstract

With the progress of technology, the realization of a symbiotic society with human beings and robots sharing the same environment has become an important subject. An example of this kind of systems is soccer game. Soccer is a multi-agent game that requires strategies by taking into account each member's position and actions. In this paper, we discuss the results of the development of a learning system that uses SOM to select behaviors depending on the situation.

Keywords: Strategy, Self-Organizing Map, team behavior, Tensor SOM, multi-agent system, Human-Robot cooperation.

1. Introduction

Recently, the implementation of robots in society has become a possible solution to many problems, such as ensuring the safety of a sustainable society, responding to a rapid population aging and population decrease. Moreover, robots will represent the foundation of the future industry. To properly implement a robotized society, it is necessary to conduct and to present research outcomes in a manner that is easy to understand, avoiding differences between social expectations and the direction of research and development. Therefore, it is essential to discuss how to achieve coexistence with robots and what a symbiotic society should look like. In such a society, humans and robots interact with each other and they are capable of mutual understanding, so not only of their own actions, but also with all the other agents, where agent means each of the active subject involved. Aim is to develop a suitable algorithm which allows to create intelligent robots able to share the environments with humans. Since soccer involves strategies, cooperation, unpredictable movements and common targets, it represents a good test bed for developing such algorithm.

Tensor Self Organizing Map (Tensor SOM)[1] is used for this scope.

1.1. Cooperative Behavior

Cooperative behavior becomes a crucial aspect when different autonomous agents interact while performing a common task. Often a single agent is not much effective in accomplishing a task, and in the last years many researchers have been studying multi agent systems (MAS) to solve difficult problems. They interact to each other and with the environment by taking real time decisions based on the data acquired from the sensors.[2, 3] As a test bed of MAS, RoboCup encourages the cooperation of multi-agents using learning methods, such as reinforcement learning and neural networks. RoboCup is a project aimed to win the soccer World Cup against humans[4]. According to Sandholm and Crites[5], reinforcement learning can be used successfully for the iterated prisoner's dilemma, if sufficient measurements data and actions are available. In addition, Arai[6] compared the Q-learning and Profit Sharing methods about the Pursuit Problem in a multi-agent system, when the environment is modelled as a grid, and showed that cooperative behaviors emerge clearly among Profit Sharing. However, these studies have not yet considered applications for robots that operate in a real environment.

As mentioned above, a humans-robots cooperative society is a topic of interest for many researchers and several works have been carried out. So, in such a symbiotic system, it is mandatory for the robots to understand and interpret humans' behavior and act accordingly. However, the target in research is often about the robots' behavior only, or a behavior based on an interaction defined precisely between a specific number of humans and robots. Also, in a multi-agent system the communication between humans and robots is usually carried out by means of some kind of interface, for example command voice or gestures. But this might not be enough, especially when the system is very complex. Accordingly, the robot should be able to understand some situations and adapt its behavior in a predictive manner, just as humans do. The main goal of this research is to consider cooperative behaviors between humans and apply them to robots (see Fig. 1).

The input vector elements for a neural network is chosen so as to achieve the lowest possible gap between humans and robots' actions. To obtain this, it is important at first to understand humans' behavior and develop robots able to understand such behaviors and imitate them. This paper focuses on the study of the humans' and robots' behavior in a futsal game, because this represents a good test bed, with a dynamic environment, several constrains and it requires a real-time planning. These are the characteristics of a common symbiotic system in which the robots may operate in the future. In analyzing a soccer futsal game, at first the elements to be observed are decided, depending on their importance and ease of evaluation. Many valuable info can be obtained from the simulation game: the position coordinates and the velocity of the players and the ball, the elapsed time of the game and the score. In the analysis performed in this work with status of players such as the positions coordinates and the players moving speeds and ball were considered. The Input data was used to train the neural network. In analyzing a soccer futsal game, at first the elements to be observed are decided, depending on their importance and ease of evaluation. Many valuable info can be obtained from the simulation game: the position coordinates and the velocity of the players and the ball, the elapsed time of the game and the score.

In the analysis performed in this work with status of players such as the positions coordinates and the players moving speeds and ball were considered. The Input data was used to train the neural network. The simulator was made by RoboCup project and which also consider about the exercise model of the robots and the ball. The futsal

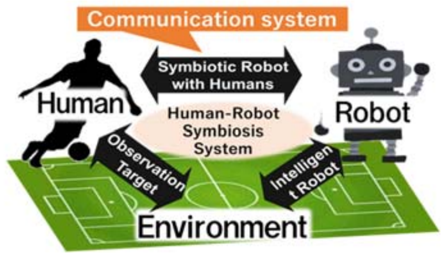


Fig. 1. A human–robot symbiotic system.

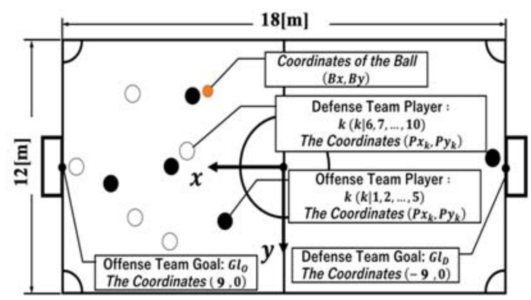


Fig. 2. The field of the game.

game was conducted with five versus five robots, and the coordinates of all players and the ball were observed (Fig. 2). The state vector is expressed by Eq. (1). T he player position coordinate $\boldsymbol{P_k}$ and the ball position coordinate $\boldsymbol{B}$ in the field are shown in Eqs. (2) and (3).

$$\boldsymbol{A} = [\boldsymbol{P_1}, \boldsymbol{P_2}, \cdots, \boldsymbol{P_{10}}, \boldsymbol{B}\ ] \tag{1}$$

$$\boldsymbol{P_k} = [Pk_x, \quad Pk_y] \tag{2}$$

k indicates the ID number identifying a player. We observed a game consisting of a team of $k = 1, 2, 3, 4, 5$ players and a team consisting of $k = 6, 7, 8, 9, 10$ players. By chance, $k = 5, 10$ are the goalkeepers of the two teams respectively.

2. Tensor Self-Organizing Map

Self-organizing map (SOM) is an unsupervised learning technique and is known as competitive learning, similar to information processing via neural circuits. Competitive learning is an important concept in hierarchical neural networks. Each input neuron in the input layer (input space) is connected to all output layer neurons, the strength of the relation between two neurons in consecutive layers is decided by a specific weight assigned to each connection. The input vectors "compete" with each other so as to find the best output vector. With best output, it is meant the output vector whose elements are as similar as possible to the input vector elements. To establish this similarity the Euclidean distance is used. By doing this, the clusters into the output map are obtained. The cooperative hierarchy deals with the output vectors. While performing the algorithm, the elements value of each vector is adjusted. If a high change of an element occurs in a vector, the vectors around it change accordingly, but in a way that the strength of this change decreases with the distance in the map from the first vector. In particular, the amount of change of each element is decided by the Gaussian function. Since each vector modifies itself depending on the surrounding vectors, this behavior is said cooperative. Input vectors can have high dimension but the output results are showed in a two-dimensional map. So, with this structure multi-dimensional input vectors can have reduced dimensions in the output layer, and the features are clustered, so as to have similar features in the same area of the map. An applicative example for which SOM is suitable is the visualization of complex scenarios, since the output results are shown in a two-dimensional form. Furthermore, the cooperative hierarchy and the smoothing process performed by the Gaussian function allow to synthesize elements and contents that were not clearly specified in the input vectors. SOM has a high interpolation performance and it can generalize concepts starting from a limited amount of input data. Because of these characteristics, it is well suited for unsupervised learning and highly valid outputs are expected. In this paper, in order to cope with high-dimensional data, analysis is performed using tensor SOM which is SOM expanded to tensor.

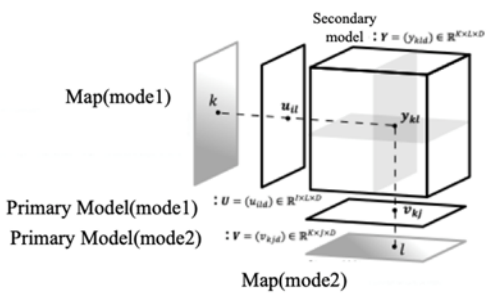


Fig. 3. Overview of the Tensor SOM algorithm.

2.1. Algorithm

The tensor SOM algorithm proceeds according to the following steps (see Fig. 3):

2.1.1. Choose Best Matching Unit(BMU)

The selection of the best matching unit (BMU) k_i^* , j_j^* are carried out for each input vector u_{ild} , v_{kjd} with the following equation, in which the Euclidean distance is calculated.

$$k_i^* = \arg\min_k \sum_{l=1}^{L} \sum_{d=1}^{D} (u_{ild} - y_{kld})^2 \tag{4}$$

$$l_j^* = \arg\min_l \sum_{k=1}^{K} \sum_{d=1}^{D} (v_{kjd} - y_{kld})^2 \tag{5}$$

2.1.2. Calculation neighbor radius

$$\alpha_{k_i} = exp\left[-\frac{1}{2\sigma^2}\left\|\zeta^{(1)}_{k_i^*} - \zeta^{(1)}_{k}\right\|^2\right] \tag{6}$$

$$\beta_{l_j} = exp\left[-\frac{1}{2\sigma^2}\left\|\zeta_{l_j^*}^{(2)} - \zeta_l^{(2)}\right\|^2\right] \tag{7}$$

2.1.3. Update secondly model

$$y_{kld} = \frac{1}{g_k g_l}\sum_{i=1}^{I}\sum_{j=1}^{J}\alpha_{k_i}\beta_{l_j}x_{ijd} \tag{8}$$

$$g_k = \sum_{i=1}^{I}\alpha_{k_i} \tag{9}$$

$$g_l = \sum_{j=1}^{J}\beta_{l_j} \tag{10}$$

2.1.4. Update primary model

$$u_{ild} = \frac{1}{g_l}\sum_{j=1}^{J}\beta_{l_j}x_{ijd} \tag{11}$$

$$v_{kjd} = \frac{1}{g_k}\sum_{i=1}^{I}\alpha_{k_i}x_{ijd} \tag{12}$$

3. Experiments

The experiment was conducted for a total of three experiments, and the parameters shown in Table. 1. The input data consists of the player's state (position, speed, etc.), player ID, and the action at that time (see Fig. 4).

Table. 1. Parameter settings for the experiment.

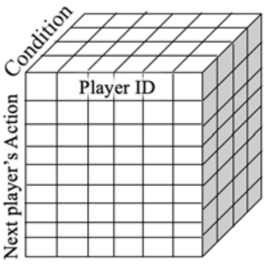


Fig. 4. Image of input data.

Number of Iterations	n	10
Map Size		100 × 100
Max Neighboring Radius	σ_{MAX}	2.0
Minimum Neighboring Radius	σ_{min}	0.2

3.1. Experimental Results and Discussion

The results are analyzed by means of the Unified Distance Matrix (U-Matrix)[7] and the Component Plane Matrix (see Fig. 5). First one aims to find the output vectors with a common b o s p w b

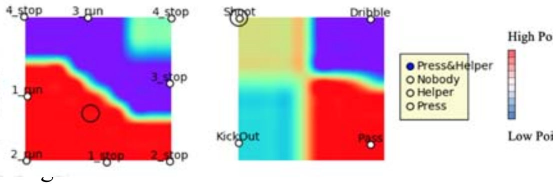


Fig. 5. Result of Tensor SOM expressed by the U-Matrix after learning.

4. Conclusions

In this study, Action selection system was developed using Tensor SOM. In the future, further improvements of the input vector are required to obtain some more detailed clusters. This work showed the promising ability of Tensor SOM algorithm to develop a symbiotic system where each agent is able to understand the scenario in which it is operating and work in a predictive manner together with the other agents of the system.

References

1. T. Iwasaki, T. Furukawa, "Tensor SOM and tensor GTM: Nonlinear tensor analysis by topographic mappings", Neural Networks, Volume 77, May 2016, pp. 107-125.
2. P. Stone and M. Veloso: "Multiagent Systems: A Survey from a Machine Learning Perspective", Autonomous Robots, Vol. 8, No. 3, pp. 345-383, 2000.
3. K. Ueda: "Souhatsu to multi agent system [Emergence and multi agent system] ", Baifukan, 2007.
4. Behnke, S., Sheh, R., Sariel, S., Lee, D.D. (Eds.): RoboCup-2016: Robot World Cup XX, Springer-Verlag (2017)
5. T. W. Sandholm, R. H. Crites: On Multiagent Q- learning in a Semi-Competitive Domain, Workshop Notes of Adaptation and Learning in Multiagent Sys-tems Workshop, IJCAI-95, 1995.
6. S. Arai, K. Miyazaki, S. Kobayashi: Methodology in Multi-Agent Reinforcement Learning : Approaches by Q-Learning and Profit Sharing, Journal of Japanese Society for Artificial Intelligence, 13(4), pp.609-618, 1998.
7. Ultsch, A. and Siemon, H.P.: Kohonen's Self Organizing Feature Maps for Exploratory Data Analysis, INNC'90, pp.305–308 (1990).

Development of the auto measurement system for cedars in a forest using a drone

Keiji Kamei
Department of Production Systems, Graduate School of Engineering, Nishinippon Institute of Technology
1-11 Aratsu, Kanda, Miyakogun, Fukuoka 800-0394, Japan

Masahiro Kaneoka, Ken Yanai, Masaya Umemoto, Hiroki Yamaguchi, Kazuki Osawa
Department of Information Systems, School of Engineering, Nishinippon Institute of Technology
1-11 Aratsu, Kanda, Miyakogun, Fukuoka 800-0394, Japan
E-mail: kamei@nishitech.ac.jp, s162410@nishitech.ac.jp, s162436@nishitech.ac.jp, s162406@nishitech.ac.jp, s162437@nishitech.ac.jp, s162408@nishitech.ac.jp
www.nishitech.ac.jp/~kamei/

Abstract

Drones have been used in many purposes for a long time. Especially, development of the automatic observation systems such as measurement using drones for the primary sector of industry have been frequently researched in recent. The measurement of a tree growth in a forest is also one of the aim for a drone application. In this study, our aim is to develop the automatic measurement system for size of a tree in a forest. The difficulties are that a drone has to recognize trees, to construct a map of a forest and to measure the size of trees from a front camera. To overcome those difficulties, we propose that a drone recognizes trees based on SSD, constructs a map from SLAM and measures a tree by SfM. Experimental results from the drone competition show that a drone has been able to recognize a tree and to fly safety.

Keywords: Deep Learning, Region Convolutional Neural Networks, Structure from Motion, Simultaneously Localization and Mapping, Recognition of trees, Drone in a forest

1. Introduction

Forests occupy about 70% of Japanese whole land area, and artificial forests are 30% in Japanese forests. Maintenance and management of forests is important because rundown forests cause of disasters; however, grooming chores of forests are considered to become difficult in near future due to acceleration of demographic aging in Japan. Hence, forestry companies and its workers strongly demand the autonomous maintenance and management system[1].

The features of Japanese forests are that trees grow on mountains, so that there are many rocks and ground inclines in a forest. In addition, ground in a forest usually covered by underbrush. For this, a robot with wheels or crawlers is difficult to apply for maintenance and management system. To solve this difficulty, we propose to apply a drone to developing autonomous maintenance and management system in a forest. The abilities of a drone have been made impressive progress in recent years, so that drones are able to fly stably in any situations. On the other hand, drones have several weak points if those fly in a forest; therefore, drones are radio-controlled by someone in general. Most of drones use Global Positioning System(GPS) for position estimation and measure height between a drone and ground from altimeters or ultrasonic sensors. The signal of GPS may be blocked by trees or be altered by reflected signal(multipath problem); therefore, receiving accurate GPS signal for construction of a map and estimating positions is difficult in a forest. In addition to problem of flight control, a problem of development for autonomous maintenance and management system is that a drone autonomously has to recognize trees, to construct map without GPS signal and to measure the size of a tree. We have proposed a method to recognize a color marker of target tree from a camera on a drone based on particle filter algorithm[2]. Our proposal have succeeded in recognition for the certain tree; however, our proposal was not able to construct map and to recognize general trees without color marker. Anzai et. al have proposed to combine a Laser Range Finder(LRF) with a camera on a drone[3]. Their proposal chooses candidates of trees and measures of that based on the information from LRF, then recognizes a color marker of target trees from a camera. In addition to recognition of tree, their proposal constructed simultaneously a map of a forest. The problem of their approach is that recognition method needs color markers; namely, a drone does not to recognize category of tree. In this paper, we propose to apply "Object Category Recognition" based on Convolutional Neural Network(CNN) to recognition of trees, "Structure Recognition" based on Structure from Motion(SfM) to measurement of size of a tree and "Map Construction" based on Simultaneous Localization and Mapping(SLAM) to construction of forest map. This paper is constituted as follows. The section 2, 3, 4 show the details of CNN, SfM and SLAM, respectively. The section 5 shows the experiments and its results. Finally, the section 6 concludes this paper.

2. Convolutional Neural Network

Convolutional Neural Network(CNN) is the one of Deep Learning Algorithm[4]. It can be said that the success of CNN becomes a springboard for attracting attention for multi(deep) layered neural network. A notable success of CNN is that CNN is able to classify with high rate probability object category from an image. We describe the details of CNN in first, then describe the details of Region CNN for recognition of trees.

2.1. *The basic CNN*

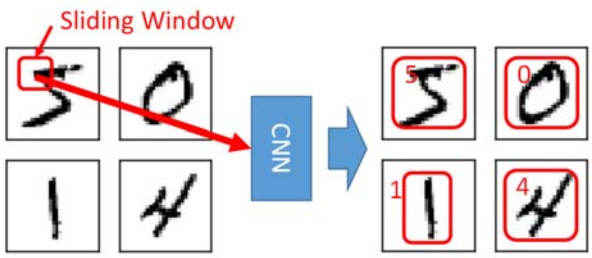
The structure of the basic CNN is that CNN has 2 types of layer, i.e., convolutional and pooling layer and those layers are used in pair. Convolutional layers learns the feature of small area in images, and pooling layers integrate outputs from convolutional layers. From the structure, the basic CNN learns the category of object in an image. For example, when there are dog or cat images, CNN classify to 2 classes those images. To realize our aim, the classifier has to recognize where target objects are in an image; hence, the basic CNN is not able to be applied for our aim.

2.2. *The Region CNN(R-CNN) and SSD*

In contrast to the basic CNN, an approach for recognition of each object in an image based on CNN is said the Region CNN(R-CNN)[5]. Fig.1 illustrates the structure of a typical R-CNN.

Fig.1 The structure of a typical R-CNN

The algorithm of R-CNN split an image to small region by using sliding window as shown in Fig.1. The size and position of sliding window is variable, so many small regions are generated. A CNN in R-CNN learns the feature of small regions in an image for classification of objects, and so R-CNN is able to recognize several types



of objects in an image. A typical R-CNN is known to indicate the very good performance; on the other hand, it needs the very high computational capability due to huge number of input data(small regions). Several approaches for cost down of R-CNN have been proposed, e.g., Fast R-CNN. We apply Single Shot MultiBox Detector(SSD)[5] to recognition of trees in this research because SSD is low computational cost. The characteristic of SSD is that an image is split into fixed size of grids alternative to using sliding window, and SSD learns the features of objects in the certain grids. For this, computational cost is drastic decrease. However, recognition rate of SSD is lower than other R-CNN methods about 10% according to paper 5, so we will validate whether the enough ability of tree recognition or not in the experiment.

3. Structure from Motion

Measurement of size of trees is important for autonomous maintenance and management system in a forest. We propose to apply Structure from Motion(SfM)[6] to generation of 3D tree models for measurement of the size of trees. The measurement algorithm is that it compares a criterial tree with target tree based on generated 3D tree models. SfM algorithm is as below. Firstly, it discovers corresponding points from a series of images based on Optical Flow algorithm. Secondly, sets of corresponding points are merged to a set of corresponding points, and point group data is generated from those points. Finally, a 3D image model is generated by pasting a series of images to point group data. Note that point group data correspond 3D models, and so measurement of size of trees does not need 3D image models.

4. Simultaneous Localization and Mapping

A drone must have a map of forest in advance to measure the size of trees. However, it is difficult to prepare the exact map of a forest in advance because states of a forest vary depending on seasons. For this, construction of exact map for a forest is also important for autonomous maintenance and management system. We propose to apply Simultaneous Localization and Mapping(SLAM)[7] for construction of exact map. SLAM is a method to estimate positions(localization) and to construct map(mapping) simultaneously by only using sensory information and robot control signals. Several types of SLAM have been proposed, we apply Visual SLAM to construction of map. Visual SLAM constructs a map and estimates position based on a series of images, so it is said that Visual SLAM is as same as SfM on generating of point group data; namely, SfM uses point group data for construction of 3D image models, and Visual SLAM uses that for localization and mapping. There are several methods for Visual SLAM, we utilize ORB-SLAM[8] because ORB-SLAM shows good performance for mapping and localization.

5. Experiments

We conduct 2 kinds of experiments, i.e., in laboratory and in"The 4th Forestry Drones & Robots Competition[1]." This section shows the details of a drone, then experimental results. From the competition, our target tree is cedar, and so our final goal is construction of the autonomous cedar maintenance and management system.

5.1. *Drone: Tello*

In this research, we use "Tello[9]" drone. Tello has a front camera, which is 5[M]Pixel and 82.6[deg]FOV. The weight of Tello is about 80[g], so Tello is very light drone. Our system controls Tello from a laptop PC via WiFi.

5.2. *Experiment 1: In Laboratory*

Firstly, we show the results of R-CNN. In this study, we use Google Object Detection API[10] for R-CNN to accelerate of development. In R-CNN experiment, we took a video of cedars in advance in a forest where is near our university. We use that video to training of R-CNN and its evaluation. Fig.2 shows the recognition result.

Fig.2 Recognition result of a cedar.

From the result, R-CNN found a cedar in region enclosed by green rectangle; hence, R-CNN succeeded in recognition a cedar although a part of cedar is hidden by other trees. For this reason, it can be said that SSD algorithm has enough ability for recognition and finding cedars. Secondly, we show the result of SfM. 3D image model of SfM was constructed in real time processing with flying Tello. In this study, we use OpenMVG, OpenMVS and Meshlab[11]. The aim of SfM in here is to construct 3D image model of a pillar of our university as Fig.3(a), and Fig.3(b), (c) show the results of SfM.

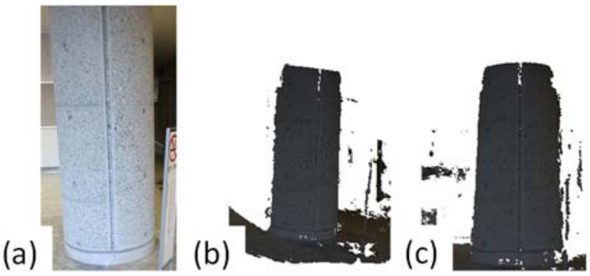


Fig.3 Result of SfM. (a)target pillar, (b),(c)3D image model

From the Fig.3(b) and (c), we were able to generate 3D images instead of planar image, so that we succeeded in construction for 3D image model by using SfM.

Finally, we show the result of SLAM. We use open library of ORB-SLAM[8] as same as R-CNN and SfM to construct map of the hall. Fig.4(a) and (b) show detection of corresponding points in an image and results of construction the hall map, respectively.

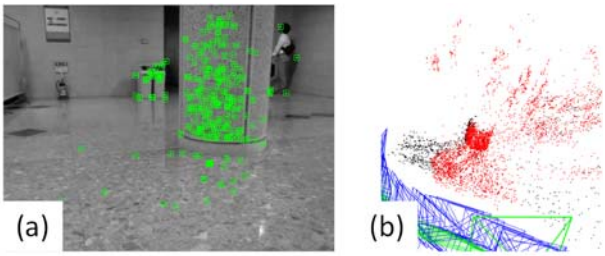


Fig.4 Result of SLAM. (a)detection of corresponding points, (b)generated hole map

Results of SLAM show that corresponding points were exactly detected in a series of images from Fig. 4(a), and a map surrounding a pillar in hall was correctly generated from Fig. 4(b); therefore, we succeeded in construction of a map based on SLAM.

5.3. *Experiment 2: Competition*

We integrated R-CNN, SfM and SLAM in order to develop the autonomous maintenance and management system; then, we participated in drone competition to evaluate our system. Fig.5 shows scene of flight of a drone with our system. From the Fig. 5, a drone succeeded in going to near a criterial tree because R-CNN and SfM module correctly operated; however, a drone were not able to construct a forest map due to problem of SLAM module. As a result of drone competition, it can be said that we succeeded in development a part of autonomous maintenance and management system.

Fig.5 The scene of flight of a drone(In competition)

6. Conclusions

In this study, our aim is to develop the automatic measurement system for size of a tree in a forest. The difficulties are that a drone has to recognize trees, to construct a map of a forest and to measure the size of trees from a front camera. To overcome those difficulties, we propose that a drone recognizes trees based on R-CNN, constructs a map from SLAM and measures a tree by SfM. As a result, it can be said that we succeeded in development a part of autonomous maintenance and management system. Precise integration of R-CNN, SfM and SLAM is left for further study.

References

1. Center for Socio-Robotic Synthesis, Forestry Drones & Robots Competition (in Japanese), 2019. www.lsse.kyutech.ac.jp/~sociorobo/ja/forestdrone2019
2. K. Kamei et. al,Adaptive Color Tracking for Particle Filter to Control a Drone in a Forest, in *Proc. SCIS&ISIS2016*, 2016, pp.512-517
3. T. Anzai et. al,Sensing and Navigation of Aerial Robot for Measuring Tree Location and Size in Forest Environment, in *Proc. SISA2017*, 2017, pp.55-60.
4. A. Krizhevsky et. al, ImageNet Classification with Deep Convolutional Neural Networks, *Advances in Neural Information Processing Systems,* 25, 2012, pp. 1097-1105
5. L. Wei et. al, SSD: Single Shot MultiBox Detector, Lecture Notes in Computer Science, Springer, 2016, pp.21-37
6. F. Bethmann et. al, Object-based Multi-Image Semi-Global Matching - Concept and first results, in Proc. *ISPRS2014*, 2014, pp.93-100
7. S. Thrun et. al, Probabilistic Robotics, MIT Press, 2006
8. ORB-SLAM Proj., webdiis.unizar.es/~raulmur/orbslam/
9. Ryze Tech: Tello, www.ryzerobotics.com/jp/tello
10. Google Object Detection API, github.com/tensorflow/models/tree/master/research/object_detection
11. OpenMVG, OpenMVS, Meshlab, github.com/openMVG/openMVG/, github.com/cdcseacave/openMVS/, www.meshlab.net/

Report on the 5th Tomato-harvesting Robot Competition

Yasunori Takemura
Department of Engineering, Nishinippon Institute of Technology, 1-11 Aratsu, Kanda, Miyako, 800-03945, Japan

Takemyuki Matsuo
National Institute Technology, Kitakyushu College
*2-20-1, Shii, Kokuraminami-ku, Kitakyushu-shi, Fukuoka, Japan**

Takashi Sonoda
Department of Engineering, Nishinippon Instituite of Technology
1-11Aratsu, Kanda, Miyako, Fukuoka, Japan

Kazuo Ishii
Kyushu Institute of Technology
2-4, Hibikino, Wakamatsu-ku, Kitakyushu city, Fukuoka, Japan

E-mail: takemura@nishitech.ac.jp, matsuo@kct.ac.jp, sonoda@nishitech.ac.jp, ishii@brain.kyutech.ac.jp
http://www.lsse.kyutech.ac.jp/~sociorobo/ja/tomato-robot2018

Abstract

Tomato is one of the important fruit vegetables and most tomatoes are produced in the greenhouses, or large-scale farms, where the high temperature and humidity, and long harvest age force the farmer heavy works. To develop the tomato harvesting robot, many research issues exist such as manipulator design, end-effector design, collaborative behavior, artificial intelligence, motor control, image processing, target recognition and so on. With an aim to promote the automation of tomato harvesting, we have organized the tomato harvesting robot competition since 2014. In this paper, we report on the results of 5th tomato harvesting robot competition in 2018.

Keywords: robots for socio synthesis, robot competition and agriculture robot

1. Introduction

Recently, the aging and depopulation of famers grow worse in Japanese agriculture, as results the shortages of future famers and manpower become big problems. Ministry of Agriculture, Forestry and Fisheries of Japan reported that Japanese self-sufficiency ratio for food is about 40 percent, which is lowest level among developed countries. As one of solutions for the problems, the implementation of robot technology into the agriculture is expected.

Most of commercialized robots are for industry and robots for agriculture, forestry and fisheries are under developing, however, not commercialized yet. The reasons for the delay are cost-efficiency of the robotization, safety of the works using robots, difficulty of outdoor operations, and knowledge transfer problem from farmers to computer, etc. If we can overcome these difficulties and implement the robots into agricultural fields, robots can contribute to the laborsaving, improvement of production, production line automation. Also, the management of agricultural products such as quality, quantity, and condition of environment become possible and the smart-agriculture will be realized. Tomato is one of important fruit vegetables and most tomatoes are produced in the greenhouses, or large-scale farms, where the high temperature and humidity, and long harvest age force the farmer heavy works.

To develop the tomato harvesting robot, many research issues exist such as manipulator design, end-effector design, collaborative behavior, artificial intelligence, motor control, image processing, target recognition and so on. With an aim to promote the automation of tomato harvesting, we have organized the tomato harvesting robot competition since 2014[1,2].

Several research results have been published though tomato harvesting competitions. For example, in mechanics designs and system designs, Yamaguchi and et., al., proposed about rotational plucking gripper for harvesting [3], B.Li and et., al. also proposed end effector design for harvesting[4], and Fujinaga and et., al. proposed system design about Tomato harvesting robot and software[5].

In this way, many new agricultural robot design methods and recognition algorithms have been born. One of the major goals of this tournament is to create new technological evolution and knowledge sharing through the tournament.

In this paper, we report on the results of 5th tomato harvesting robot competition in 2018.

2. Competition regulations

The tomato harvesting robot competition consists of two leagues which are he Senior League and the Junior League. The target competitors for Senior League are supposed to the team with automated and remote-controlled robots, and the Junior League are for high school or junior high school students who build robot using LEGO Mindstorm. In this section, competition regulations and new rules of both reargues will be introduced.

2.1. *Senior League*

The Senior League supposes that teams are not restricted to university nor company, just each team should have own tomato harvesting robot(s). Two kind of competition field are designed, the one is rail-style area and the another one is fee-style area as shown in Fig.1 and Fig.2. The rail-style area is designed to have the similar environment with the tomato factory with heat pipes for warming greenhouse. Free-style area is for the robots of open-field culture environment.

Fig.1 Free-style area[2]

Fig.2 Rail-style area[2]

Fig.3 1st Stage [2]

Fig.4 2nd Stage [2]

Fig. 5 Final Stage [2]

Tomatoes are hanged on height from 800[mm] to 1200 [mm] referring to tomatoes arrangement of Hibikinada Saen Co., Ltd. Regarding the sizes of robot, the projected area of the robot to the ground is within the 800 × 800 mm^2 and no height limitation. For safety reason, the robots should have an emergency stop switch on the easy-to-find position. As the recommendations, the weight of the robot is <50 kg and the electric power of motor is <70 W/each.

The first stage is intended as the inspection of basic functions needed for tomato harvesting, so that single tomatoes are suspended as shown in Fig.3. The team succeeded to touch the tomato, moves to the second stage. In second stage, some sets of bunches of tomatoes are suspended as shown in Fig.4. The five high score teams go up to the final stage. The score is calculated based on the number of successfully harvested tomatoes and damages to tomatoes. Also, the unripe tomatoes are counted as the damaged tomatoes. Final score of second stage is the score multiplied the basic score and the coefficient decided by a combination of choice between the control method which are remotely and autonomous control and the areas which are the rail area and the free style area. In final stage, the robots harvest tomato from plant body as shown in Fig.5.

Table 1 The class number and coefficient at choosing remote control

Method of View	Directly		Indirectly	
Area	Rail	Free	Rail	Free
Number of Category	T1	T2	T3	T4
Coefficient C	1	2	2	4

Table 2 The class number and coefficient at choosing autonomous control

Area	Rail	Free
Number of Category	T5	T6
Coefficient C	8	16

The robots are classified mainly into two types, manual control and autonomous control, and the former robots are classified by whether the operator observes tomato directly or indirectly using cameras mounted on robots, and by robot locomotion whether the robot uses rail or not. Totally, the robots are categorized into six types depending on operation and locomotion method as shown in Tables 2 and 3. The success points of one-tomato-harvesting change depending on the robot category, e.g., the point of one-tomato-harvesting for T-1 is 2 × 1 = 2 points, and that of category T-6 is 2 × 8 = 16 points[2].

The 5th tournament, 4 rules are changed by organizing committee. At first, in free style section, we add the slope (height is 100 [mm] and slope distance is 900 [mm].) on the field. 4th competition, we added the step on the field. However, there was no development of robot design with rules with steps. Therefore, we changed step to slope. Second, In the tomato self, we put tomato and black board for recognizing tomato easily in image processing. In this year, the black board has been removed. Third, 4th competition, scores are evaluated as shown in Eq.(1)

$$P = C(2\alpha + \beta) - 2(\gamma + \delta) - \varepsilon \qquad (1)$$

where P is score, C is coefficient of magnification depending on selecting class as shown in Table 1 and 2, α is the number of tomatoes which is no damaged and correct color, β is the number of damaged tomatoes, γ is the number of drop tomatoes and γ is the number of damaged tomatoes which are not harvested. ε is a deducted point when robots damage stalks of tomato plants and the point is deducted in Final Stage. However, if the teams select high class such as T5 and T6, Total score P is almost not influenced from γ, δ and ε. This means that teams selected high class can get high score if robot damage tomatoes and drop tomatoes. In actual tomato factory, damaged tomatoes and drop tomatoes have no commercial value. However, this competition is aim to carefully dealing with tomatoes. Therefore, we proposed new equation for next competition as shown in Eq.(2).

$$\boldsymbol{P = \frac{\alpha}{\alpha+\beta+\gamma+\delta} C(2\alpha + \beta) - 2(\gamma + \delta) - \varepsilon} \qquad (2)$$

Equation (2) is employed harvest rate and if robot damage and drop tomatoes magnification C is decreased. At last, number of robots are not limited.

2.2. *Junior League*

In Junior League, the subject is to carry small size tomatoes to assigned positions. Students should develop the robot using LEGO Mindstorm with the functions such as line trace, color recognition, end-effector with mechanism design and motors with control, and their programming[1].

The basic specification for robot is that the size of the robot is within 300mm x 300mm on ground [2].Height is no limited. Students should develop the robot using LEGO Mindstorm. Competition subjects include Line Trace Challenge, Color Identify Challenge, Mechanism Design and Control Challenge and Object Detection Challenge.

In Line Trace Challenge, robots should detect white line in the competition area and move along the line using a color sensor. The robot starts from the starting point. In the middle of the course, the tomato harvesting field (harvest field) exists, where tomatoes are arranged. The robot must move to the harvest field in order to get the 6 tomatoes. In Color Identifying Challenge, robots should explore and recognize color signs in the middle of the course and the same color of tomatoes. As guidance to harvest field, red, yellow and blue lines are drawn in the field. It is necessary for the robots to detect guidance line by color sensor. The robot gets the tomatoes of the same color. Along the black line for line trace, each color is signed. The robot can move to the harvest field from the lone detecting the color. In Mechanism Design and Control Challenge, robot should manipulate tomatoes using manipulator and carry to the storage location. To pick up the tomato box by using a manipulator equipment which is made by each team, participants are expected to design and make a device to get tomatoes on their idea. The robot is required to store, transport and relocation depending on tomatoes in each color. After picking up the tomatoes, the robot should return to the course. Then, the robot carries the tomatoes to the specific storage location. In Object Detection Challenge, robots should detect a battery charging station and stop there. The 5th tournament, a rule are changed by organizing committee. 4th tournament, tomatoes were stored in a box called "tomato box". However, when using the tomato box, the robot could easily carry tomatoes, so it conveys that it was important to handle plant carefully. Therefore, we decided to use raw tomatoes and make the size irregular. We decided to encourage a new mechanical design by changing rules. And, some points have been changed by changing tomato harvesting rules. Details are given in the rule book [6]. Also, in order to treat the tomatoes carefully, a rule was added to deduct points if the tomatoes are damaged.

3. Results

Teams which could proceed to final stage are 5 teams including HAYASHI-LAB (Category T5, Kyushu Institute of Technology), Nagasaki GANBARANBA (Category T4, Nagasaki Institute of Applied Science), Taccary (Category T5, Kyushu Polytechnic college) and Nishikodai Robo Ken, NIT-LAB(Category T3,Nishinippon Institute of Technology). 1st place of rail-style division is HAYASHI-LAB, 2nd place is NishikodaiRoboKen and 3rd place NIT-LAB. In free style devision 1st place is Taccary and 2nd place is Nagasaki GANBARANBA, and 3rd are no eligible teams. Finally, Overall winner is HAYASHI-LAB (Table 3-5). In Junior League, number of the participated teams were 18 teams and the result is Table 6 and Table 7.

Table 3 Final result of the Senior League

Ranking	Team
Overall winner	HAYASHI-LAB
Overall runner-up	Taccary

Table 4 Result of rail-style division

Ranking	Team
1st place	HAYASHI-LAB
2nd place	NishikodaiRoboKen
3rd place	NIT-LAB

Table 5 Result of free-style division

Ranking	Team
1st place	Taccary
2nd place	Nagasaki GANBARANBA

Table 6 Result of rail-style division

Ranking	Team
1st place	NiAScience (High school attached Nagasaki Institute of Applied Science)
2nd place	Sponge pob (Hita Rinko High School)
3rd place	SORA (Hita Rinko High School)

Table 7 Special award

Ranking	Team
Best presentation award	J-Advance1 (Fukuoka Joto high school)
Technical award	J-Advance2 (Fukuoka Joto high school)
Special award	Kaho Total High school
Challenge award	Kako-1 go (Kashii Technical High School)

Acknowledgements

This works was supported by Kyushu Institute of Technology, Center for Socio-Robotics Synthesis, The University of Kitakyushu, Nippon Bunri University, Nishinippon Institute of Technology, Kitakyushu College of Technology, Nagasaki Prefecture University, Kyushu Polytechnic Collage, Hibikinada-Saien, City of Kitakyushu, The Kitakyushu Chamber of Commerce and Industry and Afrel Co., Ltd.

References

1. K.Ishii, Y.Takemura, T.Matsuo, T.Sonoda, *Tomato Harvesting Competiton*, CD-Proc., 2016 Joint 8th International Conference on Soft Computing and Intelligent Systems (SCIS) and 17th International Symposium on Advanced Intelligent System (ISIS), Sapporo, 2016
2. T.Matsuo, T.Sonoda, Y.Takemura, M.Sato, K.Ishii, *Toward Smart Tomato Greenhouse : The Fourth Tomato Harvesting Robot Competition*, Journal of Robotics, Networking and Artificial Life, Vol.6, Issue 2, pages 138-142, 2019
3. H.Yamaguchi, K.Nagahama, T.Hasegawa, M.Inaba, *Development of an autonomous tomato harvesting robot with rotational plucking gripper*, Proc. of 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems, Daejeon, South Korea, 2016
4. B. Li, S. Yasukawa, T. Fujinaga, and K. Ishii, "End-effector for robotic harvesting of a tomato fruit with calyx", The 2019 International Conference on Artificial Life and Robotics, pp. 544-547, 2019.
5. T. Fujinaga, S. Yasukawa and K. Ishii, "System Development of Tomato Harvesting Robot Based on Modular Design", 2019 International Workshop on Smart Info-Media Systems in Asia , pp. 1-6, 2019.
6. The 5th Tomato robot competition junior rule, "http://www.lsse.kyutech.ac.jp/~sociorobo/wp-content/uploads/2018/10/b90edfbcbe92cb25d524f674988f0109.pdf", 2018, 10

Reports of 7th underwater robot festival in Kitakyushu

Yuya Nishida
Khushu Institute of Technology, 2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196, Japan

Takashi Sonoda
Nishinippon Institute of Technology, 1-11 Aratsu, Kanda, Miyako, Fukuoka 800-0394, Japan

Takayuki Matsuo
National Institute of Technology, Kitakyushu College, 5-20-1 Shii, Kokuraminami, Kitakyushu, Fukuoka 802-0985, Japan

Shinsuke Yasukawa
Khushu Institute of Technology, 2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196, Japan

Masanori Sato
Nagasaki Institute of Applied Science, 536 Abamachi, Nagasaki 851-0193, Japan

Yasunori Takemura
Nishinippon Institute of Technology, 1-11 Aratsu, Kanda, Miyako, Fukuoka 800-0394, Japan

Kazuo Ishii
Khushu Institute of Technology, 2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196, Japan
E-mail: ynishida@lsse.kyutech.ac.jp, ishii@brain.kyutech.ac.jp

Abstract

For enhancement of oceanic engineering technology and researchers, underwater robot competition has been held since 2016. Seventh competition in this year consists AUV league that university's vehicles automatically cruise at field and junior league that underwater craft is made. The paper reports competition regulations of AUV and junior league and results of the competition held in October 2019.

Keywords: Robot competition, Autonomous underwater vehicle, junior league

1. Introduction

To advance field robotics technology, the center for socio-robotic synthesis of Kyushu institute of technology organizes a several competitions every year [1-2]. We held 7th underwater robot festival that consists of AUV and junior leagues for providing a place to present research results, at Kitakyushu in October 2019. The AUV league is the competition that autonomous underwater vehicles (AUVs) developed by research institutes automatically navigate at field and perform tasks by using mounted sensors and camera. High school students and below can join the junior league and take the lectures on mechanical design, electrical circuits and programing required to make unmanned surface vehicle (USV). After the lectures, the students join the game that check the performance of handmade USV.

The paper explains the regulations of the AUV and junior leagues and reports he competition that was held at Aso sports center in Kitakyushu from October 19 to 20.

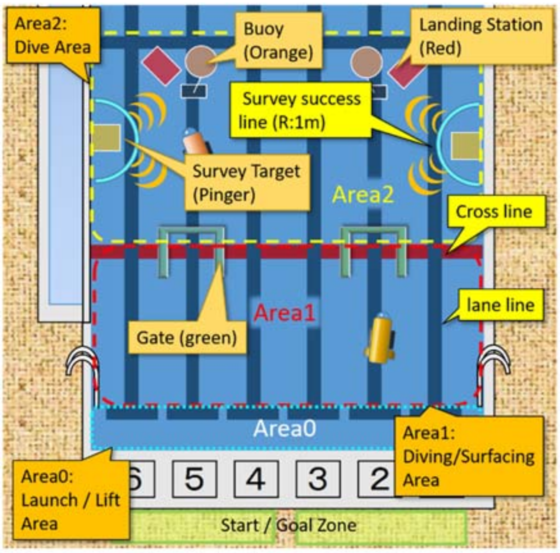


Fig.1 Game field for AUV league

2. Outline of the competition

2.1. *Committee member*

Committee member of the competition consists around professors that belong the university near Kitakyushu. Executive chairman is prof. Nishida belonging Kyushu institute technology of host university, prof. Sato of Nagasaki Institute of Applied Science and prof. Sonoda of Nishinippon Institute of Technology take AUV league, prof. Takemura of Nishinippon Institute of Technology and prof. Matsuo of National Institute of Technology take junior league, and prof. Yasukawa of Kyushu Institute of Technology is home page designer for the competition. Other committee member supports two league and invited participant to the competition.

2.2. *AUV league regulation*

Participants use developed AUV for the game of the AUV league, the game is judged on three items that are presentation (total 40 points), wet test (total 430 points) and deployment technology (total 30 points) inspections. In presentation inspection, university professors evaluate participants presentation with structure, speaking, explicitness, technology, answering and presentation time. Wet test inspection is done at the field shown Fig.1, and each team has five minutes for preparing the AUV and ten minutes for six underwater tasks that are Gate Pass to through the gate (max. 20 points), Buoy Touch to attack colored ball (max. 80 points), Investigation

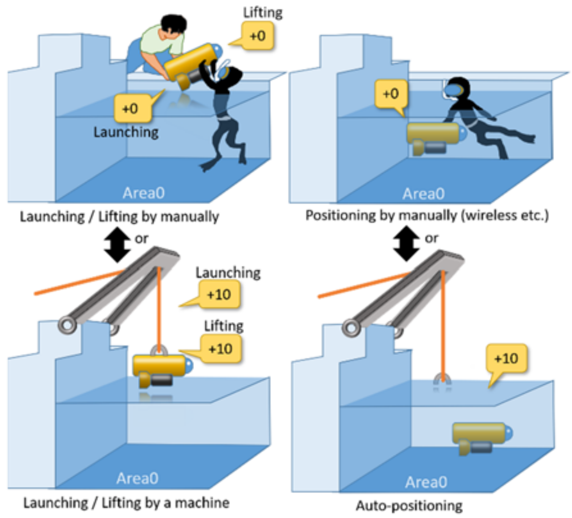


Fig.2 Deployment technology inspection

Mission to search and touch the pinger which nobody know its position (max. 200 points), Landing Mission to put on the bottom table (max. 100 points), Returning Mission to return back to start point (max. 10 points) and Drop Mission to drop marker on the landing table (max. 20 points). Wet test regulation decided with reference the regulation of other international underwater competition, and the Investigation Mission assuming actual survey by the AUV was introduced to the competition since last year. The deployment technology inspection evaluates deployment method which is the most important for actual AUV survey. When the AUV is put in the water by using developed deploy device without diver supports as shown in Fig.2, the team gets points which depend on its deployment technology. Even if the AUV deployment is succeeded, if the game field is corrupted or its deployment method is a danger, the score of its team deducted.

2.3. *Junior league regulation*

The purpose of the junior league is that participants acquires the minimum knowledges required for development of underwater vehicle, and they take the lectures on mechanism, electrical circuit and programing. The committee provides all teaching aids and tools for the lectures, and university professors teach carefully from the basics. Therefore, every high school student and below can join the league and make the robot. Surface vehicle shown in Fig.3 was selected as a teaching aid assuming that participants that can play it at home. The

Fig.3 Surface vehicle made in the junior league

Fig.4 Presentation inspection in the AUV league

vehicle has two thrusters for heading and surge controls and a microcomputer with wireless module for controlling the thrusters by remote control. After the lectures, participants play the game that the surface vehicle navigates to the goal avoiding obstacles. The shape and program of the surface vehicle can be freely modified by participants.

3. Competition results

3.1. *Competition participants*

Six university teams for the AUV league, 14 teams for the junior league and 40 people for tour joined the competition. 15 staff members in addition to the committee member managed the competition, 130 people that is largest in last few years came to the competition venue. Because the expensive AUV must be developed for participate to the AUV league, only the same university team joined the AUV league every year. However, this year new teams such as Hiroshima Institute of Technology and Osaka Prefecture University

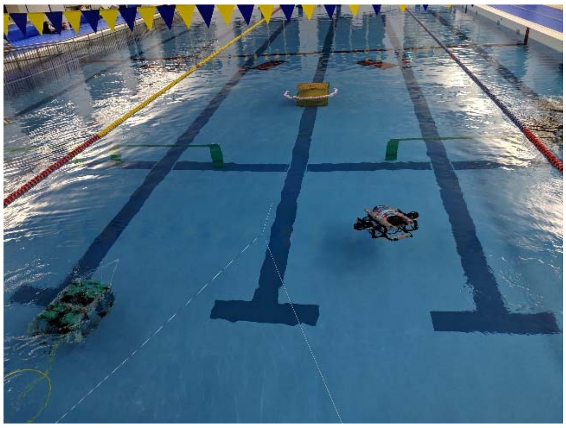


Fig.5 AUVs that cruise in the competition field

were more than usual. The participating schools in the junior league was the same as usual, but many teams entered the league.

3.2. *AUV league results*

Each team played a round robin tournament in last year with a few participant teams, but qualifying was separated two leagues and the final tournament was round robin tournament that consist of top four teams in this year with six participant teams. Methods for the tasks of wet test inspection has been diversifying with improvement of underwater robot technology, each team explained original AUV technology and wet test strategy as shown in Fig.4 in wet test inspection. Because only students developed their AUV without researchers and professors, The AUVs of participant teams often got stuck by various problems every year. However, in this year, the AUVs of all teams cruised in the competition field as shown in Fig.5, and each team competed fiercely. As a result of all inspections, the team of Hiroshima Institute of technology was the champion in the AUV league, the second place was the team of Kyushu Polytechnic College and third place was the teams of the university of Tokyo and Tokyo Institute of Technology.

3.3. *Junior league results*

Almost participant students in junior league was not used to robot making, and there were problems that such as electronic component failure, thruster failure and surface vehicle damage. However, all teams enjoyed making the surface vehicle and was able to complete the vehicle through good offices of the professors and the assistant staffs, as shown in Fig.6. Because the professors

Fig.6 The participants in the junior league during the lectures

explained waterproof structure principle and operational notes of the surface vehicle in the lectures, the vehicles of all teams can cruise on the surface without water leakage. After the lectures, the students actively participated in the game using the vehicle and adjusted their vehicles each in their own way as shown in Fig.7. As results of the game held on the last day, the team of National Institute of Technology, Kitakyushu College was the champion in the junior league and the teams of Jyoto High School were the second and third place.

Fig.7 Their vehicle adjustment by the participants

Fig.8 Group photo of everyone participated in the competition

4. Conclusions

We had a very hard time arranging a venue where the competition could be held even in rainy weather, but the underwater robot festival could be held with safety by the generous cooperation of Kitakyushu City and the facility staff. By actively support of committee members and assistant staffs, we smoothly organized the competition, and all participants of the AUV and junior league enjoyed played the game. We were able to gather many participants through the good offices of the committee members, but there were few general participants who observe the competition. In order to gather more the general participants from next year, glamorous events will be added in the competition.

Acknowledgements

The underwater robot festival in this year was hosted by Center for Socio-Robotic Synthesis, Kyushu Institute of Technology, Japan Underwater Robot Network, the Japan Society of Naval Architects and Ocean Engineers, IEEE/OES Japan Chapter, Marin Technology Society Japan Section and Techno-Ocean Network, and Kitakyushu city and Kitakyushu Convention & Visitors Association were supporting the competition. The staff of Aso Sports Center in Kitakyushu cooperated to the competition. The authors would like to offer my special thanks to all of the above.

References

1. Takayuki Mtsuo, et al., The 4th tomato harvesting robot competition, Proc. of The International Conference on Artificial Life and Robotics, DOI: 10.1109/SCIS-ISIS.2016.0118, Beppu, 2019
2. Forestry Drones & Robots Competition website, Center for Socio-Robotic Synthesis, Kyushu Institute of Technology, http://www.lsse.kyutech.ac.jp/~sociorobo/ja/forestdrone2019, *in Japanese*

Development of IoT Module with AI Function Using STM32 Chip

Jr. Hung Guo[1], Evgeni Magid[2], Kuo-Hsien Hsia[1], Kuo-Lan Su[1]
[1]*Department of Electrical Engineering, National Yunlin University of Science & Technology,*
123 University Road, Section 3,Douliou, Yunlin 64002, Taiwan, R.O.C.
[2]*Intelligent Robotics Department, Kazan Federal University,*
18 Kremlyovskaya street, Kazan 420008, Russian Federation
E-mail: g9710801@yuntech.edu.tw, magid@it.kfu.ru, khhsia@yuntech.edu.tw, sukl@yuntech.edu.tw
www.yuntech.edu.tw

Abstract

The application of Internet of Things (IoT) has been widely used in our lives with the advancement of related software and hardware technologies. In order to make these IoT modules more intelligent, many IoT modules have begun to incorporate artificial intelligence algorithms. Therefore, this paper develops IoT module with STM32 chip as main controller. This module uses fuzzy Analytic Hierarchy Process (fuzzy-AHP) and Adaptive Fusion Method(AFM) to improve the correctness and self-learning ability of the sensor. In terms of communication, the IoT module has Ethernet, Wi-Fi, LoRa, etc. communication interfaces. We also built a web server on this module, so the IoT module can operate directly in the browser. Finally, we developed a monitoring system. Through this monitoring system, multiple IoT modules can be constructed into a sensor network. This monitoring system can also use same algorithm to correct and isolate data from modules or sensors in the network to make this module more intelligent and applicable in different areas.

Keywords: Internet of Things (IoT), Fuzzy-AHP, Adaptive Fusion Method(AFM), Sensor network.

1. Introduction

With the rapid development of various hardware and software, IoT has been widely used in our daily life. As the application of IoT becomes more and more common, we update our requirements on IoT devices and systems. In previous research, we could found the development trend of IoTs has changed. This paper uses STM32 chip to develop IoT system. Since the computing power of this series of chips is sufficient, and for smaller devices that require battery power, there are related solutions, so it is very suitable for applications in the IoTs.

We have integrated various communication interfaces such as wired network (LAN), wireless network (Wi-Fi) and LoRa (Long Range). These communication interfaces can be used at the same time, and it makes each IoT module have powerful communication capabilities. In addition, we used Advanced Encryption Standard (AES) to encrypt the message. Through the powerful communication interface and encryption mechanism, IoT module is more completed in monitoring.

In the sensors part, a variety of interfaces are provided to connect the sensors, includes A/D inputs which can connect 21 inputs. Besides, through I2C, SPI, etc., it can connect 240 sensors. And the digital I/O can connect 24 sensors, so it can provide powerful sensor data processing capabilities.

Finally, we also used fuzzy-AHP and AFM algorithms in this IoT module to process sensor signals and events. This ensures the correctness of sensor signals and categorizes the types of events. The types and correctness data of these sensors and events will also be used to evaluate the correctness and stability of IoT module, so the entire IoT module can be more intelligent.

2. System Architecture

The IoT module use 32-bit STM32 series microcontroller. This series of microcontrollers is based on ARM Cortex™ -M3. It combines high performance, real-time functions, digital signal processing, low power consumption and low voltage operation. In this IoT module, the microcontroller is STM32F103VET6, and is a mid-level chip in STM32 series.
In communication interface part, IoT module can connect LoRa, Wi-Fi, Ethernet, SPI/I2C, USB, Bluetooth and a wired UART (universal asynchronous receiver transmitter). These interfaces can be used simultaneously on the IoT module. The data or calculation results of sensors can also be returned to the host computer through communication interface. Because the communication interface is completed, it makes the application and operation of IoT module easier and more convenient. It can also ensure information can be sent out immediately when the event occurs. This module can be used alone, or through a communication interface, multiple modules can be used to form a sensor network. This allows the entire monitoring system being flexible. The block diagram of IoT module is shown in Fig. 1.

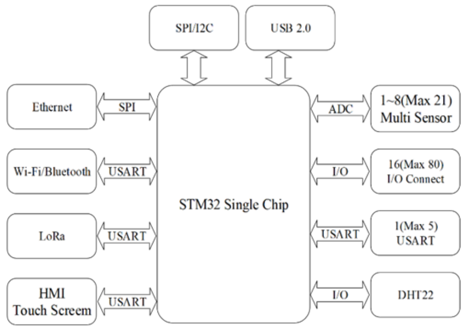


Fig. 1. Block diagram of IoT module.

3. Algorithms

In this paper, we use fuzzy-AHP algorithm and AFM to analyze and judge sensor data. We use fuzzy-AHP for sensor data processing and decision making, then use AFM to analyze these data and decision results. In this way, the false positive rate of IoT module or sensor can be obtained, and the analysis result of these false positive rates is used as one of the weights in the fuzzy-AHP processing. This processing can reduce misjudgment rate of IoT module or sensor, and early detection of possible faulty IoT modules or sensors allows each IoT module to be more accurate and intelligent. Because each module may have a different number of sensors, or even different types of sensors, there are three processing methods in fuzzy-AHP.

3.1. *Use a single sensor*

This is the most common and cheapest mode. Since it only has one sensor, it is easy to cause wrong event messages due to sensor problems. In order to reduce the probability of sensor errors, we continuously read the sensor data, the number of readings is j. We set a threshold has the judgment of sensor data. Therefore, we can get S after comparing the measured value v of the sensor with the threshold value which is shown in following equations:

$$\begin{cases} If\ v_j > h\ then\ S_j = 1 \\ If\ v_j \leq h\ then\ S_j = 0 \end{cases} \quad (1)$$

Because there is only one sensor to do multiple sampling, the result is one-dimensional array, so whether or not an event occurs can be evaluated using Eqs. 2-7. We add the values of 1 in the array and divide by the number of samples. For this calculation, we define a fuzzy-logic to determine the severity of the event.
Define E as the influence of sensor on detection event. For example, we can define E=1~5 to represent the severity of the event as the corresponding degrees normal, possible, warning, danger, emergency, and probability. Then we can define the judgment formula as shown in Eq. 4, we can find out the degree or correlation of the sensor's influence on the event.

$$C = \sum_{k=1}^{j} S_k \quad (2)$$

$$e = C/j \quad (3)$$

$$\begin{cases} if\ \ e \geq 90\%\ \ then\ E = 5 \\ if\ \ e \geq 70\%\ then\ E = 4 \\ if\ \ e \geq 50\%\ then\ E = 3 \\ if\ \ e \geq 30\%\ then\ E = 2 \\ if\ \ e < 30\%\ then\ E = 1 \end{cases} \quad (4)$$

$$\begin{cases} If\ \ E \geq 3\ then\ r = 1 \\ If\ \ E < 3\ \ then\ r = 0 \end{cases} \quad (5)$$

$$M = r \quad (6)$$

$$\begin{cases} If\ \ M = 1\ \ then\ T = true \\ If\ \ M = 0\ \ then\ T = falde \end{cases} \quad (7)$$

3.2. *Use multiple identical sensors*

This is an advanced usage model to avoid event misjudgments. We set i as the number of sensors, j as the number of samples, h as the threshold value, v as the measurement value of the sensor, and w as the weight value. This weight value is used to evaluate whether the sensor is normal. The default value is 1, and the weight value is automatically adjusted through AFM. Next, we will explain how to process and judge sensor data. If the sensor measurement value is less than or equal to the product of the threshold value and the weight value, Si=0; if sensor measurement value v is greater than the product of the threshold value and the weight value, Si=1. We can read multiple data for each sensor, and the number of readings can be set. So, we can rewrite Eq. 1 as below:

$$\begin{cases} If\ \ v_{ij} > h \times w_i \ \ then\ S_{ij} = 1 \\ If\ \ v_{ij} \le h \times w_i \ \ then\ S_{ij} = 0 \end{cases} \tag{8}$$

After collating all comparison results, the following results can be obtained.

$$A = \begin{bmatrix} S_{11} & S_{12} & S_{13} & \dots & S_{1i} \\ S_{21} & S_{22} & S_{23} & \dots & S_{2i} \\ S_{31} & S_{32} & S_{33} & \dots & S_{3i} \\ \vdots & \vdots & & \ddots & \vdots \\ S_{j1} & S_{j2} & S_{j3} & \dots & S_{ji} \end{bmatrix} \tag{9}$$

Then we add 1 in each column of above formula which is the following equation.

$$C_i = \sum_{k=1}^{j} S_{ki}, i = 1..n \tag{10}$$

Then we can get.

$$\mathrm{a} = [C_1\ C_2\ C_3 \dots\ C_i] \tag{11}$$

Similarly, we divide each value in Eq. 11 by the number of measurements.

$$e_i = C_i/j \tag{12}$$

Then bring the result of Eq. 12 into Eq. 4, use the following equation to confirm whether the event occurred.

$$\begin{cases} If\ E_i \ge 3\ then\ r_i = 1 \\ If\ E_i < 3\ \ then\ r_i = 0 \end{cases} \tag{13}$$

$$M = \sum [r_1, r_2, r_3, \dots, r_i] \tag{14}$$

$$\begin{cases} If\ M \ge i/2\ \ then\ T = true \\ If\ M < i/2\ \ then\ T = false \end{cases} \tag{15}$$

This method can avoid erroneous event messages, but it will cause higher costs and more calculations.

3.3. *Use multiple different sensors*

This method can judge and distinguish events more detailed through the sensing characteristics of different sensors. For example, to detect a flame, an ultraviolet sensor, an infrared sensor, a smoke sensor, and a temperature sensor can be used. Because above sensors can detect some of the characteristics of a fire when a fire occurs, but it may not be able to clearly confirm the occurrence of a fire. Therefore, the use of many different sensors can alert the fire event in advance and clearly indicate the status of the fire.

Similar to previous method of using multiple identical sensors, we set i as the number of sensors, j as the number of samples, h as the threshold, v is the measurement value of the sensor, and w is the weight value. We explained how to process and judge the sensor data. If the sensor measurement value is less than or equal to the product of the threshold value and the weight value, then Si=0; if sensor measurement value v is greater than the product of the threshold value and the weight value, Si=1. We can read data multiple times for each sensor, and the number of reads can be set. Therefore, read value of each sensor is compared with threshold value, we can rewrite Eq. 1 as follows:

$$\begin{cases} If\ \ v_{ij} > h_i \times w_i \ \ then\ S_{ij} = 1 \\ If\ \ v_{ij} \le h_i \times w_i \ \ then\ S_{ij} = 0 \end{cases} \tag{16}$$

Then substitute the result of Eq. 16 into Eqs. (9) ~ (12). Because different sensors are used, we can confirm the occurrence of event and also distinguish the status of the event. The judgment formula used here is Eq. 4. For example, when we use four kinds of sensors: ultraviolet, infrared, smoke, and temperature, the E of the sensor is "ultraviolet E=3, infrared E=2, smoke E=4, temperature E=3". Since each sensor has different relationship to fire, we can define a weight g for each sensor. This weight can be set by user, then we can get the fuzzy evaluation equation as the following:

$$u_0 = \sum_{i=1}^{n} (E_i * g_i) \tag{17}$$

By using fuzzy evaluation equation, we can know how likely a fire is. As above example, we can describe the event as “a fire may occur, and there is already smoke, so we must pay attention to the temperature”. This can accurately describe the event situation, and provide event handling methods or as a reference for escape decision. We use fuzzy-AHP on the module to reduce the false positive rate of the sensor, and use AFM to adjust the weight value. This algorithm is very suitable for digital detection signals. The sensor data in IoT module is digital data after being processed by fuzzy-AHP, so it can be analyzed by using AFM. The structure of the entire system is shown in Fig. 2.

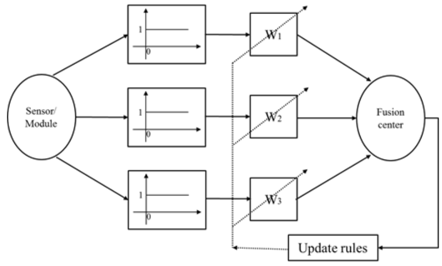
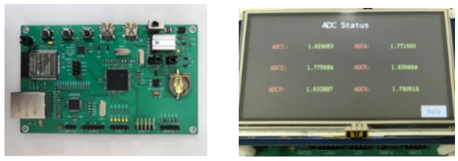


Fig. 2. Fusion architecture.

4. Experimental Results

After completing the design of entire IoT module and the derivation of algorithm, we completed actual IoT module. The actual completed module is shown in Fig. 3. Fig. 3 (a) is the circuit board of Internet of Things module, and Fig. 3 (b) is finished product after adding touched LCD. Users only need to connect the power. Finally, we used three IoT modules which were connected to sensors to verify Fuzzy-AHP and AFM for analysis. As shown in Fig. 4, the overall test results were good.



(a) IoT module main board (b) IoT module with LCD

Fig. 3. IoT module.

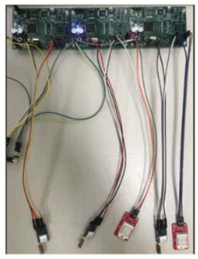


Fig. 4. Experimental Results.

5. Conclusions

We use a 32-bit STM-32 single chip to design IoT module which can connect many sensors and control I/O. In addition, we use fuzzy-AHP algorithm and AFM to process sensor’s data or IoT module. Using these two methods allows IoT modules confirm events and make decisions. IoT module also can learn and adjust the weight itself or the sensor, so it would be more intelligent. In the future, we will continue to develop various sensor modules and related graphic control systems, so this system can be applied in various fields.

References

1. Poniszewska-Maranda, Aneta, et al., Studying usability of AI in the IoT systems/paradigm through embedding NN techniques into mobile smart service system. Computing, 101.11,2019. pp. 1661-1685.
2. Pan, Jianli, and Zhicheng Yang Cybersecurity challenges and opportunities in the new edge computing+ IoT world. In: Proceedings of the 2018 ACM International Workshop on Security in Software Defined Networks & Network Function Virtualization. ACM, 2018. pp. 29-32.
3. Salah, Khaled, et al. "Blockchain for AI: review and open research challenges." IEEE Access 7, 2019, pp.10127-10149.
4. Alam, Furqan, et al. "Data fusion and IoT for smart ubiquitous environments: a survey." IEEE Access 5, 2017, pp. 9533-9554.
5. Canedo, J., & Skjellum, A. (2016, December). Using machine learning to secure IoT systems. In 2016 14th Annual Conference on Privacy, Security and Trust (PST), 2019, pp. 219-222.
6. Luo, Ren C., Kuo L. Su, and Kuo H. Tsai. "Fire detection and isolation for intelligent building system using adaptive sensory fusion method." Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on. Vol. 2. IEEE, 2002.
7. N. Ansari, J.G. Chen, Y.Z. Zhang, “Adaptive Decision Fusion for Unequiprobable Sources,” IEE Proceedings Radar, Sonar and Navigation, 1997, pp.105.
8. N. Ansari, E.S.H. Hou, B.O. Zhu and J.G. Chen, “Adaptive Fusion by Reinforcement Learning for Distributed Detection Systems,” IEEE Transactions on Aerospace and Electronic Systems,1996, Vol. 32, No. 2, pp.524-531

Markerless Indoor/Outdoor Augmented Reality Navigation Device Based on ORB-Visual-Odometry Positioning estimation and ORB-Visual-Mapping Image Registration

Chian C. Ho, Guan-Jie Wang
Department of Electrical Engineering, National Yunlin University of Science and Technology
Douliou, Yunlin County 64002, Taiwan
E-mail: futureho@yuntech.edu.tw, M10512077@yuntech.edu.tw
www.yuntech.edu.tw

Abstract

For markerless indoor/outdoor Augmented Reality Navigation (ARN) technology, camera pose is inevitably the fundamental argument of positioning estimation and pose estimation, and floor plane is indispensably the fiducial target of image registration. Based on Oriented FAST and Rotated BRIEF (ORB) feature with descriptors, this paper proposes ORB-visual-odometry positioning estimation and ORB-visual-mapping image registration to improve camera pose estimation and floor plane detection for making ARN more precise and reliable with real-time performance. Experimental results show both ORB-visual-odometry positioning estimation and ORB-visual-mapping image registration have higher accuracy and reliability than conventional well-known camera-pose-based positioning estimation and floor-plane-based image registration methods, respectively, for ARN. On the other hand, markerless indoor/outdoor ARN technology with proposed two methods have seamlessly been implemented on the portable Android platform and have smoothly been verified to co-work well on the portable Android platform.

Keywords: augmented reality, positioning estimation, image registration, navigation.

1. Introduction

With the rapid development and wide deployment of emerging technologies in ubiquitous computing field, personal navigation system that can accommodate to perform well indoors and outdoors have drawn more and more interest. This is because personal navigation device offering personal destination awareness anywhere and anytime is essential to versatile ubiquitous computing applications, like indoor/outdoor path directions, indoor/outdoor information guide, indoor/outdoor marketing advertisement, indoor/outdoor social networking, and so on.

However, conventional 2D (birdview) or 3D (overlook) virtual-model navigation devices are not intuitive enough to guarantee the proper perception alignment between virtual-model navigation scenario and real-world navigation situation. Besides, conventional 2D (birdview) or 3D (overlook) virtual model navigation devices fail to be evolved into the ever-growing wearable see-through devices, like eyeglasses, helmet, or goggles. More seriously, staring at 2D (birdview) or 3D (overlook) virtual model navigation devices on driving or on foot might cause careless accidents or dangerous occurrences.

Augmented Reality Navigation (ARN) technology is the best alternative choice and becomes more and more fascinating in ubiquitous computing applications since it can lay real-world navigation directions over what users are actually seeing in front of themselves in the true world.[1-8] But quite a few faults still exist in conventional ARN technologies summarized as Table 1. In general, according to the object-tracking principles, ARN technologies can be divided into three main categories: 1)

Table 1. Summarized comparison between various conventional ARN technologies and this work.

ARN	Positioning Estimation	Path Planning	Image Registration	Pose Estimation
Kim at al. [1] in 2008	Marker (Landmark, color)	Node Graph (Distance matrix)	Marker	None
Mohareri at al. [2] in 2011	Marker (Landmark, audio)	Predefined (no replanning)	Marker (Square pattern)	Marker (Square pattern)
DiVerdi at al. [3] in 2008	Markerless (optical flow) & Positioning (GPS, beacon)	Predefined (no replanning)	Markerless & Positioning	None
Hile at al. [4] in 2008	Markerless (Floor-to-wall points or corners)	None	Map, markerless, and pose	Markerless (Floor-to-wall points or corners)
Oskiper at al. [5] in 2015	Markerless (Scene landmark) & Positioning (GPS)	None	Marker (3D features)	Marker (3D features)
Cheok at al. [6] in 2011	Positioning (Fluorescent lamps, GPS)	None	Positioning (Fluorescent lamps, GPS)	None
Hervas at al. [7] in 2014	Positioning (GPS) without indoor positioning	Google Directions	Positioning (GPS)	Positioning (GPS)
Thomas at al. [8] in 2009	Positioning (GPS) without indoor positioning	None	Positioning (GPS)	Positioning (GPS)
Our proposed work	Markerless (ORB-Visual-Odometry)	Node Graph (Dijkstra)	Markerless (ORB-Visual-Mapping)	Markerless (ORB-Visual-Odometry)

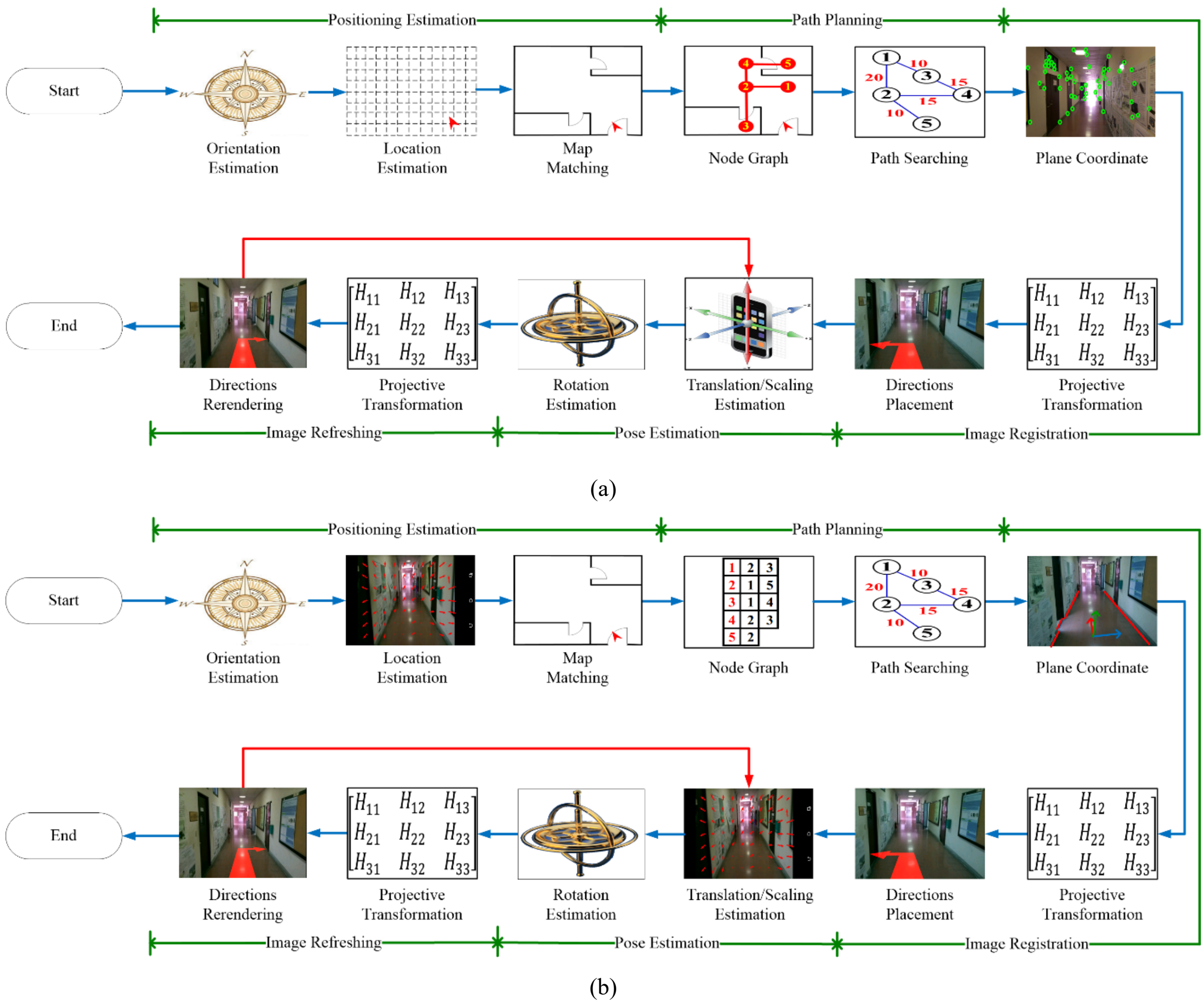


Fig. 1. Flowchart of (a) typical and (b) proposed markerless indoor/outdoor ARN technologies.

marker, 2) markerless, and 3) positioning. The marker ARN technology is the most reliable one tracking the registered objects and projects the computer graphics onto the registered objects based on distinctive markers or landmarks.[1,2] Nevertheless, a large amount of deployment of distinctive markers or landmarks for marker ARN technology are very expensive and unrealistic. The markerless ARN technology simply on the cost-effective basis of natural or plain visual features, like points, lines, corners, textures, is actually recognized as the most practical but difficult one.[3-5] Because the markerless ARN technology depends only upon elaborate computer vision algorithms, not upon marker or external positioning infrastructure. The positioning ARN technology, as its name implies, relies upon the available location coordinate from outdoor Global Positioning System (GPS) infrastructure or specific indoor Wireless Sensor Networking (WSN) infrastructure without specific markers or landmarks,[6-8] but these indoor/outdoor infrastructure is power-consuming and cost-expensive, especially is error-prone to the radio signal's drift effect resulting from external electromagnetic interference or internal multipath echoes.[9]

Due to the markerless characteristic of main playground of outdoor and indoor real-world environments, most of recent research activities and interests in ARN technologies focuses on markerless issue, especially for indoor scenarios.[10-12] Figure 1(a) shows the algorithm flowchart of typical markerless indoor/outdoor ARN device. In Figure 1(a), first of all, the indoor/outdoor ARN device finishes the estimation of orientation angle through fusion of digital compass and Inertial Measurement Unit (IMU) sensors at the step of orientation estimation, in the stage of positioning estimation. Meanwhile, the indoor/outdoor ARN device also has to finish the estimation of location coordinate through external radio positioning infrastructure (e.g., WiFi, Bluetooth LE, RFID) or internal inertial positioning unit at the step of location estimation, in the stage of positioning estimation. Then, the indoor/outdoor ARN device matches the estimated location coordinate and orientation angle of the targeted device onto the built 2D map at the step of map matching, in the stage of positioning estimation. Next, the indoor/outdoor ARN device has to build the critical nodes on the built 2D map at the step of node graph, in the stage of path planning, and has to finish the shortest or fastest navigation path evaluation to the destination on the built 2D map at the step of path searching, in the stage of path planning. After accomplishing the floor plane detection (including camera pose initialization) at the step of plane coordinate, in the stage of image registration, and Homography-based 2D-to-3D projective transformation of the map and navigation path at the step of projective transformation, in the stage of image registration, the indoor/outdoor ARN device can accurately project the ARN path directions onto the markerless real-world floor plane through elaborate floor region segmentation at the step of directions placement, in the stage of image registration. Afterward, when the camera (viewing) pose of the indoor/outdoor ARN device varies, the translation/scaling and rotation variations of camera (viewing) pose in real-world coordinate can be properly evaluated at the steps of translation/scaling estimation and rotation estimation, respectively, in the stage of pose estimation. Finally, the ARN path directions on the screen of the indoor/outdoor ARN device can be rotated and deformed properly in object (image) coordinate at the step of projective transformation, in the stage of image refreshing, so as to be closely stuck onto the markerless real-world floor plane through elaborate camera pose estimation at the step of directions rerendering, in the stage of image refreshing.

Among stages of the algorithm flowchart of the typical markerless indoor/outdoor ARN device in Figure 1(a), the accuracy and reality of ARN path directions is actually sensitive and vulnerable to these three stages of positioning estimation, image registration, and pose estimation, especially to these three steps of location estimation, plane coordinate, and translation/scaling estimation. For markerless indoor/outdoor ARN technology, camera pose is inevitably the fundamental argument of positioning estimation and pose estimation, and floor plane is indispensably the fiducial target of image registration. Thus this paper concentrates to address on issues of camera-pose-based positioning estimation and floor-plane-based image registration in the subsequent sections.

2. Conventional Methods

There are many conventional camera-pose-based positioning estimation methods, like frame-to-frame planar homographies and feature-based optical flow field, and floor-plane-based image registration methods, like like Otsu and moving-average binarization floor detection and wall-floor-boundary floor detection, applied to the indoor/outdoor ARN device in the past 10 years. But these conventional methods are rightly the dominant factors to degrade the accuracy and reality of markerless indoor/outdoor ARN technologies.

3. Proposed Methods

This paper proposes ORB-visual-odometry positioning estimation and ORB-visual-mapping image registration to improve these three stages of positioning estimation, image registration, and pose estimation of typical markerless indoor/outdoor ARN technologies, especially to improve these three steps of location estimation, translation/scaling estimation, and plane coordinate, illustrated in Fig. 1(b).

4. Experimental Results

This paper adopts the Android-based smartphone of Asus ZenFone 4 Pro as the experimental platform for comparison of various camera pose estimation methods and various floor plane detection methods. There are 5 indoor/outdoor scenes of (a) classroom, (b) corridor, (c) lobby, (d) fire lane, and (e) square, in the experiment. From the experimental results, it is evident that proposed ORB-visual-odometry positioning estimation has much higher localization accuracy and reliability than conventional positioning estimation or pose estimation methods, at similar execution latency, especially under cluttered environments. Besides, it is obvious that proposed ORB-visual-mapping image registration has much higher alignment accuracy and less execution latency than conventional floor-detection-based image registration methods, especially under high-contrast environments.

5. Conclusions

Although lots of visual SLAM alternatives, like direct type or semi-direct type, also can improve the issues of positioning estimation, image registration, and pose estimation, this paper proposes and implements two effective and efficient methods, ORB-visual-odometry positioning estimation and ORB-visual-mapping image registration, to make markerless ARN technology more precise and more reliable with real-time performance.

References

1. Jongbae Kim and Heesung Jun, Vision-based location positioning using augmented reality for indoor navigation, IEEE Transactions on Consumer Electronics, vol. 54, no. 3, Aug. 2008, pp. 954–962.
2. O. Mohareri and A. B. Rad, Autonomous humanoid robot navigation using augmented reality technique, in Proceedings of 2011 IEEE International Conference on Mechatronics (ICM), Apr. 2011, pp. 463–468.
3. S. DiVerdi and T. Hollerer, Heads up and camera down: a vision-based tracking modality for mobile mixed reality, IEEE Transactions on Visualization and Computer Graphics, vol. 14, no. 3, May–Jun. 2008, pp. 500–512.
4. H. Hile and G. Borriello, Positioning and orientation in indoor environments using camera phones, IEEE Computer Graphics and Applications, vol. 28, no. 4, Jul.–Aug. 2008, pp. 32–39.
5. T. Oskiper, M. Sizintsev, V. Branzoi, S. Samarasekera, and R. Kumar, Augmented reality binoculars, IEEE Transactions on Visualization and Computer Graphics, vol. 21, no. 5, May 2015, pp. 611–623.
6. A. D. Cheok and Li Yue, A novel light-sensor-based information transmission system for indoor positioning and navigation, IEEE Transactions on Instrumentation and Measurement, vol. 60, no. 1, Jan. 2011, pp. 290–299.
7. R. Hervas, J. Bravo, and J. Fontecha, An assistive navigation system based on augmented reality and context awareness for people with mild cognitive impairments, IEEE Journal of Biomedical and Health Informatics, vol. 18, no. 1, Jan. 2014, pp. 368–374.
8. B. H. Thomas and C. Sandor, What wearable augmented reality can do for you, IEEE Pervasive Computing, vol. 8, no. 2, Apr.–Jun. 2009, pp. 8–11.
9. C. R. Pratiwi1, P. Kristalina, and A. Sudarsono, A performance evaluation of modified weighted pathloss scenario based on the cluster based-PLE for an indoor positioning of wireless sensor network, International Journal of Engineering and Technology Innovation, vol. 9, no. 1, 2019, pp. 61 – 74.
10. A. I. Comport, E. Marchand, M. Pressigout, and F. Chaumette, Real-time markerless tracking for augmented reality: the virtual visual servoing framework, IEEE Transactions on Visualization and Computer Graphics, vol. 12, no. 4, Jul.–Aug. 2006, pp. 615–628.
11. Jongbae Kim and Heesung Jun, Vision-based location positioning using augmented reality for indoor navigation, IEEE Transactions on Consumer Electronics, vol. 54, no. 3, Aug. 2008, pp. 954–962.
12. O. Mohareri and A. B. Rad, Autonomous humanoid robot navigation using augmented reality technique, in Proceedings of 2011 IEEE International Conference on Mechatronics (ICM), Apr. 2011, pp. 463–468.

Apply Model-Free Adaptive Control Approach for Mobile Robot Path Following

Chun-Chi Lai
Bachelor Program in Interdisciplinary Studies, National Yunlin University of Science and Technology, 123 University Road, Section 3, Douliou, Yunlin 64002, Taiwan, R.O.C.

Chia-Jen Lin
TECO Electric & Machinery Co., Ltd., 4F, No.3, Park St., Nan-Kang, Taipei 11503, Taiwan R.O.C.

Kuo-Hsien Hsia
Bachelor Program in Interdisciplinary Studies, National Yunlin University of Science and Technology, 123 University Road, Section 3, Douliou, Yunlin 64002, Taiwan, R.O.C.

Kuo-Lan Su
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: cclai@yuntech.edu.tw, jeromy@teco.com.tw and khhsia@yuntech.edu.tw

Abstract

The state of the art of mobile robot path planning is composed with global planner and local planner. For example, the global planner majorly establishes a suitable path from a knowing map based on the shortest path. And the local planner which computes the velocity command that includes obstacle clearance and progress towards the goal. However, due to the 2D laser ranger on a mobile service robot just detects the part of the meal table or misses it, this will cause robot often hits the meal table. In this work we apply adaptive control method to feed the goals for DWA's computing. The experimental result shows that DWA will follow the pre-defined path closely and smoothly.

Keywords: Mode- Free Adaptive Control, Dynamic Window Approach, Path follower, Mobile Service Robot.

1. Introduction

With the rapid development of robotics and artificial intelligence technology, mobile service robots have begun to be used in our lives. Based on the design of the automatic navigation function, the mobile service robot is allowed to move autonomously and perform daily human services, such as automatic cleaning and automatic delivery. Based on simultaneous localization and mapping [1] (SLAM), the autonomous navigation of mobile robot will become a reality. If exclude the SLAM, the remaining works of autonomous navigation are including path planning and path following.

The state of the art of mobile robot path planning is composed with global planner and local planner. For example, the global planner majorly establishes a suitable path from a knowing map. And the local planner not only follows the global path but also obstacle avoidance consideration. In this work, we apply the pre-defined path and design a path follower to replace the autonomous global planner to overcome the actual navigation problem when a mobile service robot is delivering the meal in a fast food restaurant.

2. Mobile Robot Path Planning

Path planning is the most basic part of mobile robot navigation. It refers to how the robot can find collision free waypoints as an appropriate motion path from the starting point to the end point in a working environment. In general, the path planning mainly involves these two major issues:

- Obstacles avoidance
- Optimizing the path

According to the environmental information, mobile robot path planning can be divided into global path planning and local planning.

2.1. *Global Planner*

The global planner is a pre-planning method that depends on the analysis of known map. For example, the Artificial Potential Field (APF) [2] algorithm was proposed in 1989. Its' concept is to apply a virtual attraction and repulsive force to a robot. The target gives the robot attractive force, and the obstacle is the repulsive force. After combining the force, a moving path of the robot to the target point can be obtained. The Rapid Exploration Random Tree (RRT) [3] method was originally proposed 1998. Its' idea is to growth with the initial point as the root in the space and add child nodes by random sampling. Connect the closest child node to the root node until the child gradually reaches the target point, and then a feasible path could be selected. The other well-known methods of safe path creation are such as the Generalized Voronoi Diagram (GVD) [4] and the Probabilistic Roadmaps [5] (PRMs). The PRMs applies a sampling to the environment in order for a graph to be created, each edge of which is considered safe for robot traversing and a graph search algorithm such as Dijkstra [6] or A* [7] is applied to the resulting graph to determine a path between the starting and goal. In nowadays, the occupancy grid map [8] is popular to represent the posterior probability model of environment. And the graph search algorithm Dijkstra or A* is also suit for finding the optimal path which is a collection of cells. After the success of finding the global path from the start to the goal, all the selected cells are translated into positions as the point vector.

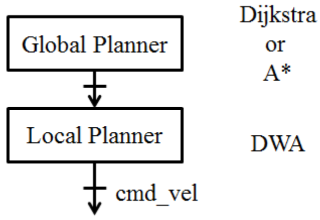


Fig. 1. Mobile robot path planning flow diagram

2.2. *Local Planner*

In order to transform global path into a suitable waypoints, local planner applies sensors to detect the robot's environment to obtain pose or geometric information of obstacles. Therefore, using the updated local sensing information, the local planner generates an obstacle avoidance strategy and attempts to match the trajectory provided by the global planner. Some local trajectory generation of local planner may apply Bezier lines [9], arcs, segments [10], or splines [11]. In nowadays, one of the most popular local path plan method is dynamic window approach (DWA) [12], the basic idea of the dynamic window approach algorithm is as follows:

- The DWA method is a local response obstacle avoidance controller, which searches for the best action (translation velocity v and angular velocity ω) in the velocity space
- Considering the dynamic constraints of the robot to reduce the sample space.
- The trajectories corresponding to different velocities can be represented by a series of arcs with different curvatures.
- In order to obtain a fast response, it is assumed that the speed of the robot is constant in all future intervals.

The sampled trajectories are evaluated by an objective function *O(v,w)*, which is consisted with the goal heading, obstacle clearance, and velocity as below criteria:

$$O(v,\omega) = \sigma(\alpha \times \mathrm{heading}(v,\omega) + \beta \times \mathrm{dist}(v,\omega) + \gamma \times \mathrm{velocity}(v,\omega) \quad (1)$$

The $\mathrm{heading}(v,\omega)$ is the angle difference between robot current head orientation and robot to goal direction, the $\mathrm{dist}(v,\omega)$ is the distance to the closest obstacle, and the

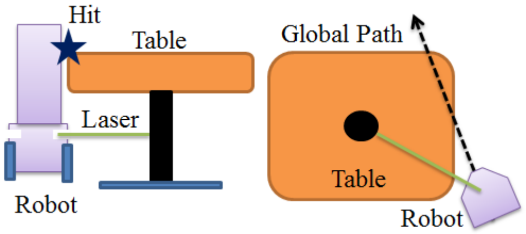


Fig. 2. Auto global path planning problem: the shortest global path often causes mobile service robot to hit the meal tablet.

velocity(v, ω) presents the forward velocity of the robot with kinematic constraints.

2.3. *Problem in Actual Navigation*

In practical applications, the navigation package released in ROS [13] framework are applied to an automatic meal delivery robot in a fast food restaurant. Due to the pedestal column of meal table in the fast food restaurant is shown in Fig.2. Although this pedestal column could be sensed via the laser ranger equipped on robot platform, but sometimes it misses the detection and the upper body of robot hits the corner of table that causes the meal deliver task to be failed. In this work, we apply the pre-define path and path follower to replace the autonomous global planning to overcome the actual navigation problem which causes by the meal table.

3. Path Follower Design

There are four sequences for path follower:

- Get the pre-defined path from file or Topic.
- Prune the Waypoints: Erase the unsuitable waypoints in the path.
- Segment Regulation: Adjust the segment length between waypoints.
- Send out or publish the modified path for DWA.

3.1. *Prune the Waypoints*

The first work in prune the waypoint is to check if the waypoint should be discarded. Assume for a given path with five waypoints $P_0 \sim P_4$, if the initial waypoint P_0 is in front of robot as shown in Fig. 3(a) the angle θ between $\overrightarrow{v_1}$ and $\overrightarrow{v_2}$ will greater than $\pi/2$, where $\overrightarrow{v_1}$ is the vector of robot pose and P_0 and $\overrightarrow{v_2}$ is the vector of P_0 and P_1. On the other hand, if the angle θ between $\overrightarrow{v_1}$ and $\overrightarrow{v_2}$ are smaller than $\pi/2$ as shown in Fig. 3(b), then the P0 point will be deleted and the point data will be recorded as P_{past}. In Fig. 3(c), L_1 is the distance between robot center pose and P_0. L_2 is distance between robot center pose and P_{past}. When $L_1 > L_2$ but robot cannot turn to P_0 due to the local obstacle. In this situation the P_0 point will also be deleted and the point data will be recorded as P_{past}.

3.2. *Segment Regulation*

At this stage, we need an indicator as the criterion for judging how close that robot is following the original path. Calculate the vertical distance "height1" as shown in Fig.3 (d) between the robot pose and the segment of P_{past} to P_0 as the performance criteria for following the line. The "height1" distance can be calculated as below equation:

$$\overrightarrow{V_3} \times \overrightarrow{V_4} = |\overrightarrow{V_3}||\overrightarrow{V_4}| \sin\theta \qquad (2)$$

$$\text{height1} = \frac{\overrightarrow{V_3} \times \overrightarrow{V_4}}{|V_3|} \qquad (3)$$

- Model-free adaptive control

Model-free adaptive control was proposed by Hou [14] which can realize the parameter adaptive control and structure adaptive control of the controlled system which makes it more applicable for many practical plants.

$$u(k) = u(k-1) + \frac{\rho\phi(k)}{\lambda + \phi^2(k)}\left(y^*(k+1) - y(k)\right) \qquad (4)$$

$$\phi(k) = \phi(k-1) + \frac{\eta \Delta u(k-1)}{(\mu + \Delta u^2(k-1))(\Delta y(k) - \phi(k-1)\Delta u(k-1))} \qquad (5)$$

$$\phi(k) = \phi(1), \text{ if } \phi(k) \le \varepsilon \text{ or} |\Delta u(k-1)| \le \varepsilon \qquad (6)$$

From the above equation (4), (5) and (6), consider the

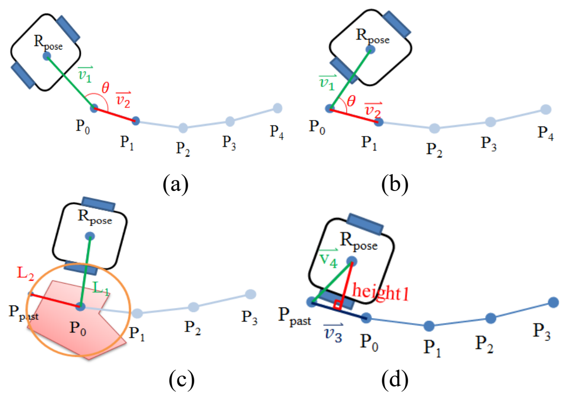


Fig. 3. Prune the waypoints of default path.

previously calculated height1 as y(k) and u(k) as control variables to regulate the length of the path.

4. Experiment Results

The experiment is firstly simulated and demonstrated in ROS rviz user interface for verifying the path following result. In Fig. 4 (a), the pre-defined path is shown as the green circle which is consisted with many piecewise segments. In Fig. 4 (b), the blue circular path represents the path following result with smoothly. In a real fast food restaurant, the pre-defined path is established as the green segment as shown in Fig. 4 (c). In Fig. 4 (d), mobile robot follows the path well and prepares to reach the goal.

5. Conclusions

The state of the art of mobile robot path planning is composed with global planner and local path planner. However for meal deliver application in a real fast food restaurant, the mobile service robot often hits the corner of the meal table due to the 2D laser ranger just detects the part of the meal table or misses it. In this work, we apply the pre-defined path and path follower to replace the autonomous global planning to overcome the actual navigation problem which causes by the meal table. The experimental result shows that DWA can also follow the pre-defined path closely and smoothly.

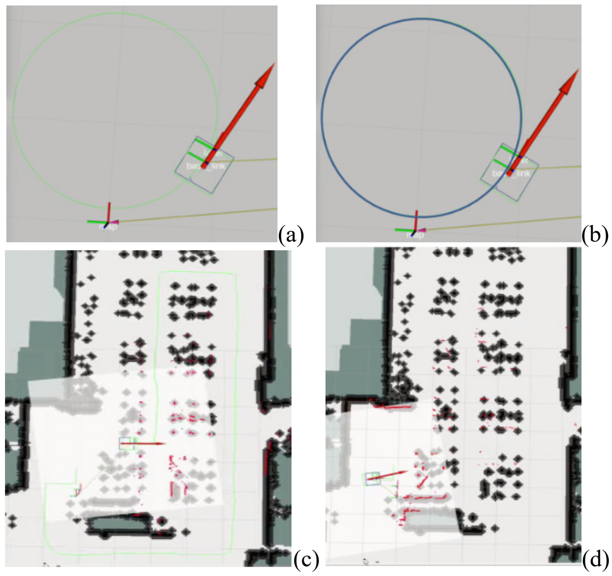


Fig. 4. Robot path following demo: (a) circular path generation (b) circular path following simulation (c) path generation in a fast-food restaurant (d) robot follows the path to the final goal.

References

1. H. Durrant-Whyte and T. Bailey, Simultaneous localization and mapping: part I, *IEEE Robotics & Automation Magazine*, 3(2), June 2006, pp. 99-110.
2. O. Khatib, Real-time obstacle avoidance for manipulators and mobile robots, *IEEE International Conference on Robotics and Automation*, 1985, pp. 500-505.
3. S. M. LaValle, Rapidly-exploring random trees: A new tool for path planning, 1998.
4. S. Garrido, L. Moreno, M. Abderrahim and F. Martin, Path Planning for Mobile Robot Navigation using Voronoi Diagram and Fast Marching, *International Conference on Intelligent Robots and Systems*, 2006, pp. 2376-2381.
5. L. E. Kavraki, M. N. Kolountzakis and J. -C. Latombe, Analysis of probabilistic roadmaps for path planning, *IEEE Transactions on Robotics and Automation*, 14(1), 1998, pp. 166-171.
6. D. E. Knuth, A generalization of Dijkstra's algorithm, *Information Processing Letters*, 6(1), 1977, pp. 1-5.
7. P. E. Hart, N. J. Nilsson and B. Raphael, A Formal Basis for the Heuristic Determination of Minimum Cost Paths, *IEEE Transactions on Systems Science and Cybernetics*, 4(2), 1968, pp. 100-107.
8. S. Thrun and A. Bücken. Integrating grid-based and topological maps for mobile robot navigation, *Proceedings of the National Conference on Artificial Intelligence*, 1996.
9. J. P. Rastelli, R. Lattarulo and F. Nashashibi, Dynamic trajectory generation using continuous-curvature algorithms for door to door assistance vehicles, *IEEE Intelligent Vehicles Symposium Proceedings*, 2014, pp. 510-515.
10. J. A. Reeds and L. A. Shepp, Optimal paths for a car that goes both forwards and backwards, *Pacific Journal of Mathematics*, 145(2), 1990, pp. 367–393.
11. A. Piazzi, C. G. Lo Bianco, M. Bertozzi, A. Fascioli, and A.Broggi, Quintic G2-splines for the iterative steering of vision-based autonomous vehicles, *IEEE Transactions on Intelligent Transportation Systems*, 3(1), 2002, pp. 27–36.
12. D. Fox, W. Burgard and S. Thrun, The dynamic window approach to collision avoidance, *IEEE Robotics & Automation Magazine*, 4(1), 1997, pp. 23-33.
13. http://wiki.ros.org/move_base
14. Z. S. Hou and W. Huang, The model-free learning adaptive control of a class of SISO nonlinear systems, *American Control Conference*, 1, 1997, pp. 343-344.

MALWARE CLASSIFICATION USING DEEP LEARNING

CHENG-HSIANG LO, TA-CHE LIU, I-HSIEN LIU, JUNG-SHIAN LI*
Department of Electrical Engineering / Institute of Computer and Communication Engineering, National Cheng Kung University, Tainan City, 70101, Taiwan

CHUAN-GANG LIU
Department of Applied Informatics and Multimedia, Chia-Nan University of Pharmacy and Science, Tainan City, 71710, Taiwan

CHU-FEN LI
Department of Finance, National Formosa University, Yunlin County 632, Taiwan
E-mail: {chlo, tcliu, ihliu}@cans.ee.ncku.edu.tw, chgliu@mail.cnu.edu.tw, jsli@mail.ncku.edu.tw
www.cans.ee.ncku.edu.tw

Abstract

We'll display two different kinds of experiments, which are NIDS-based and Dynamic-based analysis shows how artificial intelligence (AI) helps us detecting and classify malware. On the NIDS-based intrusion detection, we use CICIDS2017 as a research dataset, embedding high dimensional features and find out redundant features in the raw dataset by Random Forest algorithm, reach 99.93% accuracy and 0.3% of the false alert rate. We extract the function calls in malware data by the method proposed in this paper to generate text data. The algorithm n-gram and TF-IDF are used to process text data, converts them into numeric features, and by another feature selection methods, we reduce the training time, achieve 87.08% accuracy, and save 87.97% training time in Dynamic-based analysis.

Keywords: *NIDS, Dynamic analysis, Deep Learning.*

1. Introduction

In recent years, with the rapid development of Internet technology, people can find the information easily on the Internet or share information with others. Cloud technology, 5G and other emerging network technologies have also made data exchange faster and more convenient. At the same time, cyber security has become one of the biggest concerns nowadays, traditional equipment has less ability to deal with diverse and complex attack techniques. How to promote equipment and software, and improve the confidentiality, integrity and accessible of users has become an important issue. Fortunately, today's hardware is better than pass, huge amount of data can be stored and analyzed, especially in internet, leads the application of artificial intelligence technology to cyber security.

2. Background and Related work

2.1. *Static Analysis v.s Dynamic Analysis*

There are two main methods when we want to analysis the malware comes into our local devices like computers or servers, static analysis and dynamic analysis. Static analysis is a white box analysis method. As the name implies, when you analyze, it doesn't execute the executable file of the malicious program. Instead, it directly analyzes the internal process operation or data usage based on the binary executable file or the original code, because of its analysis way, the advantage of static analysis is that it has low infection opportunity. However, static analysis often requires reverse engineering to disassemble of the executable file. Even some malicious programs that have been protected by a shell must be unpacked by a specific tool before reverse engineering. The features commonly used in static analysis methods are the following, such as Operation Codes [1] [2]and byte sequences, or extracting useful features from portable executable files. On the other hand, dynamic analysis is a black box analysis method, which means that it is necessary to start a malicious program during execution, it will be executed in a virtual environment, and record the behavior of the malicious program such as access file writing and deleting, network connection, Mutexes, Registry Keys modification and API function calls, etc.

2.2. *NIDS v.s. HIDS*

Another way to detect the malware is building the intrusion detection system(IDS), which can be divided into two different kinds by their main function: Network-based IDS (NIDS) and Host-based IDS (HIDS). NIDS mainly detects the attack by network flow, whereas HIDS detects abnormal user behavior on local host computer. both of them compare the log file to their database, the detection method can also separate into two ways: Misuse-based and Anomaly-based. Misuse-based, as known as Signature-based, collect the signature of malware constantly first, and then build a malware signature database, if a network traffic flow or behavior matches the signature in malware database, it will be identified as abnormal. On the other hand, Anomaly-based predefines the normal signature to detect the attack. In this study, we use CICIDS2017[3] as experiment dataset, which collected the data from NIDS.

3. Experiment Architecture

In this part, we'll introduce the methods we used in this paper, for both dynamic-based analysis and NIDS-based malware classification. Our dynamic-based experiment architecture is shown in Figure 1, and NIDS-based is shown in Figure 2.

3.1. *Data Pre-Processing*

Since the original malware data has some noise or untrainable features that makes model predict result worse, so we should take them off. The first challenge we face is that some data has no label, which makes supervised learning impossible. To deal with this problem, we proposes a method that uses a variety of anti-virus software as the basis and produces a final label by majority decision. Although the method is more complicated, it obtains a more credible label than a single anti-virus software.

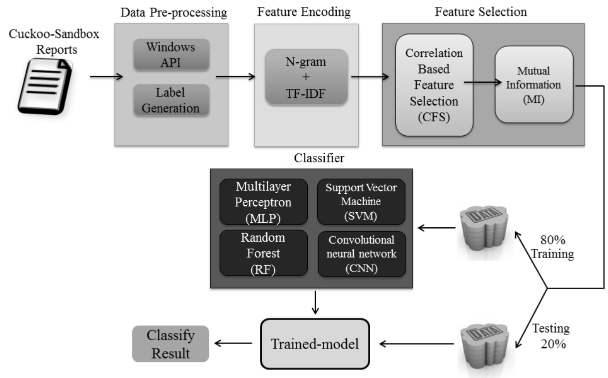


Fig.1. Dynamic-Based Experiment Architecture

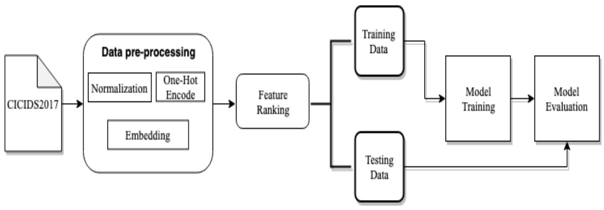


Fig.2. NIDS-based Experiment Architecture

The second challenge is that some non-numerical features such as strings or symbols are exist, which are untrainable features. We use different encode methods for different area to solve this problem. In dynamic-based analysis, the TF-IDF algorithm is used It is a weighting technique that is often used in data mining and information retrieval as a statistical method. In contrast to Bag-of-Word, TF-IDF in addition to counting the frequency of occurrence of words in a single text, it is also used to assess how important a word is to each text, the length of a single text is also considered. The equation of TF-IDF are as Eq.(1), (2) and (3). The reason we use TF-IDF is that the content of raw data comes from the dynamic analysis include too many irrelevant information, and we decided to extract the api function calls of each malware. These api function calls are built on lots of words, so we take advantage of TF-IDF on processing texts, turn the api function call into numerical features.

$$tf_{t,d} = \frac{n_{t,d}}{\sum_{k=1}^{T} n_{k,d}} \quad (1)$$

$$idf_t = \log\left(\frac{D}{d_t}\right) \tag{2}$$

$$score_{t,d} = tf_{t,d} \times idf_t \tag{3}$$

On the other hand, in the NIDS-based intrusion detection, we use one-hot encode for low dimensional non-numerical features, and embedding for the high dimensional ones. One-hot encode can expand different m categories in the feature to m independent two-bit features, and mark the features they represent as 1 and the rest are 0. But one-hot encode will cause the dimension disaster for high-dimensional features, so we decide to use embedding, which can map high-dimensional features into low dimension properly, by optimizing the mapping matrix. For instance, the feature names "Source Port" has 52,554 different category attributes, we use embedding method to project them into 2-dimensional space, as shown in Fig. 3.

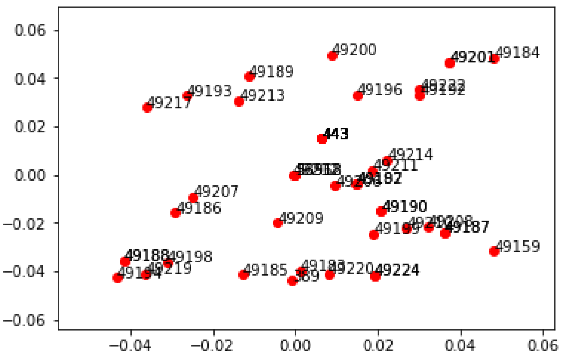


Fig.3. The result of embedding the part of attributes in Source Port

3.2. *Feature Selection*

Researchers are working hard to find a good methods to discard the redundant features in dataset, which has significant influence on model performance. The reason we use different feature selection methods for different area is that the features of dynamic-based malware analysis dataset are words, and most of NIDS-based intrusion detection dataset are numbers. If we use Random Forest for dynamic-based malware analysis dataset, it will take too much time to calculate the importance of features.

3.3. *Application of Deep Learning Model*

In the dynamic-based malware analysis, we use deep learning model such as Convolution Neural Network(CNN) [6], Multi-Layer Perceptron(MLP) [7] to classify the malware sample These model are well-known in neural network study, so we don't give unnecessary details.

In the NIDS-based intrusion detection, we take model named Sequence-to-Sequence, which was proposed by I. Sutskever et al. in 2014 [5]. The simple architecture is shown in Fig.4 The arrow between c and <EOS> is used as the boundary, left of arrow is the encoder, and the decoder for the right, both of which are composed of LSTM. The main task of the encoder is to compress the input sequence vector into a content vector v (context vector) with a much smaller dimension. This content vector is also the hidden layer output of the encoder in the last layer, which represents the model's understanding of a sequence. On the decoder, there are two input sources. The first source is the content vector from the encoder, and the second source is the delay of sequence we tempt to predict. After receiving both, the decoder begins decoding and outputs the specified sequence.

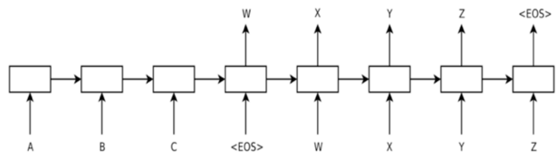


Fig.4. Architecture of Sequence to Sequence[5]

4. Resault and Discussion

We used multiple measures to get a more persuasive result. In this part, we'll introduce the evaluation metrics we used in this study.

- True Positive (TP) : malware sample that is correctly classified as malware.
- False Positive (FP) : benign sample that is incorrectly classified as malware.
- True Negative (TN) : benign sample that is correctly classified as benign.
- False Negative (FN) : malware sample that is incorrectly classified as benign.

The accuracy means the proportion of the total number of correct classifications：

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FP} \qquad (4)$$

In order to verify that the feature selection method proposed in this study does improve the model performance, we will compare four classification algorithms, both machine learning and deep learning: RF, SVM, MLP and CNN. To verify whether the high-dimensional feature embedding method and the Sequence to Sequence model has ability to improve the capability of intrusion detection, we convert high-dimensional non-numerical features of CICIDS2017 dataset: Source Port and Destination Port, into low-dimensional features by using the embedding method, and add them into training data to train the Sequence to Sequence model. As shown in Fig.5, we can obtain a better evaluation result based on embedding method and Sequence to Sequence model. The accuracy, prediction, recall and f1-score are 99.93%, 99.8%, 99.87% and 99.84, respectively, which achieve an ideal result.

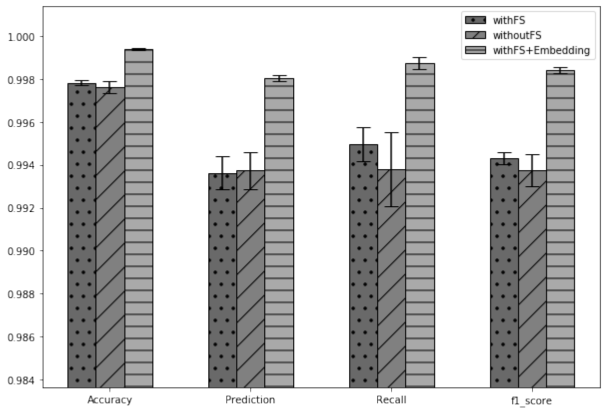


Fig.5. Comparison of Using Random Forest Feature Selection and Embedding or not

5. Conclusion

This study proposes a method of text processing as the main idea to extract, encode and adjust the weight of this feature of the Windows operating system application interface call, and then use the feature selection to drop redundant features step by step. In addition to the ability to reduce lots of features, the progressive feature selection method proposed in this paper can keep the information of the original features. We hope that these methods can be further used in real-time analyze.

Acknowledgements

This work was supported by the MOST, Taiwan under contracts numbers MOST 108-2221-E-006-110-MY3 and MOST 108-2218-E-006-035-.

References

1. D. Bilar, Opcodes as predictor for malware, *International Journal of Electronic Security and Digital Forensics*, vol. 1, no. 2, pp. 156-168, 2 5 2007. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
2. I. Santos, B. Felix and J. Nieves, Idea: Opcode-Sequence-Based Malware Detection, *International Symposium on Engineering Secure Software and Systems*, vol. 5965, pp. 35-43, 2010.
3. I. Sharafaldin, A. Lashkari and A. Ghorbani, Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization, *ICISSP*, pp. 108-116, 22-24 Jan 2018..
4. L. Breiman, Random Forests, Jan 2001. [Online]. Available: https://www.stat.berkeley.edu/~breiman/randomforest2001.pdf. [Accessed 10 Feb 2019].
5. I. Sutskever, O. Vinyals and Q. V. Le, Sequence to Sequence Learning with Neural Networks, *NIPS'14 Proceedings of the 27th International Conference on Neural Information Processing Systems*, vol. 2, pp. 3104-3112, 8-13 Dec 2014.
6. A. Krizhevsky, I. Sutskever and G. E. Hinton, ImageNet classification with deep convolutional neural networks, *International Conference on Neural Information Processing Systems*, vol. 1, pp. 1097-1105, 3-6 Dec 2012.
7. FionnMurtagh, Multilayer perceptrons for classification and regression, *Neurocomputing*, vol. 2, pp. 107-183, July 1991

Robust Control of Nonholonomic Wheeled Mobile Robot with Hybrid Controller Approach

Ho-Nien Shou
Department of Aviation & Communication Electronics, No.198, Jieshou W. Rd., Gangshan District
Kaohsiung City, 82047, Taiwan (R.O.C.)
E-mail: honien.shou@gmail.com

Abstract

This study proposes a control strategy to solve the nonholonomic mobile robot trajectory tracking problem on the basis of Cerebellar Model Articulation Controller (CMAC). Mobile robot needs two controllers to provide the control demands. One controller is mathematically described in terms of robot's kinematics; while the other is given by dynamics equations. To implement the speed control to track the reference trajectory, we apply the Lyapunov theory to obtain the virtual speed control command. On the other hand, we use cerebellum controller to approach to the non-linearity and uncertainty of the dynamics model. Furthermore, we combine the speed error to construct a torque controller, which can online real-time compensate the influences made by uncertainties. The observer is used to estimate the external disturbance, so that the controller has more ability to reject external disturbance. The convergence and stability of the system is determined by the Lyapunov stability criterion after linearizing the system. Our simulation is performed in Matlab/Simulink environment, and the results verify the effectiveness of the controller algorithm.

Keywords: nonholonomic mobile robot; Cerebellar Model Articulation Controller; Lyapunov stability criterion; disturbance observer.

1. Introduction

Many literatures which study nonholonomic mobile robots are usually focus on wheeled mobile robots (WMR). In the theoretical research of WMR's motion control, in general, only pure rolling condition is considered. In other words, it is assumed that no slip condition (including lateral and longitudinal sliding) occurs. This ideal constraint is essentially a kind of nonholonomic constraint, therefore, WMR is a typical case of the nonholonomic system. In this paper, the control of WMR is studied. According to different control objectives, the control problems of nonholonomic systems can be divided into three categories[1] [2]: position stabilization, trajectory tracking and path following. However, in the content of research report[3], the emphasis the Brockett's theorem proves that there is no asymptotic stable fixed point where pure state feedback does not exist [1]. In the report, the nonholonomic mobile robot control is divided into two categories, one for fixed point asymptotic stabilization relying on highly nonlinear techniques, and the other based on more classical linear and nonlinear techniques for asymptotic stabilization of feasible and persistently exciting trajectories.

Position stabilization control refers to the design of a feedback controller, which can actuate and stabilize the system from a given initial state to an arbitrary target state. It is also called the posture stabilization, posture regulation, and set-point regulation in some references. The position, state, posture, set-point, and so on, describe position and attitude of a mobile robot by a set of generalized coordinates. In research, the origin point usually be set as the target state.

Trajectory tracking control refers to a controller design which can command the robot arriving and tracking a specific trajectory from a given initial state in an inertial coordinate system. On the other hand, path following control means to control the robot to arrive and keep following a specific geometric path from a given initial state in an inertial coordinate system. Frankly speaking, trajectory tracking control demands the robot to track a specific time-varying trajectory, however, path following control only follows a designed path regardless of the arriving time of the specific position. Hence, we can regard the path following problem as a special case of the trajectory tracking problem since the former is much easier to be deal with compared to the later.

The organization of this paper is in the following. Section 1 introduces the historical review of nonholonomic WMR motion control, including research significance and the main difficulties. The current studies about point stabilization, trajectory tracking, and path following control of nonholonomic WMR are summarized. The research motivation and main structure of this paper are illustrated in the final paragraph. These two algorithms are used to design the online compensation system for uncertainties and the controller. An innovative control concept is proposed in Section 2, which combines kinematics control and dynamics control. The cerebellar neural network weight updating algorithm and kernel space algorithm are introduced in Section 3. The effective control results given by the robust cerebellar neural network self-adjusting trajectory tracking controller are demonstrated in Section 4. Section 5 gives the conclusions and the future work.

2. Kinematic Model and Controller Design

2.1 Model of WMR

The dynamics equations of a nonholonomic mobile robot with n-dimensional state space and subject to m constraints read [4-7]

$$M(\boldsymbol{q})\ddot{\boldsymbol{q}}(t)+C(\boldsymbol{q},\dot{\boldsymbol{q}})\dot{\boldsymbol{q}}+G(\boldsymbol{q})+F(\dot{\boldsymbol{q}})+\boldsymbol{\tau}_d=B(\boldsymbol{q})\boldsymbol{\tau}-A^T(\boldsymbol{q})\lambda \quad (1)$$

$$\dot{\boldsymbol{q}}=S(\boldsymbol{q})\boldsymbol{u} \quad (2)$$

where $M(\boldsymbol{q})\in\mathbb{R}^{n\times n}$ is a symmetric, positive definite inertia matrix, $C(\boldsymbol{q},\dot{\boldsymbol{q}})\in\mathbb{R}^{n\times n}$ is the centripetal and Coriolis matrix, $G(\boldsymbol{q})\in\mathbb{R}^{n}$ is the gravitational vector, and $F(\dot{\boldsymbol{q}})\in\mathbb{R}^{n}$ is the surface friction. To a mobile robot moving on a smooth plane, vectors $G(\boldsymbol{q})$ and $F(\dot{\boldsymbol{q}})$ are equal to zero. $\boldsymbol{\tau}_d\in\mathbb{R}^{n}$ denotes bounded unknown disturbances, $\boldsymbol{\tau}\in\mathbb{R}^{p}$ denotes the control input, $B(\boldsymbol{q})\in\mathbb{R}^{n\times p}$ denotes the input transformation matrix, $\lambda\in\mathbb{R}^{m}$ is the constraint force vector, and $A(\boldsymbol{q})\in\mathbb{R}^{m\times n}$ is the constraint matrix.

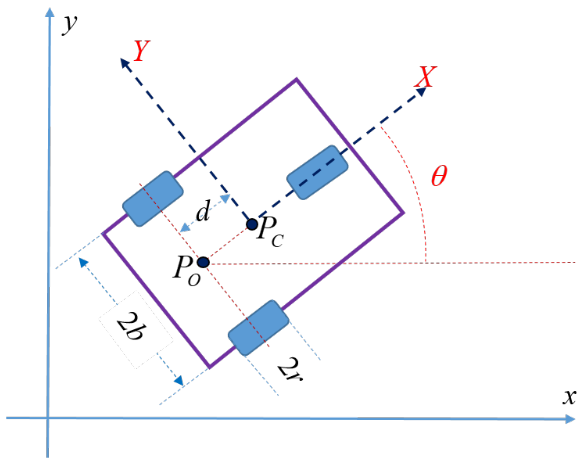


Figure 1 WMR nonholonomic mobile robot.

Figure 1 illustrates a three wheels mobile robot, in which d is the distance between robot's mass center P_C, P_O denotes the geometric center, $2b$ is the distance between two driving wheels, and r denotes the wheel radius. $\boldsymbol{q}=[x \quad y \quad \theta]^T$ represents robot's position and orientation, $\boldsymbol{u}=[v \quad \omega]^T$ represents velocity and angular velocity, $\boldsymbol{\tau}=[\tau_1 \quad \tau_2]^T$ is control torque, J denotes the inertia moment, and m_a denotes the mass of the WMR. The parameters in each matrix appeared in Eqs. (1) and (2) are

$$M(\boldsymbol{q})=\begin{bmatrix} m_a & 0 & m_a d\sin\theta \\ 0 & m_a & -m_a d\cos\theta \\ m_a d\sin\theta & -m_a d\cos\theta & J \end{bmatrix},$$

$$C(\boldsymbol{q},\dot{\boldsymbol{q}})=\begin{bmatrix} 0 & 0 & m_a d\dot{\theta}\cos\theta \\ 0 & 0 & m_a d\dot{\theta}\sin\theta \\ 0 & 0 & 0 \end{bmatrix}, \quad B(\boldsymbol{q})=\frac{1}{r}\begin{bmatrix} \cos\theta & \cos\theta \\ \sin\theta & \sin\theta \\ b & -b \end{bmatrix},$$

$$A^T(\boldsymbol{q})=\begin{bmatrix} -\sin\theta \\ \cos\theta \\ -d \end{bmatrix}, \quad S(\boldsymbol{q})=\begin{bmatrix} \cos\theta & -d\sin\theta \\ \sin\theta & d\cos\theta \\ 0 & 1 \end{bmatrix}$$

We can express nonholonomic constraints as

$$A(\boldsymbol{q})\dot{\boldsymbol{q}} = 0 \quad . \tag{3}$$

Set a *n-m* dimensional full rank matrix $S(\boldsymbol{q})$ as a base set in null space $\boldsymbol{A}(\boldsymbol{q})$ such that

$$A(\boldsymbol{q})S^T(\boldsymbol{q}) = 0 . \tag{4}$$

We can obtain an auxiliary velocity control input $\boldsymbol{u} \in \mathbb{R}^{n-m}$ from Eqs. (2) and (4),

$$\dot{\boldsymbol{q}} = S(\boldsymbol{q})\boldsymbol{u} . \tag{5}$$

From Eq. (5) we can have

$$\ddot{\boldsymbol{q}} = S(\boldsymbol{q})\dot{\boldsymbol{u}} + S(\dot{\boldsymbol{q}})\boldsymbol{u} \quad . \tag{6}$$

By inserting Eq. (6) into Eq. (1) and multiplying by $S^T(\boldsymbol{q})$ to cancel the constraint matrix $A^T\lambda$, we obtain the dynamic equation of the nonholonomic mobile robot:

$$\boldsymbol{S}^T M\boldsymbol{S}\dot{\boldsymbol{u}} + \boldsymbol{S}^T\left(M\dot{\boldsymbol{S}} + C\boldsymbol{S}\right)\boldsymbol{u} + \boldsymbol{S}^T\boldsymbol{\tau}_d = \boldsymbol{S}^T B\boldsymbol{\tau} \tag{7}$$

After variable replacements, Eq. (7) becomes

$$\bar{M}(\boldsymbol{q})\dot{\boldsymbol{u}} + \bar{C}(\boldsymbol{q},\dot{\boldsymbol{q}})\boldsymbol{u} + \bar{\boldsymbol{\tau}}_d = \bar{\boldsymbol{\tau}} \tag{8}$$

where $\bar{M}(\boldsymbol{q}) \in \mathbb{R}^{p\times p}$ is the symmetric positive definite inertia matrix, $\bar{C}(\boldsymbol{q},\dot{\boldsymbol{q}}) \in \mathbb{R}^{p\times p}$ is matrix for centripetal and Coriolis forces, $\boldsymbol{S}^T\boldsymbol{\tau}_d = \bar{\boldsymbol{\tau}}_d \in \mathbb{R}^p$ contains unstructured unmodeled dynamic bounded unknown perturbations, and $\boldsymbol{S}^T\boldsymbol{B\tau} = \bar{\boldsymbol{\tau}} \in \mathbb{R}^p$ is control input matrix ; in which $\bar{M} = \begin{bmatrix} m_a & 0 \\ 0 & J - m_a d^2 \end{bmatrix}$, $\bar{C} = 0_{2\times 2}$, $\bar{\boldsymbol{\tau}}_d = S^T(\boldsymbol{q})\boldsymbol{\tau}_d$, and $\bar{\boldsymbol{\tau}} = S^T(\boldsymbol{q})B(\boldsymbol{q})$.

2.2 Kinematic Controller Design

$\boldsymbol{q}_r = \begin{bmatrix} x_r & y_r & \theta_r \end{bmatrix}^T$ is the reference orientation of the WMR

$$\begin{bmatrix} \dot{x}_r \\ \dot{y}_r \\ \dot{\theta}_r \end{bmatrix} = \begin{bmatrix} \cos\theta_r & -d\sin\theta_r \\ \sin\theta_r & d\cos\theta_r \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_r \\ \omega_r \end{bmatrix} \tag{9}$$

where x_r , y_r , θ_r , v_r and ω_r are the expectations of x, y ,θ, v and ω.

A. Kinematics Controller Design

Let us define tracking error as

$$\begin{bmatrix} x_e \\ y_e \\ \theta_e \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_r - x \\ y_r - y \\ \theta_r - \theta \end{bmatrix} = T_e \begin{bmatrix} x_r - x \\ y_r - y \\ \theta_r - \theta \end{bmatrix} \tag{10}$$

where $$T_e = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{11}$$

With the differentiation of Eq. (10), we can have the attitude error with respect to time expressed as:

$$\begin{bmatrix} \dot{x}_e \\ \dot{y}_e \\ \dot{\theta}_e \end{bmatrix} = \begin{bmatrix} v_r\cos\theta_e \\ v_r\sin\theta_e \\ \omega_r \end{bmatrix} + \begin{bmatrix} -1 & y_e \\ 0 & -x_e \\ 0 & -1 \end{bmatrix} \begin{bmatrix} v \\ \omega \end{bmatrix} \tag{12}$$

From the above analysis, we can define the trajectory tracking of the mobile robot kinematics model as finding the bounded input v and ω such that for any initial error, system (12) is bounded by the control input $(x_e, y_e, \theta_e)^T$ and satisfies the condition $\lim_{t\to\infty}\left\|(x_e, y_e, \theta_e)^T\right\| = 0$.

Lemma 1: To any $\varphi \in \mathbb{R}$ and $\|\varphi\| \in \mathbb{R}$, the condition $f(\varphi) = \varphi\sin(arc\tan(\varphi)) \geq 0$ is true, if and only if $\varphi = 0$.

B. $\bar{\tau}_d$ Disturbance Observer Design

There are two main factors affecting the robustness of the WMR system. One is the interference from the outside world; while the other is the uncertainty of the internal parameters of the whole system. In fact, the WMR will be affected by friction and various noise such that there are some differences compared to the ideal system. The disturbance observer is designed to estimate the external disturbance, $\bar{\tau}_d$. The external disturbance estimator designed for the WMR system is usually in the form of [3,6,7]:

$$\begin{cases} \hat{\bar{\boldsymbol{\tau}}}_d = z + \boldsymbol{L}\boldsymbol{u} \\ \dot{z} = \boldsymbol{L}\bar{M}^{-1}z - \boldsymbol{L}\left(-\bar{M}^{-1}\boldsymbol{L}\boldsymbol{u} - \bar{M}^{-1}\bar{C}\boldsymbol{u} + \bar{M}^{-1}\bar{\boldsymbol{\tau}}\right) \end{cases} \tag{13}$$

where $\hat{\bar{\boldsymbol{\tau}}}_d$ denotes the estimation of the unknown external disturbance $\bar{\boldsymbol{\tau}}_d$, z is the internal state of the nonlinear estimator, $\boldsymbol{L}$ is the parameter of the nonlinear estimator needed to be solved, which is usually expressed in terms of the constant matrix.

Some assumptions are needed to control the WMR to track the target trajectory. Let us state these assumptions in the following.

Assumption 1: The disturbances $\|\bar{\boldsymbol{\tau}}_d\|$ and $\|\dot{\bar{\boldsymbol{\tau}}}_d\|$ are bounded, and $\bar{\boldsymbol{\tau}}_d$ is constant in the steady state, i.e., $\lim_{t\to\infty}\|\dot{\bar{\boldsymbol{\tau}}}_d\| = 0$.

Assumption 2: The first derivative of the reference linear velocity v_r and angular velocity ω_r are bounded and $v_r > 0$.

Lemma 2: $L\bar{M}^{-1}$ is the Hurwitz matrix, and $\tilde{\bar{\boldsymbol{\tau}}}_d$ is asymptote convergence, then we express $\bar{\boldsymbol{\tau}}_d = \hat{\bar{\boldsymbol{\tau}}}_d$.

C. The Output Layer Setting of the Cerebellar Model Controller

Consider Eqs. (12) and (14), and introduce a new control variable, $\bar{\boldsymbol{u}} = \boldsymbol{u}_c - \boldsymbol{u}$

$$\begin{aligned}\bar{M}\dot{\bar{\boldsymbol{u}}} &= -\bar{C}(\boldsymbol{q},\dot{\boldsymbol{q}})\bar{\boldsymbol{u}} + \left(\bar{M}\dot{\boldsymbol{u}}_c + \bar{C}(\boldsymbol{q},\dot{\boldsymbol{q}})\boldsymbol{u}_c\right) + \bar{\boldsymbol{\tau}}_d - \bar{\boldsymbol{\tau}} \\ &= -\bar{C}(\boldsymbol{q},\dot{\boldsymbol{q}})\bar{\boldsymbol{u}} + \Gamma(\boldsymbol{u}_c,\dot{\boldsymbol{u}}_c) + \bar{\boldsymbol{\tau}}_d - \bar{\boldsymbol{\tau}} \qquad (14)\end{aligned}$$

where $\boldsymbol{\Gamma}(\boldsymbol{u}_c,\dot{\boldsymbol{u}}_c) = \bar{M}\dot{\boldsymbol{u}}_c + \bar{C}(\boldsymbol{q},\dot{\boldsymbol{q}})\boldsymbol{u}_c$ is the nonlinear function of the mobile robot, function $\boldsymbol{\Gamma}(\boldsymbol{u}_c,\dot{\boldsymbol{u}}_c)$ includes many parameters of the mobile robot such as mass, rotation inertia and so on. It is very difficult to determine these parameters, hence we use cerebellar model controller to approach, which can be expressed as

$$\boldsymbol{\Gamma}(\boldsymbol{u}_c,\dot{\boldsymbol{u}}_c) = Y(\boldsymbol{u}_c,\dot{\boldsymbol{u}}_c)\boldsymbol{W} \qquad (15)$$

where $Y(\boldsymbol{u}_c,\dot{\boldsymbol{u}}_c)$ denotes the cerebellar network output, and $\boldsymbol{W}$ is the connection weighting of the cerebellar network output. Then we can rewrite Eq. (14) as

$$\bar{M}\dot{\bar{\boldsymbol{u}}} = -\bar{C}\bar{\boldsymbol{u}} + YW + \bar{\boldsymbol{\tau}}_d - \bar{\boldsymbol{\tau}} \qquad (16)$$

3. Simulation Results and Analysis

In order to verify the actual control performance provided by the proposed algorithm, we use MATLAB / SIMULINK to execute the simulation. Parameters of the WMR are list as follows: the wheel radius, $r = 0.12\ m$, the distance between two driving wheels, $2b = 0.6\ m$, the mass of the WMR, $m_a = 4\ kg$, the distance between robot's mass center and geometric center, $d = 0.25\ m$ and the inertial moment, $J = 0.25\ kgm^2$. In the simulation, the parameters of the observer and controller are designed as $\boldsymbol{L} = \begin{bmatrix} L_1 & 0 \\ 0 & L_2 \end{bmatrix}$, where $L_1 = L_2 = 12m$, and $k_x = k_y = k_\theta = 1$, $k_u = 0.3$ and $k_s = 0.8$. The initial position of the reference input is $[x_r(0) \quad y_r(0) \quad \theta_r(0)] = [0 \quad 0 \quad 0]$. The reference velocities are set as $v_r = 0.1\ m/s$, $\omega_r = 0.01\ rad/s$. The initial state of the WMR is given as $[x(0) \quad y(0) \quad \theta(0)] = [39.5m \quad 0.1m \quad 89.427\deg.]$ .The external disturbance $\bar{\boldsymbol{\tau}} = \begin{bmatrix} 1.2\sin(0.3t - \frac{\pi}{2}) & 0.51\sin(0.3t + \frac{\pi}{2}) & -1.31\sin(0.5t) \end{bmatrix}^T$ occurs after the WMR is moving. Learning rate of CMAC controller is $\boldsymbol{\eta} = 0.5$, and the inertia coefficient is $\boldsymbol{\alpha} = 0.5$.

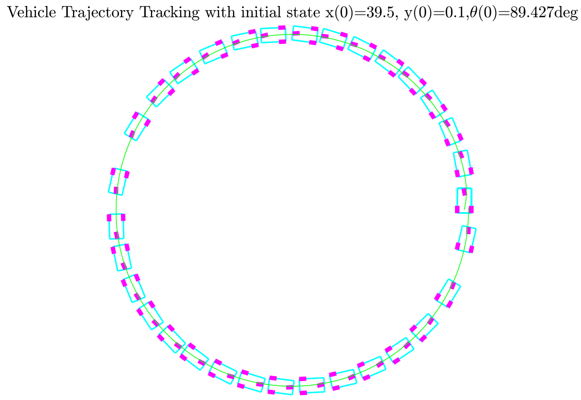


Figure 1 WRM trajectory tracking

References

1. Brocket, R. W. *Asymptotic Stability and Feedback Stabilization*, In: Brockett, R. W.; Millman, R. S.; Sussmann, H. J. eds., Differential Geometric Control Theory, Birkhauser, Boston, 1983, pp. 181-208.
2. Alessandro De Luca, Giuseppe Oriolo, Marilena Vendittelli, *Control of Wheeled Mobile Robots: An Experimental Overview*, RAMSETE, Berlin: Springer, 2001, 181-226.
3. Morin P.; Samson, C. *Trajectory tracking for nonholonomic systems*. Theoretical background and applications, 2008.
4. Campion, G. Structural Properties and Classification of Kinematic and Dynamic Models of Wheeled Mobile Robot. *IEEE Transactions on Robot and Automation*. 1996, 12(1), 47-61.
5. Fierro R.; Lewis F. L. Control of a nonholonomic mobile robot using neural networks. *IEEE Transactions on Neural Networks.* 1998,9(4), 589-600.

6. Fierro R., Lewis, F. Control of a nonholonomic mobile robot: backstepping kinematics into dynamics. *Journal of Robotic Systems*. 1997, 14(3), 149-163.
7. Kukao T.; Nakagawa H.; Adachi N. Adaptive tracking control of nonholonomic mobile robot. *IEEE Trans Robot Automat*. 2000, 16(5), 609-615.

Attitude control of microsatellite using PWM method: verification using PIL

Ho-Nien Shou
Department of Aviation & Communication Electronics, No.198, Jieshou W. Rd., Gangshan District
Kaohsiung City, 82047, Taiwan (R.O.C.)
E-mail: honien.shou@gmail.com

Abstract

The missions of micro-satellites range from early military usages to weather forecast, resources exploration, communication, and scientific experiment. The advantages of micro-satellites are simple mechanical structures, great reliability, low prices, and precise researchers and equipment. Therefore, due to the limits of weight and power, this kind of micro-satellite with low design cost and high precision requires reducing unnecessary attitude sensors and controllers to ensure the precision of attitude control. Besides, it adapts the least hardware constituting attitude system, which is the future trend in the satellite engineering. In this article, an estimating technology about measuring angle speed with a gyro less is mentioned. This technology is based on the period change which the earth magnetic field gets along the track. Only by using a three axis magnetometer, it can produce the data from the micro-satellite measuring the earth magnetic field. Besides, the measures of three axis angle speed and attitude angle can be gotten through Kalman filtering. The purpose of this article is mainly to explore the problem on attitude detumbling control of micro-satellites departing away a carrier to enter a track. It is realized by a thruster to proceed the confinement of the satellite moving, the attitude stable control and processor-in-the-loop (PIL).

Keywords: three axis magnetometer, earth magnetic field, Kalman filtering, processor-in-the-loop.

1. Introduction

Before the satellite departs away the carrier into the mission track, it will tumble with a high speed rotating attitude. Or before the initial rotating speed exceeds the designed tolerant value, the function of the attitude control system lies to confine tumble and lower the angle speed for the satellite to achieve the pointed location. If the attitude control system uses a thruster to correct the attitude, the advantage is that it is beyond the limits of outer circumstances and easy to operate. In addition, it is suitable for any controlling methods, fast to correct time, and precise to point. 1) The uncertainties in the interior satellite are there is an error in the setting distance of the spray head, the inclined side of the spray head is not accurate, and the usage of gas fuel makes the quality center deflect. 2) The exterior disturbances of the satellite are the height and position to the environmental factors: gravity gradient, air resistance and magnetic forces. 3) Using the spray head easily produces bang-bang control. It is impossible as a linear switch to have high fuel efficiency, which can use the pulse-width pulse-frequency modulation to solve the thruster problem at the same time[1][2].

As to a micro-satellite with lower precision to three axis point, lighter carrier, low power consumption and limited research charge, there is undoubtedly a good inducement to merely use the magnetometer to complete the three axis attitude and the speed of the attitude angle. Therefore, it is practically important to develop this engineering. In the past, the magnetometer is often used as an assistance to meet with other parts in the satellite attitude confirming system. It is less to use a single magnetometer to complete attitude confirmation. The main reason is that the attitude precision is limited by using this method.

It cannot get rid of the error limitation from incorrect terrestrial magnetism. This common method is to use the star pattern to get the track parameter and then pass through International Geomagnetic Reference Field (IGRF) to find our local geomagnetic field[3], which compares with the magnetometer measuring value to get the attitude angle by the attitude algorithm. However, by doing so, it only can get two axis attitude angles, but it cannot provide the message of the angle efficiency. To get satellite upper angle efficiency, it is necessary to use a gyroscope. However, the function of the gyroscope will be worsened with time, and its price is high and the equipment is complicated. Therefore, it may be a good choice to adapt attitude angle sensor instead of a gyroscope. In the article, the formula of micro-satellite and exercise is calculated by quaternion and angle speed. In it, an unscented Kalman filtering is adopted. The purpose of the article focuses on (1) Designing the thruster position, (2) Taking the satellite attitude angle speed and four-element position as an order of thruster, (3) Considering when there is an error in the distance of the thruster position, the uncertainty in the interior satellite is the quality center change by using gas fuels, (4) Exterior disturbances are: satellite height and position to magnetic force, gravity gradient, and air resistance, (5) Using the thruster with the pulse-width pulse-frequency modulation to solve the thruster problem. (6) Describing the above questions by the leaning differential inequality [4] [5], which means to reduce the exterior disturbances and limit the system parameter uncertainty of the satellite attitude controller.

The contents of the article are as follows: section II problem description, the Thrust force derived of satellite in section III, section IV Pulse-width Pulse-frequency Modulation Rules, section V Kalman Filtering and Different Calculating Methods, section VI Fulfillment of Software-in-the-Loop (SIL) and processor-in-the-Loop (PIL), and future prospects and conclusions in the last section.

2. Problem Description

Considering satellite attitude exercise formula

$$\dot{\boldsymbol{q}} = \frac{1}{2}\begin{bmatrix} -\boldsymbol{\omega}_o^{b\times} & \boldsymbol{\omega}_o^{b} \\ -\boldsymbol{\omega}_o^{bT} & 0 \end{bmatrix}\begin{bmatrix} \mathbf{q} \\ q_4 \end{bmatrix} \quad (1)$$

Satellite non-linear dynamics formula

$$\left(J + \Delta J\right)\boldsymbol{\omega}_i^{b} = -\boldsymbol{\omega}_i^{b\times}\left(J + \Delta J\right)\boldsymbol{\omega}_i^{b} + \boldsymbol{\tau}_w + \boldsymbol{\tau}_u \quad (2)$$

(i) $\boldsymbol{q} = \begin{bmatrix} \mathbf{q}^T & q_4 \end{bmatrix}^T = \begin{bmatrix} q_1 & q_2 & q_3 & q_4 \end{bmatrix}^T = \begin{bmatrix} \bar{e}\sin\dfrac{\phi}{2} & \cos\dfrac{\phi}{2} \end{bmatrix}^T$

The above is a quaternion, which is to rotate an *Euler angle* by order to a new attitude.

Suppose the specific condition is $\bar{e} = \begin{bmatrix} e_x & e_y & e_z \end{bmatrix}^T$, and the condition rotating angle is ϕ.

(ii) $\boldsymbol{\omega}_o^{b}$ is the angular velocity of the body frame with respect to the orbit frame. Among it,

$\boldsymbol{\omega}_o^{b} = \begin{bmatrix} \omega_x & \omega_y & \omega_z \end{bmatrix}^T$, $\boldsymbol{\omega}_o^{b\times} = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix}$ is a corresponding $\boldsymbol{\omega}_o^{b}$ opposing skew matrix.

(iii) $\boldsymbol{\omega}_i^{b}$ is the angular velocity of the body frame with respect to the inertial frame.

(iv) $J+\Delta J$ is inertial matrix,

$$J + \Delta J = \begin{bmatrix} J_{xx} & J_{xy} & J_{xz} \\ J_{xy} & J_{yy} & J_{yz} \\ J_{xz} & J_{yz} & J_{zz} \end{bmatrix} + \begin{bmatrix} \Delta J_{xx} & \Delta J_{xy} & \Delta J_{xz} \\ \Delta J_{xy} & \Delta J_{yy} & \Delta J_{yz} \\ \Delta J_{xz} & \Delta J_{yz} & \Delta J_{zz} \end{bmatrix} \quad (3)$$

In it, J_{ii} is i axle rotates inertial, and ΔJ_{ij} is i axle rotates inertial along j axle. When the fuels are used, they will the quality center in the satellite system, which are uncertain items in the satellite system parameter.

(v) $\boldsymbol{\tau}_w$ represents that the satellite will be disturbed by the force, including gravity gradient force, air resistance force, and magnetic disturbed force.

(vi) $\boldsymbol{\tau}_u$ represents satellite attitude controlled force. In the paper, the control theory is conducted to get the relationship $\boldsymbol{\tau}_u$ thruster matrix and output $\boldsymbol{u}$, and then is combined pulse-width modulation and the jet to produce thrust.

3. Production of Satellite Thrust

The satellite exercise frame is as figure 1. Four thrusters are put $x-y$ plane, including λangle, the jet moves toward $-z$ direction, and lines to be a square distance d. Fuel is put the jet $x-y$ plane l distance, and $+z$ direction is toward satellite center, in figure 2. $\Delta\lambda$、Δd and Δl respectively represent the jet including angle, jet relative distance, and the uncertainty of the fuel gravity position. The following is production $\boldsymbol{\tau}_u$ from the thruster:

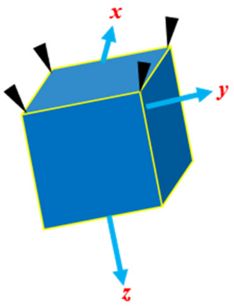


Figure 1 Satellite frame

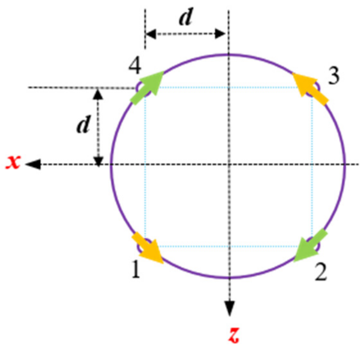


Figure 2 Satellite Jet Geometry Frame

$$\boldsymbol{\tau_u} = \begin{bmatrix} \tau_x \\ \tau_y \\ \tau_z \end{bmatrix} = \begin{bmatrix} \vec{r}_1 \times \vec{F}_1 & \vec{r}_2 \times \vec{F}_2 & \vec{r}_3 \times \vec{F}_3 & \vec{r}_4 \times \vec{F}_4 \end{bmatrix} \boldsymbol{u}$$

$$= (\boldsymbol{B} + \Delta \boldsymbol{B}) \boldsymbol{u} \qquad (4)$$

$$\vec{r} = \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} -(d+\Delta d) & -(d+\Delta d) & (d+\Delta d) & (d+\Delta d) \\ -(d+\Delta d) & (d+\Delta d) & (d+\Delta d) & -(d+\Delta d) \\ -(l+\Delta l) & -(l+\Delta l) & -(l+\Delta l) & -(l+\Delta l) \end{bmatrix} \end{array} \begin{array}{c} \\ x \\ y \\ z \end{array}$$

$$\vec{F} = \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} -\rho \sin(\lambda+\Delta\lambda) & -\rho \sin(\lambda+\Delta\lambda) & \rho \sin(\lambda+\Delta\lambda) & \rho \sin(\lambda+\Delta\lambda) \\ \rho \sin(\lambda+\Delta\lambda) & -\rho \sin(\lambda+\Delta\lambda) & -\rho \sin(\lambda+\Delta\lambda) & \rho \sin(\lambda+\Delta\lambda) \\ -\cos(\lambda+\Delta\lambda) & -\cos(\lambda+\Delta\lambda) & -\cos(\lambda+\Delta\lambda) & -\cos(\lambda+\Delta\lambda) \end{bmatrix} \end{array} \begin{array}{c} \\ x \\ y \\ z \end{array}$$

$$\alpha = \frac{1}{\sqrt{2}}, \ \beta_1 = d + \alpha\lambda l, \ \beta_2 = d - \alpha\lambda l, \ \beta_3 = \frac{\lambda d}{\alpha},$$

$$\Delta\beta_1 = \Delta d - \lambda\Delta\lambda(d+\Delta d) + \alpha(\lambda\Delta l + l\Delta\lambda + \Delta\lambda\Delta l)$$

$$\Delta\beta_2 = \Delta d - \lambda\Delta\lambda(d+\Delta d) - \alpha(\lambda\Delta l + l\Delta\lambda + \Delta\lambda\Delta l)$$

$$\Delta\beta_3 = \frac{1}{\alpha}(\lambda\Delta d + d\Delta d + \Delta\lambda\Delta d) \qquad (5)$$

When the thruster transfers matrix $\boldsymbol{B}$, the features are as follows[6].

1) $rank(\boldsymbol{B} + \Delta\boldsymbol{B}) = \min\{m, n\}$ (6)

2) $(\boldsymbol{B} + \Delta\boldsymbol{B})(\boldsymbol{B} + \Delta\boldsymbol{B})^T$

$$= \begin{bmatrix} 4(\beta_1 + \Delta\beta_1) & 0 & 0 \\ 0 & 4(\beta_2 + \Delta\beta_2) & 0 \\ 0 & 0 & 4(\beta_3 + \Delta\beta_3) \end{bmatrix} \qquad (7)$$

4. Pulse-width Modulation Rules

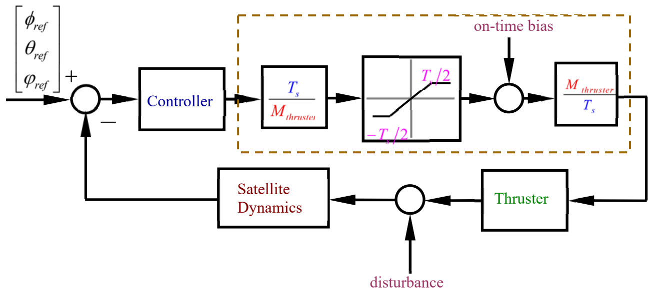


Figure 3 Micro-satellite Attitude Control Pulse-Width Modulation Thruster Block Diagram

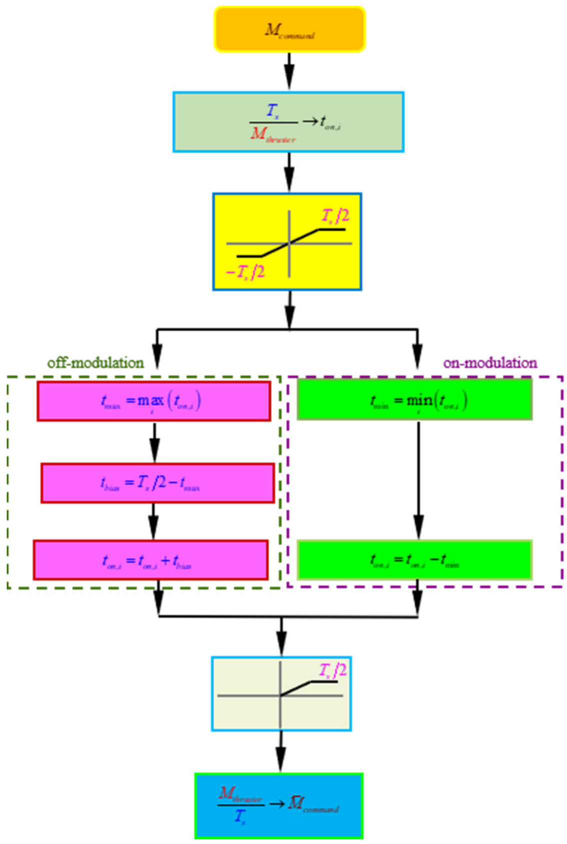


Figure 4 Pulse-width on/off modulation Control Process Chart

5. Running Heads

Because the software loop cannot be completely tested in attitude measuring elements and the thruster, the Processor-in-the-Loop in the second stage is needed [6]. In figure 5, the structure of the Processor-in-the-Loop.

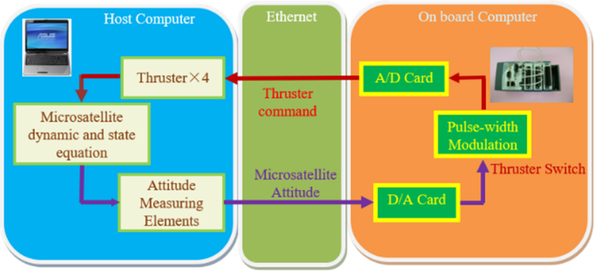


Figure 5 Processor-in-the-Loop system block diagram

TABLE I. PARAMETERS OF MICROSATELLITE

Item	Uncertainty Bound	Value	Unit
λ	$\Delta\lambda/\lambda = \pm 0.1$	5	deg
l	$\Delta l/l = \pm 0.1$	0.24	m
d	$\Delta d/d = \pm 0.1$	0.25	m
J_{xx}	$\Delta J_{xx}/J_{xx} = \pm 0.1$	5.5384	kg-m^2
J_{yy}	$\Delta J_{yy}/J_{yy} = \pm 0.1$	5.6001	kg-m^2
J_{zz}	$\Delta J_{zz}/J_{zz} = \pm 0.1$	4.2382	kg-m^2

TABLE II. INITIAL CONDITIONS OF MICROSATELLITE

Item	Description	Value	Unit
$[\varphi,\theta,\psi]$	Initial Euler Angle	[-120 30 60]	deg
$[q_1, q_2, q_3, q_4]$	Initial Quaternion	[0.0474 0.5303 -0.6597 0.5303]	
$[\omega_x,\omega_y,\omega_z]$	Initial Angular Rate	[3.5 -3.5 5.23]	rad/s
dt	Step Time	1.0	sec
T_{sim}	Simulation Time	1200	sec

The content of the experiment, will use the satellite parameters of Table I and Table II the initial conditions to simulate the satellite attitude response.

Acknowledgements

This work was supported by the National Science Council of Republic of China under contract NSC 107-2221-E-344-002 and the construction technological advice, assistance, and financial support from National Space Center research project NSPO-S-099135.

References

1. J. Yuan , D. Yang, and H. Wei, “ Flexible Satellite Attitude Maneuver Control Using Pulse-width Pulse-frequency Modulated Input Shaper,” International Symposium on Systems and Control in Aerospace and Astronautics, pp.1407-1412, 2006.
2. B. N. Agrawal, R. S. Mcclelland and G. Song, “Attitude Control of Flexible Spacecraft Using Pulse-Width Pulse-Frequency Modulated Thrusters,” Space Technolog, vol. 17, No. 1, pp.15-34, 1997.
3. H. Weidong, Z. Yulin, “Rate Damping Control for Small Satellite Using Thruster,” Acta Astronautica, vol.55, pp.9-13, 2004.
4. [4] Ho-Nien, Shou and J, C, Juang, “An LMI-Based Nonlinear Attitude Control Approach,” IEEE Transactions on Control Systems Technology, vol.11, no. 1, pp.73-83, 2003.
5. Ho-Nien, Shou and J, C, Juang, Y. W. Jan and C. T. Lin, “Quaternion Feedback Attitude Control Design: A Nonlinear Approach,” Asian Journal of Control, vol. 5, no. 3, pp. 406-411, 2003.
6. Ho-Nien Shou, “A Two-Phase Optimization Algorithm in Controller Synthesis,” Intelligent Engineering Systems through Artificial Neural Networks, Issue 5, pp.1~12, 2011.

A Study of Applying Computer-assisted Language Learning to English Course for Junior College Students in Taiwan

Shu-Hua Huang
General Education Center, Air Force Institute of Technology,
Kaohsiung City, 82047, Taiwan

I-Hsien Liu
Department of Electrical Engineering / Institute of Computer and Communication Engineering,
National Cheng Kung University, Tainan City, 70101, Taiwan
E-mail: helloferny2016@gmail.com, ihliu@cans.ee.ncku.edu.tw

Abstract

This study aims to explore the impact of the integration of Computer-assisted language learning (Call) into traditional English course on the learning effectiveness at Junior college in Taiwan. To achieve the above purpose, this study adopts the pre-experimental design with one-group pretest-posttest design. This study adopts quantitative for data analysis. The results indicated the application of Computer-assisted English learning has positive and significant progress on students' English learning effectiveness and learning motivation.

Keywords: Blended Learning, Learning Effectiveness, English course, Computer-assisted language learning (Call).

1. Introduction

Blended learning integrates the advantages of traditional face-to-face teaching and online learning, enters the mainstream course. How to combine face-to-face course and online course to enhance students' learning effectiveness? The purpose of this study will discuss what may be the suitable balance of activities in English blended learning settings. Therefore, this study aims to explore the impact of the integration of Computer-assisted language learning system into English teaching resources on the learning effectiveness at Junior college in Taiwan.

2. Literature Review

Learning in this century should not be constrained to the fence of a classroom, nor does it simply begin and end at an arranged hour. The development of information technology and the Internet, online learning has become a new emerging learning channel. The online learning has provided as alternative forms of education that has brought traditional in-class learning platforms offered in education out of the four walls of the classroom into the realm of cyberspace(Fedynich,2014). But the anecdotal evidence indicates that blended course instruction both offers more choices for content delivery and may be more effective than courses that are either fully online or fully classroom-based (Singh, 2003). According to Garnham and Kaleta (2002), blended learning is simply those in which a significant portion of the learning activities have been moved online, and time traditionally spent in the classroom is reduced but not eliminated.

The goal of blended courses is to combine the best features of in-class teaching with the best features of online learning to promote active independent learning and reduce class seat time (Garnham & Kaleta, 2002). "Blended courses offer the convenience and flexibility of wholly online courses without the loss of faculty or student interaction" (Sitter, Carter, C., Mahan, Massello, & Carter, T., 2009, p. 42).

While many of disciplines already offer online coursework, adding blended courses has the potential to meet the diverse learning needs of students and maximize available campus resources, especially in English learning. There are so many online English learning courses on the Web site. They can be used in Computer-assisted English learning system to motivate the students through interactive learning, and enhance student's learning effectiveness by immediate online learning feedback.

3. Research Method

3.1. Research Design

This study adopted the pre-experimental design with one-group pretest-posttest design.

3.2. Subject

There were one lecturer and 22 Junior college low achievement students as the subjects.

3.3. Research Materials

There was a step-by-step curriculum layout, which guided the students to arrange their own study schedule based on their own English proficiency.

3.4. Research Steps

The students would take the pretest in the beginning. Then the students experienced a 120 minutes per-week traditional face-to-face English courses for 18 weeks and finally took the posttest. The students had to report their learning journals to show their own progress every week. Finally, the posttest and a learning feedback questionnaire was conducted at the end of semester.

3.5. Data Analysis

The data were collected from pretest and posttest of English, the selection and compilation of Computer-assisted English learning resources, students' online learning records, the percentage of progress. This study adopted quantitative methods for data analysis. The SPSS statistical software package was used for quantitative analysis. First, descriptive statistics would be computed. Next, correlation and dependent sample t-test were calculated.

4. Results

4.1. Pretest and Posttest

The results indicated that pretest score on the average (M=45.09, SD=12.463) was lower than the posttest score (M=67.18, SD=12.765). The posttest was higher than the pretest score. The correlation between the pretest and posttest of English score was 0.545, ($p=0.009< 0.05$). It means the pretest and posttest of English score were significant related. Table1 shows the mean between the pretest and posttest of English score was -22.091, t =-8.539 (df =21, $p=0.000< 0.05$). There were significant differences between the pre-test and posttest English scores, and the posttest was significantly higher than the pre-test.

Table1 paired sample T-test

		Paired Differences					t	Df	Sig(2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
pair1	Pretest Posttest	-22.091	12.134	2.587	-27.471	-16.711	-8.539***	21	.000

4.2. the Percentage of Progress and Time Spent

Table2 correlation for percentage of progress and student time spent in online course

		Percentage of progress	Time spent in online course
Percentage of progress	Pearson Correlation	1	.439*
	Sig (2-tailed)		.041
	N	22	22
Student time spent in online course	Pearson Correlation	.439*	1
	Sig (2-tailed)	.041	
	N	22	22

*. Correlation is significant at the 0.05 level(2-tailed)

Table2 showed the correlation between the percentage of progress and the time students automatically spent on the Computer-assisted English learning course was 0.439 ($p< 0.05$). There was a statistically significant difference for actual time spent and for student learning, as measured by the percentage of progress of English score. It means when the students automatically spent more time on the Computer-assisted English learning course, the more percentage of progress they made.

5. Discussion

The results showed that the blended learning improved students' English proficiency. The average posttest was significant higher than the pretest. And the more time students spent on the Computer-assisted English learning course, the more progress they made. The application of Computer-assisted English learning system into traditional face-to-face course was positive and significantly effective. That is, the blended learning mode was efficient in English Learning.

5.1. Student

Garnham and Kaleta (2002) found that the principle reason that 80% of the students gave high level of satisfaction was the time flexibility provided by the blended mode. Time flexibility is defined as the ability to control the pace of one's learning, the convenience of scheduling coursework and a decrease in time spent in commuting. In this experiment, 81% students strongly agreed that one of the main benefits of blended learning was that they had more flexibility in their study time and they could arrange the pace of their own learning. An important aspect of blended learning is that students should be able to learn independently and spend more time doing self-directed learning.

Blended learning increases the opportunities for self-directed learning and develops project and time management skills (Spilka, 2002). Participating in a blended course requires students to be self-directed learners with effective time management skills. From the students' learning journals, many students reported that blended learning enhanced them in self-directed learning and improve time management skills.

But it is not generally common for Asian students. Prangpatanpon (1996) reported there is a lack of self-directed learning activities among Thai students, because they are used to authoritarian practice, and are willing to accept what their teachers said without questioning. It is different from the result of Spilka's experiment. In this case, Taiwanese students hope to have more flexible time and arrange their own learning progress.

However, a change from a face-to-face classroom to student-centered active learning can constitute a radical change for some students. If students who wish to take a traditional face-to-face study take a blended learning course, this form of learning may be disappointing and frustrating. Especially those students who are not used to being responsible for their own studies. So the student' characteristics will influence their acceptance of blended learning mode, and their learning effectiveness. It may be a new direction for future research.

5.2. Teacher

In order to teach a successful blended course, the teacher must invest a significant amount of time and effort into the redesign of the class (Garnham and Kaleta, 2002). The study found the blended instruction was a greater workload than face-to-face instruction. The lecturer indicated that she spent much more time on blended instruction and on individual communication than she had in any face-to-face instruction.

In order for true blending learning to occur, the structure of the course must be carefully evaluated to determine which instructional objectives can best be met in an online environment and which are better suited to a traditional classroom environment (Lloyd-Smith, 2010).

The blended learning needs to combine the face-to-face course and the Computer-assisted course to enhance students' learning effectiveness. The instructor needs to redesign course, and to think about what may be the appropriate balance of activities in English learning settings. It is important to take into consideration of activities which objectives can be met via Computer-assisted learning and those which will be enhanced in a traditional face-to-face format. The redesign of the curriculum also requires review of the teaching strategies and assessment models as well as the limitations of existing curriculum management techniques and systems.

The Computer-assisted English learning systems would be used to enhance students' vocabulary, pronunciation, and communicative skills. Using the Computer-assisted English learning system, the instructor had the flexibility to arrange the curriculum, exam type and difficulty levels, as shown in Figure 1. She could create exams from the test bank of the website. Through managing exam types, time and frequency, she could fully understand the students' learning status.

Fig. 1. examination management on the Web site

The students could arrange their own study schedules based on their own English proficiency and tested themselves when they were ready. The instructor could clearly know the grades, percentage of progress, correct rate, spending time, and the learning schedule of each student from the Web site, as depicted in Fig.2. The students could view their own progress and grades anytime. Besides, the instructor could really grasp the extent of the progress of the student's grades, and then modified the content and progress of the teaching. When some students had problems in certain area, the instructor could help them individually.

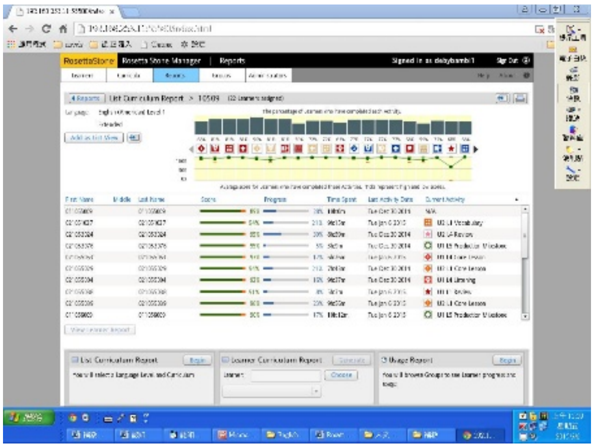


Fig. 2. the learning schedule and spending time of each student.

The instructor also indicated that she always worked alone because other colleagues were seemingly frightened by the new technology and refused to apply Computer-assisted learning to the courses. She coaxed into the teaching training, built the course and learned to use what had become one of her best teaching tools.

6. Conclusion

Entering Education 4.0, learning has become highly customized and individualized. The Computer-assisted learning system used interactive approach as instruction method to enhance students' learning effectiveness. Through the hyperlink and AI calculus analysis from the online education platform of big data, we can accurately meet the needs of students, create customized courses and textbooks, and provide students with personalized courses in language learning.

References

1. Fedynich, L. V. (2014). Teaching beyond the Classroom Walls: The Pros and Cons of Cyber Learning. Journal of Instructional Pedagogies, 13, 1. Retrieved October 12, 2018 from http://www.aabri.com/manuscripts/131701.pdf
2. Garnham, C. & Kaleta, R. (2002). Introduction to Hybrid Courses. Teaching with Technology Today, (8) 6. University of Wisconsin -Milwaukee. Retrieved September 22, 2018 from http://www.uwsa.edu/ttt/articles/garnham.htm.
3. Lloyd-Smith, L. (2010). Exploring the Advantages of Blended Instruction at Community Colleges and technical Schools. Journal of Online Learning and Teaching, 6 (2), Retrieved November 24, 2018 from http://jolt.merlot.org/vol6no2/lloyd-smith_0610.pdf.
4. Prangpatanpon, S. (1996). Higher education in Thailand: Traditional and bureaucracy. Retrieved October 2, 2018, from http://www.bc.edu/bc_org/avp/soe/cihe/newspaper/News 06/text8.html
5. Singh, H. (2003). Building effective blended learning programs. Educational Technology, (43), 6. 51-54.
6. Sitter, V., Carter, C., Mahan, R., Massello, C. and Carter, T. (2009). Hybrid course design: Faculty and student perceptions. Proceedings of the ASCUE 2009, Myrtle Beach, South Carolina. Retrieved October 2, 2018 from http://www.ascue.org/files/proceedings/2009/p40.pdf.
7. Spilka, R. (2002). Approximately ―real world‖ learning with the hybrid model. Teaching with Technology Today, 8(6). Retrieved September 20, 2018, from http://www.uwsa.edu/tt/articles/spilka.htm

Pick and Place of Large Object Based on 3D Vision

Hsien-Huang Wu
Dept. of Electrical Engineering, National Yunlin University of Science and Technology, No. 123, Sec. 3, University Road, Douliu City, Yunlin County 640, Taiwan (R.O.C.)

Jia-Kun Xie
Dept. of Electrical Engineering, National Yunlin University of Science and Technology, No. 123, Sec. 3, University Road, Douliu City, Yunlin County 640, Taiwan (R.O.C.)
E-mail: wuhp@yuntech.edu.tw

Abstract

Automation is a necessary tool to achieve unmanned factory, and machine vision plays a vital role for providing intelligent recognition in automation. In this study, technique of 3D camera is used for 3D image capturing and matching to identify, pick and place large objects automatically. The system uses a commercially available 3D stereo vision camera to build the image acquisition system, and reconstruct the large objects in 3D. This 3D image is used for classifying the objects with the 3D object recognition algorithm. After the object was identified and its 3d information was obtained, a robot arm integrated with the camera system can be used for grasping. Compared with traditional 2d image matching for 3d big object recognition, stable 2d image features are much harder to obtain due to the shadow. The 3D stereo vision camera does not require strict requirements for lighting control, and only needs stable ambient light. The reduction of the difficulty in building image acquisition environment for large 3d object and the cost of camera system provide a new option for applications that requires large object identification.

Keywords: Automated optical inspection, stereo computer vision, 3D modeling, 3D object recognition, robot vision

1. Introduction

In the production line of diversified production, the production process is planned by manpower, and the processing flow that various objects should be executed is identified. The advantage of visual inspection is that it does not require complicated mechanism design and precise and expensive loading and unloading mechanism, so that the object can be identified and the object can be transported by simple auxiliary mechanism, so the production process is relatively simple and the production cost is low. However, this method has the effect of human error, and if the object is large and difficult to handle, the labor cost is intensive and the production efficiency is reduced. The advantage of introducing AOI technology is that it can automatically detect objects, replace human visual inspection, and instantly record the current situation on the production line. By matching the robot arm, it can reduce labor demand, reduce human error, and improve production line efficiency.

The purpose of this research is to use the tool magazine's production line. The research uses three-dimensional computer vision to improve the bottleneck of traditional two-dimensional visual recognition in medium and large object recognition, and reduce the system construction cost.

In the reconstruction of the 3D model through the camera, Ref.1 and 2 uses binocular stereo to capture the same object with two cameras, and then reconstruct the 3D object; the use of laser triangulation[3,4] for 3D measurement of the surface of the moving object, such as: tire surface detection,[5] joint quality detection[6]; in the 3D reconstruction, the Kinect Fusion algorithm[7] is applied in many ways, through each depth image of the camera. Many 3D visions are also used in conjunction with robotic arms,[8,9] which are very popular and technically informative studies.

2. Architecture and Method

This system uses the 3D stereo vision camera with the robot arm and the main control computer to achieve 3D image capture and automatic identification and picking and matching of large objects. After that, grab it through the robot arm.

2.1. *Hardware architecture*

In the overall system planning, there will be two main coordinate systems (Fig. 1), which are the robot arm coordinate system and the camera coordinate system. Therefore, the two coordinate systems must be rotated and displaced to coordinate the coordinates, so that the object can be automatically picked and put after the object is recognized.

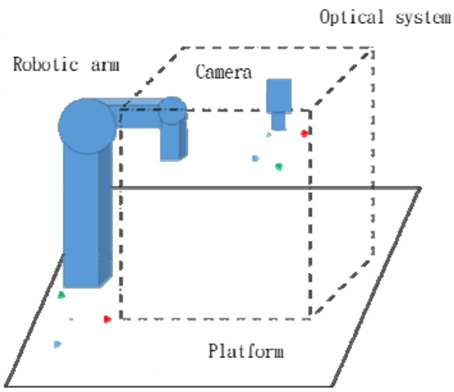


Fig. 1. System hardware architecture

2.2. *Software architecture*

The software architecture of this system is divided into three parts: depth image pre-processing and 3D modeling, 3D template matching algorithm, and setting arm coordinates. After the whole process is the tool magazine to the system, the depth image is taken through the stereo camera, and then the 3D modeling part is performed. Then 3D template matching algorithm is performed on the CAD model of the built model, and the object position is calculated. It then gives the information to the robot arm, and finally control the robot arm to move to the object position.

2.2.1. *Camera capture*

First, the stereo camera is used to capture the depth image of the object to be tested, which is a 16-bit grayscale image representing the distance between the object and the camera.(Fig. 2)

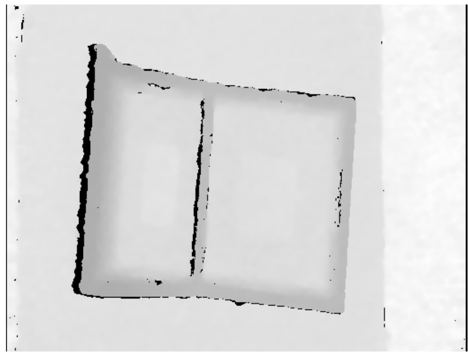


Fig. 2. Depth image by using stereo camera

2.2.2. *Image preprocessing*

After obtaining the depth image, the object area should be identified first. The noise in the image will be smoothed by the average filter to reduce the noise influence. Since the depth image represents the distance value from the camera, the area of the depth where the object may exist can be circled by simple binarization processing, and the connected area is used.

2.2.3. *3D object creation*

The grayscale value of each pixel of the depth image represents the distance from the depth camera, so three-dimensional information can be generated, which are (X, Y, Z), X and Y are pixel positions, and Z is depth. Information can be used to create 3D objects by calibration .

2.2.4. *3D Model import*

This research is to match the reconstructed objects through the existing CAD model, and there are many kinds of CAD models. In the model, it is necessary to confirm the relationship between the scale magnification of the imported model and the actual object size, and whether the coordinate direction of the reference plane of the model is consistent with the direction of the reference plane coordinate of the reconstructed object (Fig. 3). An inconsistency will result in an incorrect angle of the image matching result.

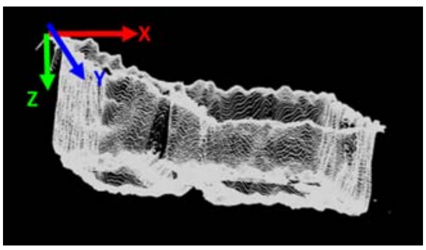


Fig. 3. Scene and model coordinates

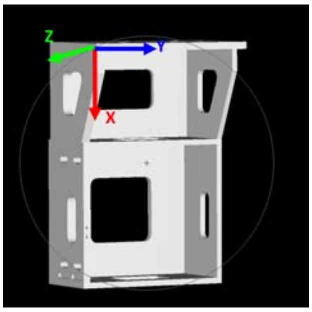


Fig. 3. (*Continued*)

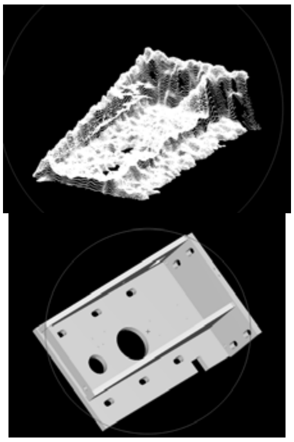
2.2.5. *Identification method*

After completing the reconstruction of the 3D object of the scene and adjusting the import model file, 3D object recognition will be performed. The method used in this study is proposed by Drost et al.[10] 3D point cloud matching method.

3. **Results**

3.1. *Reconstruction result*

This study contains images and model files used for reconstruction. Different objects should be reconstructed normally in the same shooting scene. This is the reconstruction result of the four actual object scenes and the corresponding CAD model. (Fig. 4)



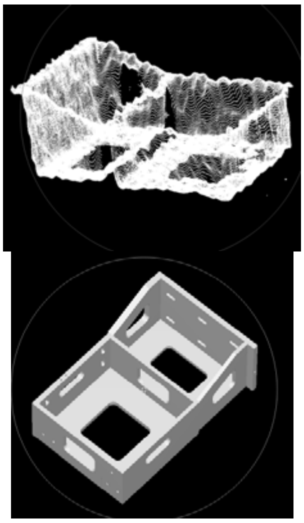


Fig. 4. Reconstruction result for each scene and CAD

3.2. *Object identification result*

The actual shooting scene is 4 kinds, and the CAD database has a total of 56 kinds of objects. In this research, the correct corresponding objects can be found in 56 kinds of CAD objects, and the object identification is completed. The figure shows the score curve of each scene for each CAD object. The blue dot in the Fig.5 is the sampling point of the scene reconstruction 3D object, and the yellow area is the matching model.

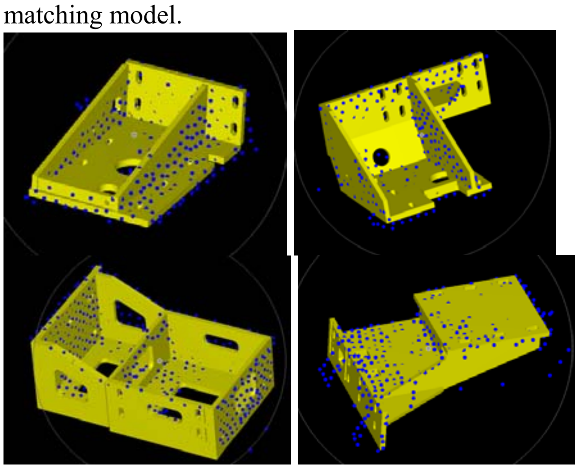


Fig. 5. Surface Match result

Fig. 6 shows the score trend graphs of each model and the same scene. It can be found that in the correct scene, the scores of the model are the highest, and the following figures are sorted from scene one to scene four.

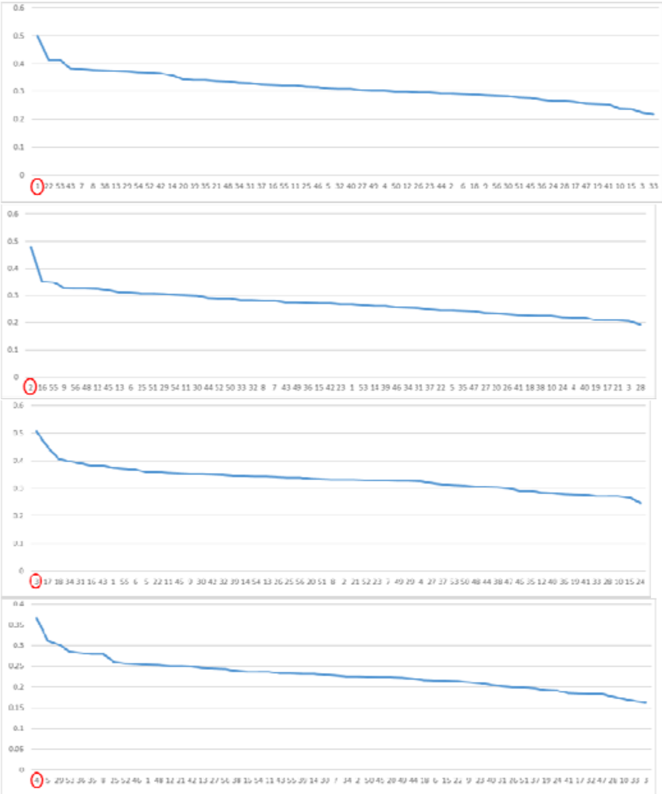


Fig. 6. Surface Match score line chart

4. Conclusions

In this paper, the automatic identification and pick-and-place system for large objects through 3D image capturing and matching can be used to solve the dilemma that large objects in the past are difficult to light in traditional image recognition, and the environment can be fully enriched by 3D cameras and light sources. To complete the object identification, if you need to add the identified object in the system in the future, you only need to import the 3D model file of the object, which does not require additional analysis of object features, increase system flexibility and reduce the technical cost of system update. The identification of the object and arm movement time is about one minute or so. The identification requirement for the application of the large object processing line is consistent. The biggest advantage is that the system is simpler than the traditional image recognition system, the shooting environment is low, and the construction cost is low. Low, is expected to lower the threshold for large object identification.

References

1. D. Bradley, T. Boubekeur, and W. Heidrich, "Accurate multi-view reconstruction using robust binocular stereo and surface meshing," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, 2008, pp. 1-8.
2. L. Yan, X. Zhao, and H. Du, "Research on 3D measuring based binocular vision," in *2014 IEEE International Conference on Control Science and Systems Engineering*, 2014, pp. 18-22.
3. L. Baozhang and H. Fengshan, "A Laser Distance Tracking 3D Coordinate Vision Measuring System," in *2011 Third International Conference on Measuring Technology and Mechatronics Automation*, 2011, vol. 1, pp. 83-87.
4. J. Park, D. Park, S. Cho, Y. Park, Y. Ko, and J. Lee, "3D Shape Restoration System for Femto-Second Laser Processing Utilizing Image Obtained by Line Scan Laser," *in 2006 SICE-ICASE International Joint Conference*, 2006, pp. 1968-1973.
5. I. Frosio, N. A. Borghese, P. Tirelli, G. Venturino, and G. Rotondo, "Flexible and low cost laser scanner for automatic tyre inspection," in *2011 IEEE International Instrumentation and Measurement Technology Conference*, 2011, pp. 1-5.
6. W. Huang and R. Kovacevic, "A Laser-Based Vision System for Weld Quality Inspection," *Sensors*, vol. 11, no. 1, 2011.
7. R. A. Newcombe et al., "KinectFusion: Real-time dense surface mapping and tracking," in *2011 10th IEEE International Symposium on Mixed and Augmented Reality*, 2011, pp. 127-136.
8. B. Drost and S. Ilic, "3D Object Detection and Localization Using Multimodal Point Pair Features," in *2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission*, 2012, pp. 9-16.
9. X. Fan, X. Wang, and Y. Xiao, "A combined 2D-3D vision system for automatic robot picking," in *Proceedings of the 2014 International Conference on Advanced Mechatronic Systems*, 2014, pp. 513-516
10. B. Drost, M. Ulrich, N. Navab, and S. Ilic, "Model globally, match locally: Efficient and robust 3D object recognition," in 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2010, pp. 998-1005.

Application of the MyRIO Based Mobile Robot Using Vision System

Bo-Jun Yang[1], Kuo-Hsien Hsia[2], Kuo-Lan Su[2], Evgeni Magid[3]
[1] *Graduate school Engineering Science and Technology, National Yunlin University of Science & Technology,*
[2]*Department of Electrical Engineering, National Yunlin University of Science & Technology,*
123 University Road, Section 3,Douliou, Yunlin 64002, Taiwan, R.O.C.
[3]*Intelligent Robotics Department, Kazan Federal University,*
18 Kremlyovskaya street, Kazan 420008, Russian Federation
E-mail: D10810211@yuntech.edu.tw, khhsia@yuntech.edu.tw, sukl@yuntech.edu.tw, magid@it.kfu.ru
www.yuntech.edu.tw

Abstract

The paper develops an auto mobile robotics that has the feature of image identification through a set of designed experiments. It is used in automatic industry and the application of AGV. The main controller of mobile robot is myRIO-1900 which is developed by America National Instruments. The inner controller is ARM Cortex-A9, it has two I/O ports, and it adopts MXP and MSP. This device includes and integrates analogy input, analogy output, digital I/O, LED, a button, accelerator, a set of Xilinx FPGA, processor, memory, hard disk, and image identification processing module. It uses Tetrix and Matrix components combining with RC motor to be an arm to solve Competition problem, and it also uses trapezoidal acceleration algorithm and PID algorithm to control more precisely. It applies NI LabVIEW Vision Assistant to do image identification processing to help mobile robot finish specified action.

Keywords: myRIO-1900, PID, Vision Assistant, Mobile Robots.

1. Introduction

With the incoming of Artificial Intelligence Era, robot education is gradually being valued. The output value of robots in Taiwan has been increasing year by year, and it is an important industry promoted by government and companies. Whether home appliances, entertainment, automatic robots in factory or medical care supply of seniors, all have a promising future, but compared with Japan and South Korea's market, the R & D and commercialization are slow. Robot education still has great room for improvement in Taiwan.

Robots have become next super-tech star after computers and Internet. All over the world are working hard to promote robot industry. Robot education has continued to take root to inspire students' interests in this field. In fact, robot education not only inspires children's interests in science and technology, but also cultivates talented people. For example, Lego cooperated with National Instruments (NI). With the promotion of Industry 4.0 and Productivity 4.0, the industry seeks to integrate intelligent automation in order to use in factory's production and transportation. Through artificial intelligence combines with robots, it can replace lots of human resources and provide more efficient production and transportation.

Behind robot education, there are a lot of creative and logical thinking. For example, students used to learn univariate quadratic equations in school, but through robot education, it can let students learn the application of these formulas which is far more effective. Extending robot education and assisting Ministry of Education to promote the science and technology education curriculum is important. As a result, children can be cultivated in their childhood to face and think about future high-tech society.

2. The Structure of System

The autonomous mobile robotics (as shown in Figure 1) consists of three parts: positioning system, motion control system, and image recognition system. The autonomous mobile robot includes power system, human-machine interface monitoring system, computer vision system, motor control system, and sensor system. The main control core is myRIO-1900 (as shown in Figure 2), and Tetrix and Matrix are the core components. The program is written by Labview2018 software which is developed by National Instruments. The human-machine interface monitoring, data transmission, image vision, and motor control are all performed by National Instruments LabVIEW2018.

The main core of myRIO controller is NI Single-Board RIO which is developed by National Instruments (NI). It is suitable for teaching, competition, and robot development with Tetrix and Matrix Base set metal kits. The embedded system of NI Single-Board RIO takes LabVIEW as the core, it allows users to integrate motors, sensor controls and creative assembly of metal kits more easily. With the expansion board, LabVIEW software writing programs is easier than traditional one. In addition, the programming language is similar to C language. The only difference is that LabVIEW integrates programming language, interface design, algorithms, visual development, measurement tools, communication modes, etc. Its functionality is relatively powerful, but its basic structure of programming language is still based on C language. LabVIEW is easier for users to do integration development than other programming languages.

LabVIEW (Laboratory Virtual Instrumentation Engineering Workbench) is a graphical program compilation platform which is developed by America National Instruments Corporation. LabVIEW was designed for automatic control of instruments in the early days, and now it has become a mature high-level programming language.

LabVIEW consists of three parts: Block Diagram, Front Panel, and Icon/Connector as shown in Figures 3 and 4. LabVIEW not only edits and designs programming, but also designs the interface. It is much easier for people

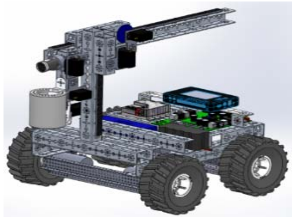


Fig. 1. Block diagram of IoT module.

Fig. 2 myRIO-1900

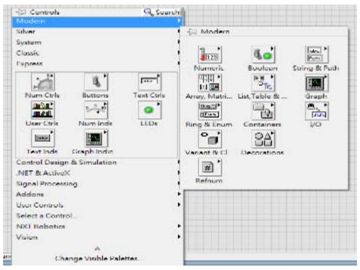


Fig. 3 front panel

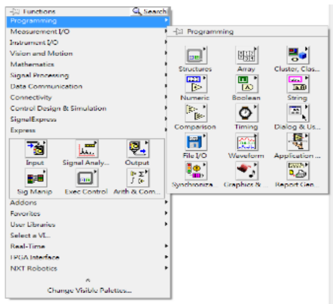


Fig. 4 edit the programming

without a programming foundation to learn than other programming languages.

3. Motion Control

PID is a common algorithm in the industry. People usually use it in the factories to control heating and cooling systems, flow control and pressure control. However, in this paper, PID is used to control the movement of robots, and with trapezoidal acceleration and deceleration described in next section, it makes the position of robots' movement more accurate. The setting value of PID controller (PV) compares with process variable (SP), then get the deviation (e).

$$e = SP|PV \tag{1}$$

Ratio control motion is K_c multiplies the deviation (e). K_c means the gain of controller. When there is only ratio control, the system will have steady-state deviation. The formula is as follows:

$$u_p(t) = K_c\, e \tag{2}$$

Ratio and integral control used to eliminate the steady-state deviation. In order to eliminate the steady-state deviation, the controller must input an integral term and accumulate the system deviation over time. T_i indicates the integration time constant, and the following represents the integration action:

$$u_I(t) = \frac{K_c}{T_i}\int_0^{\tau} e\, dt \tag{3}$$

Ratio and differential control used to eliminate transient error. Adding differential terms can accelerate system response time, but adding too much will cause system instability. T_d means differential time constant, and it is expressed as follows:

$$u_D = K_c T_d \frac{de}{dt} \tag{4}$$

In the PID loop, the response value required by above three calculation methods for motor controller is shown as below:

$$u(t) = K_c\left[e + \frac{1}{T_i}\int_0^{\tau} e\, dt + T_d\frac{de}{dt}\right] \tag{5}$$

If the control motor does not consider precise motion control, generally only needs I/O to control forward and reverse. When more precise positioning control is required, motors will have a little deviation due to some small tolerances. For example, when two same specification motors were controlled, they run at different speed, and made the car turned to left or right. In this paper, when AGV is in the motion control, we use trapezoidal acceleration and deceleration to make it move more accurately. In addition, it will make the car avoid overshoot when braking. Using trapezoidal acceleration/deceleration to control is better than using simple acceleration/deceleration without trapezoidal. For instance, when driving speed is from 0 to 100, then see the red light and do emergency braking, or do slowly brake when see the red light. Trapezoidal acceleration and deceleration use similar principles to control.

4. Image Recognition

When you need a specific color, threshold adjustment is used to analyze, process and delete unnecessary color images. Figures 5 and 6 show histograms of each plane of color images that stored in HSL format. The threshold range of grayscale representation in shaded area is the color in each plane. The pixels in color image will be set

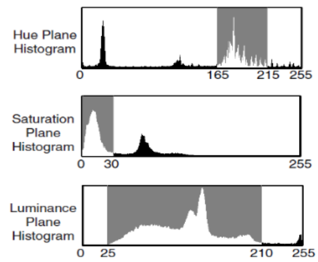


Fig. 5 HSL plane histogram.

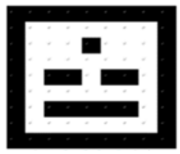


Fig. 6 Dividing into 81 points.

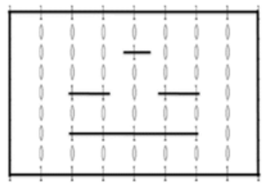


Fig. 7 images conversion.

to 1 in the binary image, and the color value is between 165 and 215, the saturation value is between 0 and 30, and the brightness value is between 25 and 210. Consequently, Hue, Saturation, and Luminance of colors can be defined more clearly, and it is different from general RGB.

The method to identify images in this paper is dividing the graphics into 81 points (Figure 6), 1 is black square and 0 is white square. After images conversion, they will be saved in database in txt file (Figure 7). Convert 16 graphics into 16 txt files, then store them in the same folder and store into the memory of myRIO, wait for recognition images comparing with the database.

In this paper, the fancy ball recognition is directly inputting RGB values, then binarizing the eight colors. This can improve the speed of image processing and the accuracy of finding balls. The recognition results can be found in the human-machine interface (as shown in Figure 8).

5. Experimental Results

Using image recognition, the area of the ball and the current position of the ball in the image can be calculated. The coordinates in the image must be converted into the coordinates of the robot, so that the coordinates of the ball can be handled well. Then find out the ball's radius through the image, you can know the seat of the ball on the court.

After the experimental measurement, the sphere image radius is about 184 at 20cm. However, the principle of farthest distortion and recent distortion must be considered, otherwise the image will be misjudged. If it is wrongly judged, you can cross-compare the measured radius and the distance from the original setting to solve the problem of misjudgement. The sphere image will shrink because of the distance, and zoom in if the distance is close. Using the ratio conversion method, you can convert the distance from the camera to the physical sphere by a few centimeters.

6. Conclusions

This paper mainly uses myRIO controller to develop competition platform and teaching materials for domestic mobile robotics competition. The robot’s movement and rotation motion control uses trapezoidal acceleration and deceleration and PID to control. MyRIO can define foot position, WIFI module and camera to let mobile robot complete the competition.

The structure of mobile robot uses Tetrix Robotics and Matrix Base Set to install. About the motor, it uses a gear ratio of 64:1 DC servo motor with high torque characteristics which developed by K-Kingdom. Therefore, it can make mobile robot apply in AGV. After experiments, it has good consequence in moving objects and image recognition of Pool.

7. Appendix

Features of the armed car in this paper:

- The control of movement uses trapezoidal acceleration and deceleration to make mobile robotics move more accurately.
- It can find color balls and two-color balls accurately, and catch the target effectively.
- The image recognition does not require complicated algorithms. Vision Acquisition of LabVIEW makes writing image programs much more efficient, and lets robot’s integration easier.
- It is effective to use single camera to estimate the distance between target object (sphere) and the vehicle.

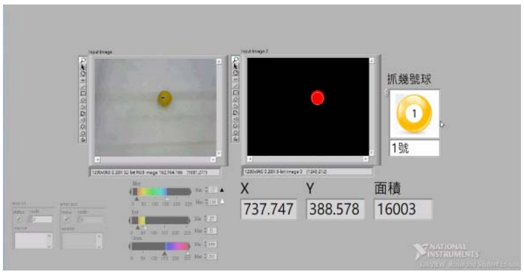
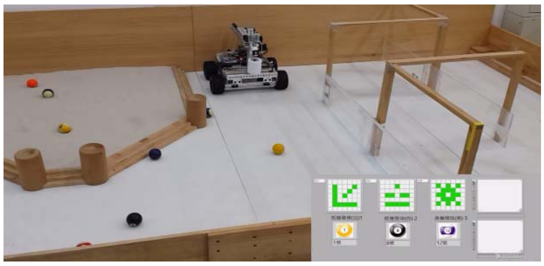
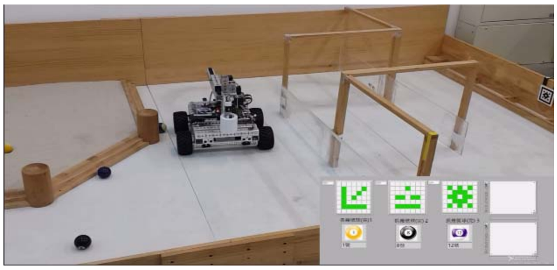


Fig. 8 Image human-machine interface.



(a)



(b)

Fig. 9 Experimental court of the mobile robot.

References

1. *The Industrial Market Capacity and Market Space Prediction of Worldwide Mobile Robotics*, http://www.51report.com/free/3036576.html, 2014.
2. *Automated Guided Vehicle*, http://en.wikipedia.org/wiki/Automated_guided_vehicle, 2014.

Real-time pattern recognition implementation on FPGA in multi-SNNs

Yang Xia
Graduate School of Engineering, The University of Tokyo, Japan
Tokyo, 4-7-1 Komaba, Meguro-ku,153-8505, Japan

Timothée Levi
LIMMS/CRNS-IIS, The University of Tokyo, Japan
Tokyo, 4-7-1 Komaba, Meguro-ku,153-8505, Japan

Takashi Kohno
IIS, The University of Tokyo, Japan
Tokyo, 4-7-1 Komaba, Meguro-ku,153-8505, Japan

E-mail: xiayang@sat.t.u-tokyo.ac.jp,

Abstract

By mimicking or being inspired by the nervous system, Neuromorphic systems are designed to realize robust and power-efficient information processing by highly parallel architecture. Spike timing dependent plasticity (STDP) is a common method for training Spiking Neural Networks (SNNs) for pattern recognition. Here, we present a real-time STDP implementation on FPGA in SNN using digital spiking silicon neuron (DSSN) model. Equipped with Ethernet Interface, FPGA allows online configuration as well as data input and output all in real-time. We show that this STDP implementation can achieve pattern recognition task and the connection between multi-SNNs enlarge the scale of networks and application.

Keywords: SNN, STDP, DSSN, FPGA, Ethernet

1. Introduction

Neuromorphic systems are designed by mimicking or being inspired by the nervous system, which aims to realize robust, autonomous, and power-efficient information. There are three common methods to realize the neuromorphic circuits, which are software[1], analog hardware[2,3] and digital hardware[4,5]. Compared to analog circuits, digital implementations generally consume higher power but are more scalable because they are less susceptible to noise and fabrication mismatch. And they are more configurable, portable, and low-cost with FPGA devices.

Silicon neuronal network (SNN) is a neuromorphic circuit that reproduces the electrophysiological activities in the nervous system.

In SNN they are modeled focusing on the spiking dynamics in the neuronal cells and their transmission via the synapses. Their application includes bio-inspired information processing such as pattern recognition[6,7] and associative memory[8] as well as neuro-prosthetic devices[9,10].

Digital SNNs are expected to achieve a very large-scale network comparable to the human brain in the future exploiting the scalability of the digital circuits. In these 5

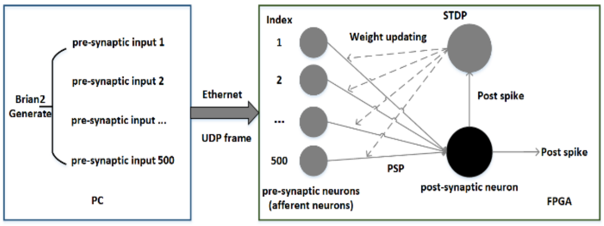


Fig. 1. Overall Architecture of STDP learning.

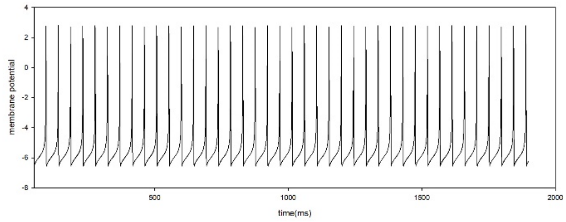


Fig. 2. Stimulation of a DSSN neuron by a constant input current (0.308).

years, several very large-scale SNNs with one million neurons were developed[11].

Spike-timing-dependent plasticity (STDP) is a well-known rule for updating the synaptic efficacy in SNNs, which uses only local information. Many biological experiments found evidence for STDP process in the synapses[12,13].

In this paper, we report an implementation of digital SNN with online STDP learning on FPGA. The model and implementation of our SNN are explained in the next section and section 3, respectively. Then results and conclusion follows.

2. Architecture of the network model

2.1. *DSSN model*

Neuronal models need to be chosen taking into account the balance between the reproducibility of neuronal activity and computational efficiency. Integrate-and-fire (I&F)-based models are able to be implemented by compact hardware, but they lack reproducibility of complex neuronal dynamics. Ionic-conductance models have high-ability of reproducing neuronal activities, but cost massive computational resources. The DSSN model is a qualitative neural model[14], which was designed for efficient implementation in digital circuits. The simplest version of DSSN model supports the Class I and II cells in Hodgkin's Classification[15].The differential equations of DSSN model are as follows.

$$\frac{dv}{dt} = \frac{\Phi}{\tau}(f(v) - n + I_0 + I_{stm}), \tag{1}$$

$$\frac{dn}{dt} = \frac{1}{\tau}(g(v) - n), \tag{2}$$

$$f(v) = \begin{cases} a_{fn}(v - b_{fn})^2 + c_{fn}(v < 0) \\ a_{fp}(v - b_{fp})^2 + c_{fp}(v \geq 0) \end{cases}, \tag{3}$$

$$g(v) = \begin{cases} a_{gn}(v - b_{gn})^2 + c_{gn}(v < r_g) \\ a_{gn}(v - b_{gn})^2 + c_{gn}(v \geq r_g) \end{cases}. \tag{4}$$

Here v represents the membrane potential, n is a variable that reflects the activities of hyperpolarizing ionic channels. a_{xy}, b_{xy}, and c_{xy} ($x = f, g$ and $y = n, p$) are parameters. The Parameter I_0 is a bias constant and I_{stm} represents the input stimulus.

The only nonlinearity in the DSSN model is a quadratic function. Thus, solving this model using Euler's method needs one multiplication operation per step if the parameters are carefully selected[8,16]. As multiplication operation requires relatively large resources in digital circuits, the DSSN model is suitable for digital silicon neuronal networks.

2.2. *Synaptic model*

Postsynaptic current (PSC) is a current inserted to the postsynaptic cell that induces temporal change in its membrane potential. The PSCs generated by a pulse stimulus to chemical synapses are able to be modeled by the alpha function with double-exponential generalization[17]. In our network model, the PSP model was simplified as follows.

$$\frac{dx}{dt} = \frac{-x}{\tau}. \tag{5}$$

Here x represents PSC generated by a chemical synaptic transmission. It is reset to w, the connection strength of the synapse, when the membrane potential of the presynaptic neuron exceeds 0. The initial value of w is $6\text{x}10^{-10}$. The PSCs in the synaptic connections to a postsynaptic are summed up as follows.

$$I_{stm} = \sum x_i. \tag{6}$$

Here i represents index of presynaptic neurons.

2.3. *STDP algorithm and architecture*

Synaptic plasticity means the magnitude of synaptic efficacy to change in response to the activities of pre and postsynaptic neurons.

The STDP is a biological rule that adjusts the strength of synaptic connections (w) based on the relative timing of the spikes in a pre and postsynaptic neurons. Recently it

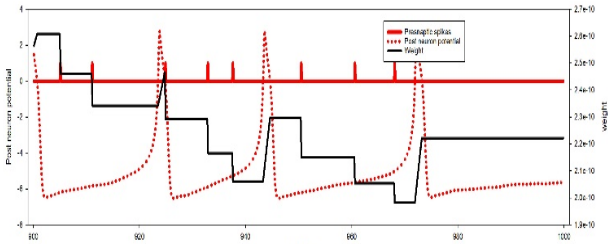


Fig.3. Waveform of STDP method implementation.

has been shown how STDP rule play a key role by detecting repeating patterns and generating selective response to them[18].

The STDP rule is a most common form of learning rules used in SNNs. Here is the standard exponential STDP equations.

$$w_{new} = w + \Delta w, \tag{7}$$

$$\Delta w = \begin{cases} a^{+} \cdot \exp\left(\frac{t_j - t_i}{\tau^{+}}\right), if\ t_j \leq t_i (LTP) \\ -a^{-} \cdot \exp\left(-\frac{t_j - t_i}{\tau^{-}}\right), if\ t_j > t_i (LTD) \end{cases}. \tag{8}$$

Here Δw is the modification of the synaptic weight. When a post-synaptic spike arises after a pre-synaptic spike ($t_j \leq t_i$), the connection is reinforced (long-term potentiation (LTP), $\Delta w > 0$), whereas in the opposite case it is weakened (long-term depression (LTD)).

3. Implementation

The overall architecture of our SNN is shown in Fig.1.The DSSN, STDP, and PSC blocks were implemented on a FGPA chip. In this chip, single postsynaptic neuron is connected with 500 input afferent neurons. All the circuits were described using VHDL. Stimulus spike trains were generated by Brian2 on PC and sent to FPGA via Ethernet connection which are explained in 3.4 and 3.5.

3.1. *Implementation of DSSN*

The DSSN model's differential equations are solved by Euler's method with the timestep of 0.1ms. The solver circuit has time division multiplexing (TDM) and pipelined architecture.

3.2. *Implementation of PSC*

As introduced in 2.2, we use a simplified exponential decay to approximate the PSCs. The connection between a postsynaptic neuron and the 500 pre-synaptic neurons was calculated efficiently by using TDM and two-stage pipeline as in the DSSN solver circuit. Operations for the 500 sets of PSCs were finished in 1000 clock cycles, by using single multiplier and single adder. The time step for the summation of PSCs is within 0.1ms if the system clock is faster than 100MHz.

3.3. *Implementation of STDP learning*

Based on the STDP rule, each presynaptic (postsynaptic) spike, induces an LTD (LTP) where the synaptic weight is updated according to Eq. (7). An example waveform of PSC is shown in Fig.3, where the solid and dotted red curve represent the presynaptic and postsynaptic spikes, as well as the black curve is the strength of synaptic connection. For simulating 500 afferents, 500 clock cycles were consumed. Calculation of exponential function and update of synaptic efficacy are executed by TDM.

3.4. *Ethernet on FPGA*

Ethernet is a computer networking technology commonly used in local area networks (LAN). Devices equipped with Ethernet interface are connectable to LAN as well as the internet by supporting common communication protocol, for instance, TCP/IP and UDP. For neuronal network implementation on FPGAs, transmission of the spikes is always an issue particularly when the number of neurons is large.

By implementing Ethernet interface and a full hardware protocol stack including IP, UDP and ARP protocol, our SNN on FPGA is capable of receiving input spike patterns from a PC by Ethernet connection and sending report frame back to the PC for monitoring the SNN working status at any time.

The configuration and parameter setting are also possible by the Ethernet connection, which contribute to take advantage of FPGA in flexibility and portability.

4. Pattern recognition

Spatiotemporal spike pattern recognition tasks proposed by Masquelier[19] were performed on our network. The stimulus pattern required in them were generated by superimposing a spatiotemporal spike pattern at many time points on a background random spike pattern. The former spike patterns (ones to be detected) were generated by PoissonGroup Function in Brian2 (a Python library) on PCs. The background random spike patterns were generated using the same library. The stimulus patterns were stored in files and sent from the PC to the FPGA via the Ethernet connection. An example of

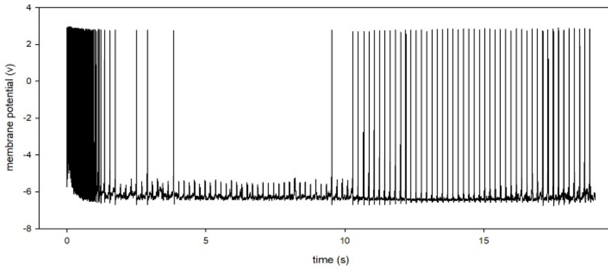


Fig. 4. Waveform of STDP learning process. The learning is finished after around 11s.

learning process is shown in Fig.6. The postsynaptic neuron started to detect the superimposed patterns at 11s.Our circuit performed this task in real-time. FPGA's system clock was 100MHz.

5. Conclusion

In this paper, a SNN on an FPGA with STDP learning capability were reported. Stimulus spikes were conveyed in real-time to the FPGA chip via Ethernet connection. It was proved that our SNN is capable of spatiotemporal spike pattern recognition in real-time.

In this work, the input patterns were generated on a PC. Generating them in real-time with another SNN is a future work. It is also planned to develop an FPGA-FPGA connection bus to expand the scalability of the FPGA-based SNNs, because the scale of the network on an FPGA chip is limited by the amount of on-chip memory used for storing the strength of synaptic connections. An SNN platform with high scalability is advantageous compared to software simulation from the viewpoint of real-time operation and power consumption.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 15KK0003.

References

1. M.L. Hines and N.T. Carnevale, *NEURON: a tool for neuroscientists*, The Neuroscientist 7, 123-135 (2001).
2. A. Natarajan ans J. Hasler, *Hodgkin-Huxley Neuron and FPAA Dynamics*, *IEEE Trans Biomed Circuits Systems*, 12(4), 918-926 (2018).
3. F. Grassia, L. Buhry, T. Levi, J. Tomas, A. Destexhe and S. Saïghi, *Tunable neuromimetic integrated system for emulating cortical neuron models*, *Frontiers in Neurosciences*, 5(134) (2011).
4. M. Ambroise, T. Levi, S. Joucla, B. Yvert and S. Saïghi, *Real-time biomimetic Central Pattern Generators in an FPGA for hybrid experiments*, *Frontiers in Neurosciences,* 7(215) (2013).
5. Grassia F, Kohno T, Levi T, *Digital hardware implementation of a stochastic two-dimensional neuron model*, *Journal of Physiology Paris*, 110(4), 409-416 (2016).
6. N. Qiao et al., *A reconfigurable on-line learning spiking neuromorphic processor comprising 256 neurons and 128K synapses*, *Frontiers in Neuroscience*, 9(141) (2015).
7. P. Merolla et al., *A million spiking-neuron integrated circuit with a scalable communication network and interface, Sciene*, 345(6197), 668-673 (2014).
8. Li J, Katori Y, Kohno T. *An FPGA-Based Silicon Neuronal Network with Selectable Excitability Silicon Neurons. Front Neurosci,* 6(183) (2012).
9. F. Broccard, S. Joshi, G. Cauwenbergs, *Neuromorphic neural interfaces: from neurophysiological inspiration to biohybrid coupling with nervous systems, J. Neural Eng.*, 14(4):041002 (2017).
10. S. Buccelli et al., *A neuroprosthetic system to restore neuronal communication in modular networks, iScience*, 19, 402-414 (2019).
11. M. Davies et al., *Loihi: a neuromorphic manycore processor with on-chip learning*, *IEEE Micro*, 38(1), 82-99 (2018).
12. Markram, H., Lübke, J., Frotscher, M., and Sakmann, B., *Regulation of synaptic efficacy by coincidence of postsynaptic aps and epsps. Science* 275(5297), 213-215 (1997).
13. Jacob, V., Brasier, D. J., Erchova, I., Feldman, D., and Shulz, D. E., *Spike-timing-dependent synaptic depression in the in vivo barrel cortex of the rat. J. Neurosci,* 27(6), 1271-1284 (2007).
14. Kohno T., Aihara K., *Digital spiking silicon neuron: concept and behaviors in GJ-coupled network, in Proceedings of International Symposium on Artificial Life and Robotics,* OS-6 (2007).
15. Hodgkin A. L.,*The local electric changes associated with repetitive action in a non-medullated axon. J. Physiol,* 107(2), 165-181 (1948).
16. Nanami T, Kohno T. *Simple Cortical and Thalamic Neuron Models for Digital Arithmetic Circuit Implementation. Front Neurosci,* 10(181) (2016).
17. Amari.S, *The Handbook of Brain Theory and Neural Networks, MIT press* (1995).
18. Grassia F, Levi T, Doukkali E, Kohno T, *Spike pattern recognition using artificial neuron and spike-timing-dependent plasticity implemented on a multi-core embedded platform, Journal on Artificial Life and Robotics, Springer Nature*, 23(2), 200-204 (2018).
19. Masquelier T,Guyonneau R,Thorpe SJ,*Competitive STDP-based spike pattern learning. Neural computation*, 21(5), 1259-1276 (2009).

Experimental results of a biomimetic analog silicon synaptic circuit.

Ashish Gautam
*Graduate School of Engineering, 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: asgautam@sat.t.u-tokto.ac.jp, kohno@sat.t.u-tokyo.ac.jp
www.u-tokyo.ac.jp*

Abstract

We present the experimental results of a biomimetic silicon synaptic circuit capable of generating both excitatory and inhibitory postsynaptic currents (EPSC's and IPSC's). All the synaptic circuits proposed till date have either excitatory or inhibitory dynamics alone. This single circuit is capable of emulating the dynamics of the major synaptic populations received by a neuron. The first-order dependence of the synaptic current on the instantaneous value of the membrane potential is also taken into account, with the synaptic current being proportional to the difference between a tunable synaptic reversal potential and membrane potential of the postsynaptic neuron.

Keywords: Silicon Neuronal Networks, Synaptic Reversal Potential, Low-power.

1. Introduction

Networks of spiking neurons implemented in silicon referred to as Silicon Neuronal Networks (SNNs) armed with local and unsupervised learning rules like STDP are a prime candidate to arrive at the goal of realizing autonomous learning machines addressing the major issues of high power consumption and von-Neumann bottleneck prevalent in contemporary computing architectures. Research labs over the globe have come with various hardware platforms to implement analog SNN[1, 2, 3]. Biologically inspired neuronal soma and synaptic circuits form the basis of this hardware platform. In this manuscript, we present the experimental results of a low-power analog silicon synaptic circuit. In the nervous system, a large number of synapses converging on to a single neuronal soma equip it to implement a wide variety of linear and non-linear operations arming a network of neurons to perform complex computational tasks. There are multitudes of chemical synapses in the nervous system comprising various neurotransmitters (glutamate, GABA, etc.) which act on specific receptors on the postsynaptic membrane with high (AMPA, NMDA) or low ($GABA_A$, $GABA_B$) reversal potentials. The presented circuit is the first low-power silicon synapse circuit that takes into account the first-order dependence of synaptic current on the difference between synaptic reversal potential and the postsynaptic membrane potential. This phenomenological description of synaptic current is of paramount importance in implementing the experimentally observed phenomenon of shunting or silent inhibition. All circuits presented to date use different circuits for implementing excitatory and inhibitory synapses and only a few[3, 4] of them take into account the effect of postsynaptic membrane potential. Special emphasis has been put in the design to minimize the static power consumption of the synaptic circuit, given that a large scale Silicon Neuronal Network will have thousands of synapses activated in an event-based manner. The proposed circuit was designed in TSMC 250nm process. The next section of this manuscript deals with a brief description of the synaptic circuit along with the setup used for measurements, followed by results, discussion, and the planned future work.

2. Synaptic Circuit

The proposed synaptic circuit (Fig.1) consists of a binary-weighted Digital to Analog Converter (M1-M10), an integrator circuit similar to the log domain integrator (C_{syn} & M11), and a transconductance stage (M13-M19) to simulate the first-order dependence of synaptic current described below.

$$\boldsymbol{I_{syn}(t) = g_{syn}.(V_{mem}(t) - E_{syn})}, \qquad (1)$$

where g_{syn} is the synaptic conductance, V_{syn} is the post-synaptic potential and E_{syn} is the synaptic reversal potential. The equations describing the circuit's operation are described in detail in Ref.5. The synaptic circuit operates as follows. An input pulse at node nV_{in} activates the DAC, charging the capacitor C_{syn} of the integrator stage through a constant current determined by the voltage sV_w and the state of the DAC. After the input pulse, the DAC turns off and the transistor M11 (in saturation) discharges C_{syn} sinking in a constant current controlled by the voltage sV_t. This node V_{syn} is then fed to the transconductance stage, which generates either an EPSC or IPSC based on E_{syn} and V_{mem}. To minimize static power consumption, a transconductance amplifier with p-type current mirror was used, which led to the use of n-type transistor in the DAC and integrator stages unlike conventional designs[6]. In contrast to the design approach in Ref.7 where significant stress is put on the accuracy of the DAC resolution, and Ref.8 where independent current sources were used to implement binary encoding leading to overhead in terms of power consumption, we designed a more compact DAC circuit (see Fig.1 caption) trading on the accuracy of the DAC resolution. It is still unknown how many bits resolution the synapses has in the brain. So, we take a midway approach making sure that the DAC output changes monotonously with the change in the corresponding input synaptic efficacy, without significant stress on the accuracy of the resolution of the DAC (i.e. value of jump in the output per unit input change). The fabricated chip consists of 128 synaptic circuits connected to a single qualitatively modeled[9] neuronal soma circuit. The synaptic circuits were laid out in groups of four. Between each of these groups, a resistor implemented using MOS transistor operating in the linear region was inserted to model passive dendritic cable properties. Another group of 16 synapses, (hereafter referred to as TEG) were connected to a high resistance circuit designed using a transconductance amplifier with source degeneration to measure the synaptic current. The synaptic weights were stored in registers and a spike address decoder was used to route spikes to desired synaptic inputs. The registers and spike address decoders in the chip were controlled using an FPGA.

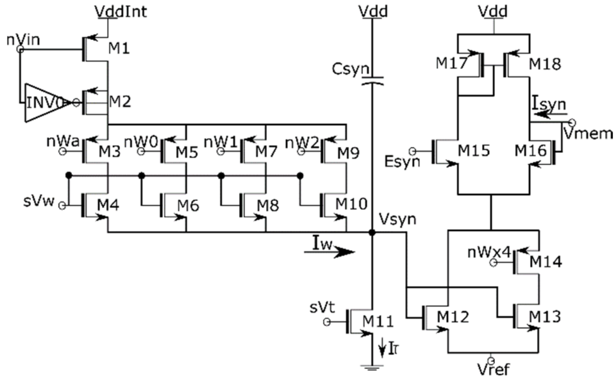


Figure. 1: Schematic diagram of the proposed synaptic circuit. Dimensions of the transistors: M4=0.3758*(w/l), M6=w/l, M8= 2*(w/l) and M10=4*(w/l). Tail transistor M13 is three times wider than M12.

3. Results

The first set of results are measured from synapses in the TEG. Unless specified, parameters listed in the caption of Fig.3 were used for measurement. The synaptic waveforms mimicking the behavior of the four major neurotransmitters are shown in Fig.2. Fig.3 shows the dependence of synaptic current on the synaptic reversal potential (E_{syn}). The postsynaptic membrane potential (V_{mem}) was fixed at 700mV. As expected from the phenomenological model in Eq.(1) the relationship is linear. The synaptic current was measured by the voltage across the integrated high resistance circuit. The zero-crossing occurs at around E_{syn}=707mV, this offset is due to the mismatch of the transistors in the transconductance stage. Fig.4 shows the variation of time constants against

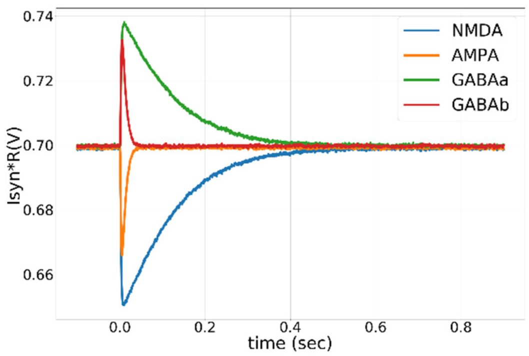


Figure 2: Synaptic currents emulating the response of AMPA, NMDA (synapse No.1) $GABA_A$, $GABA_B$ (Synapse No. 2) neurotransmitters. sV_t was set to 10mV for NMDA and $GABA_b$ mode and to 120mV for AMPA and $GABA_a$ mode.

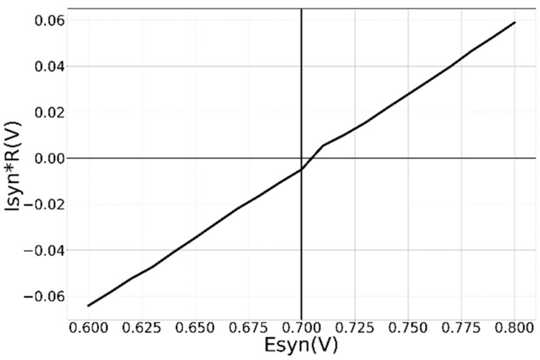


Figure. 3. Dependence of synaptic current on synaptic reversal potential. Y-axis reflects the peak value of the synaptic current. Parameters used: sV_w=140mV, sV_t=0, E_{syn}=600mV. Efficacy: nWa, nW0, nW1, nW2, nWx4=1.

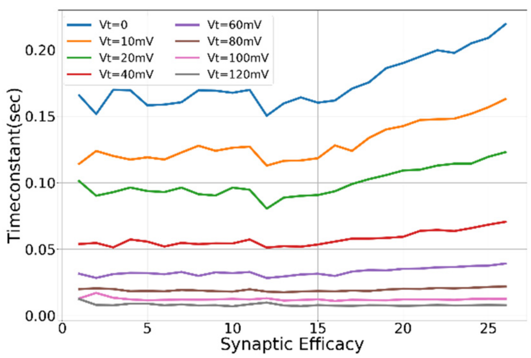


Figure. 4. Profile of time constant for various values of sVt across the range of synaptic efficacy.

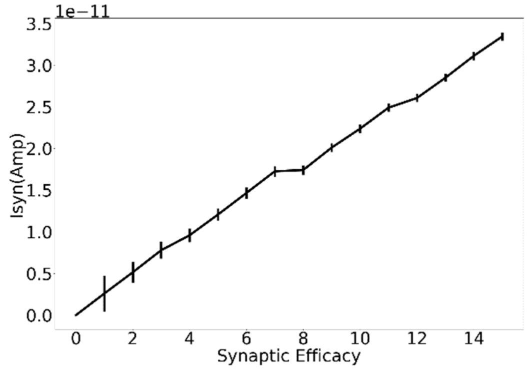


Figure. 5. Average value of the peak synaptic current for 128 synapses across the range of synaptic efficacy to characterize the DAC. Error bars capture variation across mean value of current.

the synaptic efficacy for various sV_t values. The maximum time constant obtained for the value of sV_t=0V was on average around 160ms. Mismatch in the 128 synaptic circuits connected to a neuronal soma circuit was measured. The spiking circuitry of the neuron was disabled, thus acting as a passive membrane. The value of dendritic excitatory postsynaptic potential (EPSP) from each of the 128 synaptic circuits for 16 different weight values (from 0000 to 1111 in the binary-weighted DAC) were measured giving total 2048 EPSPs. As the relation between EPSP and EPSC is not linear, an elaborate procedure involving current sources in the neuron circuit was used to construct a lookup table mapping the value of EPSP to corresponding EPSC. Fig.5 shows the average and standard deviation of the measured peak synaptic current for 128 synapses plotted against the synaptic efficacy. Out of 128 synapses a kink in the desired monotonicity was observed in 58 synapses between the synaptic efficacy value 7(1000) to 8 (0111). The maximum value standard deviation of the peak synaptic current was between 5% (maximum weight) to 20% of the mean (minimum weight) value (Error bars in Fig.5).

4. Discussion

As described above the deviation from monotonicity with respect to the value of synaptic efficacy was observed at one point for around 46% of the synapses. Around 6% of the synapses displayed deviation from monotonicity between the synaptic efficacy value of 3 & 4 and 11 & 12. The probable reason for this deviation is device mismatch combined with second-order and parasitic effects associated with the half-size transistor, but we could not confirm this quantitatively due to the unavailability of statistical mismatch data in TSMC 250nm CMOS PDK needed to perform Monte Carlo simulation. Fixing this is one of our future works. The current mirror devices of the transconductance stage were designed using relatively short transistors (l=400nm) to make sure that the response time of the inhibitory synaptic transmission (E_{syn}=800mV) was fast enough. The mismatch due to this short device contributed to the shift of zero offset in Fig.3. The variance in the peak value of the synaptic current for a particular value of synaptic efficacy was higher in the case of inhibitory synapses than in the case of excitatory synapses. This mismatch along with the Drain Induced Barrier Leakage effect (on M18) lead to leakage current being higher in the case of inhibitory synapses than in the case of excitatory synapses. Though this additional leakage current can be compensated in the soma circuit, this difference needs to be minimized. One possible way which was verified in the simulation was to separately control the bulk voltage of the current mirror transistors and fix it at a value higher than V_{dd}. We are working on a new transconductance stage to address these issues.

5. Conclusion

Experimental results of a low power biomimetic synaptic circuit fabricated in TSMC 250nm technology node were presented. The circuit is the first low-power silicon synaptic circuit that takes into account the effect of synaptic reversal potential. The design is focused on minimizing static power consumption which is a metric not yet evaluated in most contemporary CMOS based synaptic circuits. The static power consumption is evaluated to be less than 2pW and the dynamic energy consumption is measured as follows. For sVw=110mV (peak synaptic current of around 5pA) and time constant of 160ms, the input stage of our circuit consumes around 40fJ per synaptic event and the integrator and the transconductance stage consume around 500fJ (this value reduces to 112fJ for a time constant of 3ms). This first-generation chip contains only 128 synapses with just one input stage per integrator. The number of input stages can be increased to a higher number, the design is completely modular with no limit on the number of synapses that can be implemented. Issues related to device mismatch in the DAC and the transconductance stage were characterized and will be fixed in the near future.

6. Acknowledgements

This study was supported by NEC Corporation and CREST, and DLab, The University of Tokyo in collaboration with Cadence Design Systems, Inc.

7. References

1. Giacomo Indiveri, Federico Corradi, and Ning Qiao. Neuromorphic architectures for spiking deep neural networks. In *IEEE International Electron Devices Meeting (IEDM),* 2015.
2. Brink, S., Basu, A., Degnan, B.P., Hasler, P.E., Nease, S., Ramakrishnan, S., & Wunderlich, R.B. (2013). A Learning-Enabled Neuron Array IC Based Upon Transistor Channel Models of Biological Phenomena. *IEEE Transactions on Biomedical Circuits and Systems,* 7, 71-81.
3. J. Schemmel, L. Kriener, P. Muller, and K. Meier, "An accelerated analog neuromorphic hardware system emulating NMDA- and calcium-based non-linear dendrites," in Proceedings of *the International Joint Conference on Neural Networks,* 2017, vol. 2017–May, pp. 2217–2226.
4. B. V. Benjamin, J. V. Arthur, P. Gao, P. Merolla, and K. Boahen, "A superposable silicon synapse with programmable reversal potential," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS,* 2012, pp. 771–774.
5. Gautam, A., and Kohno, T. A low power silicon synapse with tunable reversal potential in *The 2018 International Conference on Artificial Life and Robotics (ICAROB2018),* Feb. 1-4, B-ConPlaza, Beppu, Oita, Japan.
6. Bartolozzi,C., and Indiveri,G . (2007). Synaptic dynamics in analog VLSI. *Neural Computation*, 19(10):2581-2603.
7. Wang, Y.,and Liu, S. Programmable synaptic weights for an aVLSI network of spiking neurons*. 2006 IEEE International Symposium on Circuits &Systems.*doi:10.1109/iscas.2006.1693637.
8. S. Moradi and G. Indiveri, "An event-based neural network architecture with an asynchronous programmable synaptic memory," *IEEE Trans. Biomed. Circuits Syst., vol. 8, no. 1,* pp. 98–107, 2014.
9. Kohno, T., and Aihara, K. A three variable ultralow-power analog silicon neuron circuit. *NOLTA 2016 (Yugawara, Japan, November 27-30, 2016).*

Towards Modeling Cholinergic Modulation for Neuromorphic Computing

Naruaki Takano
Department of Information Science and Technology, The University of Tokyo,
*7-3-1 Hongo, Bunkyo-ku, Hongo, Bunkyo-ku, Tokyo 113-8656, Japan**

Takashi Kohno
Institute of Industrial Science, The University of Tokyo, 4-6-1, Komaba, Meguro-ku, Tokyo, 153-8505, Japan
E-mail: naruakitakano@sat.t.u-tokyo.ac.jp , kohno@sat.t.u-tokyo.ac.jp
www.sat.t.u-tokyo.ac.jp/~kohno/

Abstract

Digital Spiking Silicon Neuron (DSSN) model is a qualitative neuron model specifically designed for digital circuit implementation which exhibits high biological plausibility. In this study we analyzed the behavior of an all-to-all network composed of 3-variable DSSN model which has a slow negative feedback variable corresponding to a slow calcium-dependent potassium current. We observed the network dynamics by altering the magnitude of the slow negative feedback current which is known to be controlled by cholinergic modulation, and the strength of neuronal interaction. By altering these parameters, we obtained various pattern retrieval dynamics, such as chaotic transitions within stored patterns or stable and high retrieval performance. We will briefly discuss potential applications of these results for neuromorphic computing.

Keywords: Spiking neural network, Associative memory, DSSN model, Spike frequency adaptation

1. Introduction

Neuromorphic computing is an attempt to realize a computing architecture capable of complex and robust calculations with low power consumption, which is designed by either imitating or inspired by the information processing mechanisms in the nervous system.

Unlike the conventional von Neumann computer architecture in which CPU and memory units exist separately, the neuromorphic hardware generally has highly distributed computation and memory units. Such a massively parallelized computation which overcomes the von Neumann bottleneck is expected to reduce overall energy consumption.

Silicon neuronal networks (SNNs) are a neuromorphic hardware that attempts to copy the electrophysiological activities in the nervous system by electronic circuits. Models of neurons, synaptic activities and designs of underlying algorithms are crucial factors as components in the SNNs. Many neuron models have been proposed for reproducing various traits of neurons. However, they generally undergo the problem of trade-off between implementation cost and biological plausibility.

Digital Spiking Silicon Neuron (DSSN) model is a qualitative neuron model which reproduces dynamical structure in various neuronal firing activities and is designed to be implemented efficiently on digital circuit[1].

As a fundamental model of neuromorphic computation, an all-to-all connected network composed of DSSN model has been implemented and autoassociative task was performed[2]. The network retrieved a stored pattern given corresponding corrupted input patterns, and its performance varied greatly on each neuron's parameters. It was shown that the performance is dependent on the class of the neuron model (Class 1 and Class 2 in Hodgkin's Classification).

In our study, we adopt the equivalent system and incorporate Regular Spiking class for each neuron. Regular Spiking class exhibit Spike Frequency Adaptation (SFA) which is a convergence of firing frequency from high value to a low value given a step current with sufficiently large magnitude.

Moreover, the strength of SFA is known decrease by cholinergic modulation and has been modeled on network of Hodgkin Huxley type neuron model[4,5]. We aim to apply and utilize the modulatory effect for neuromorphic computing and studied network dynamics by altering corresponding parameters of the DSSN model.

2. Neuron and Synapse Model

Two variables in the DSSN model, membrane potential v and slow variable n that represents a group of ion channels is responsible for generating neuronal spikes.

SFA is reproduced by introducing the slowest negative feedback variable q. This is a 3-variable DSSN model[2]:

$$\frac{dv}{dt} = \frac{\phi}{\tau}\left(f(v) - n - q + I_0 + I_{stim}\right)$$
$$\frac{dn}{dt} = \frac{1}{\tau}\left(g(v) - n\right)$$
$$\frac{dq}{dt} = \frac{\epsilon}{\tau}\left(v - v_0 - \alpha q\right) \tag{1}$$

where,

$$f(v) = \begin{cases} a_{fn}\left(v - b_{fn}\right)^2 + c_{fn} \ (v < 0) \\ a_{fp}\left(v - b_{fp}\right)^2 + c_{fp} \ (v \geq 0), \end{cases} \tag{2}$$

$$g(v) = \begin{cases} a_{gn}\left(v - b_{gn}\right)^2 + c_{gn} \ (v < v_g) \\ a_{gp}\left(v - b_{gp}\right)^2 + c_{gp} \ (v \geq v_g). \end{cases} \tag{3}$$

Here, parameter ϵ is the inverse time constant of the adaptation dynamics. Parameter α controls strength of SFA, which is the extent of difference of inter-spike-interval (ISI) between the initial and the converged state. When α is low, the effect of SFA is strong (see Fig. 1). Therefore, magnitude of α can be mapped to concentration of Acetylcholine during modulation.

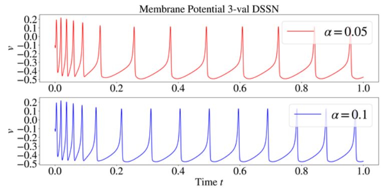


Fig. 1. Spike frequency adaptation observed in DSSN model for α = 0.5 (Top) and 0.1 (Bottom).

Synaptic current is modeled by combination of two exponential curves and is dependent upon neuron's membrane potential,

$$\frac{dI_s}{dt} = \begin{cases} \alpha(1 - I_s) & (v > 0) \\ -\beta_s I_s & (v \leq 0). \end{cases} \tag{4}$$

3. Network Configuration and Method

The network is an all-to-all connected network of 256 neurons, and synaptic efficacy in each synapse is given by synaptic weight matrix W. External input to i^{th} neuron is given as,

$$I_{stim}^i = c\sum_{j=1}^{N} W_{i,j}\, I_s^j + I_{ext}^i, \tag{5}$$

where c is a parameter that denotes the magnitude of neuronal interaction and I_{ext} is a constant bias current.

We stored 4 mutually orthogonal patterns (Fig. 2a) in the network by configuring W by the following correlation rule:

$$W_{i,j} = \begin{cases} \frac{1}{p}\sum_{u=1}^{p} w_u x_i^u x_j^u & (i \neq j) \\ 0 & (i = j). \end{cases} \quad (p = 4, 1 \leq i,j \leq 256) \tag{6}$$

Here, x represents 1 dimensional vector of size 256, where each pattern (16 × 16 pixels) are encoded by binary value: +1 for black pixel and -1 for white pixel. w_u encodes weight bias of each pattern which are $w_u = 1.0$ $(u = 1,2,3,4)$ unless explicitly noted. Based on the stored patterns, we produced sets of corrupted input data, where certain percentage of randomly selected pixels are inverted (Fig. 2b).

(a) (b)

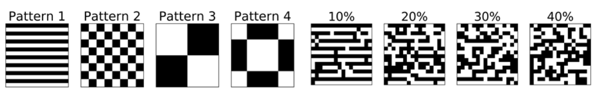


Fig. 2. (a) Stored patterns. (b) Input data (based on pattern 1).

The autoassociative task is performed as follows: First, we initialized the network state with corrupted input data by injecting step current $I_{ext} = 0.15$ for 0.5 seconds to the neurons that encode black pixel. Then, we stimulated all the neurons with $I_{ext} = 0.15$ to evoke repetitive neuronal activity.

We evaluated the network state by an overlap index M_u which quantifies the similarity between firing times of every neuron and u^{th} stored pattern. Based on the firing phase in the repetitive firing activity, overlap between the network output and the stored patterns is computed by,

$$M_u(t) = \frac{1}{N}\left|\sum_{j=1}^{N} \exp\left(i\phi_j(t)\right)\right| \quad (7)$$

(i: imaginary number),

where firing phase of neuron j is defined as,

$$\phi_j(t) = 2\pi k + 2\pi \frac{t - t_j^k}{t_j^{k+1} - t_j^k}. \quad (8)$$

(t^k: k^{th} firing time of neuron j)

4. Simulation results

4.1. *Weak neuronal interaction*

We obtained high retrieval performance at the weak neuronal interaction ($c = 0.005$) regime (Fig. 3a, b). The retrieval performance is comparable with network of class 2 neurons and can be reasoned by biphasic (type 2) phase response curve[6] (PRC) of each neuron (Fig. 4).

PRC of Regular Spiking class neuron was computed by applying the perturbation only after the adaptation has fully stabilized in order to determine a unique firing period. We found that the zero-phase-shift-crossing point is essentially altered by α. Therefore, we can design a specific PRC by the alteration of α.

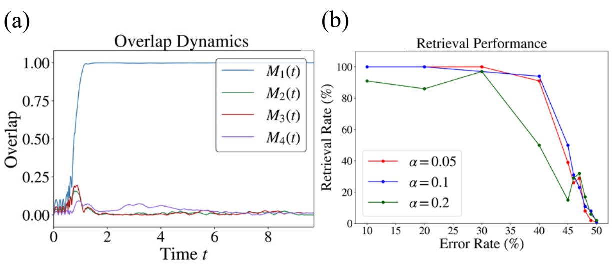


Fig. 3. (a) Overlap in a successful retrieval process. (b) Retrieval performance of 100 different trials. Successful retrieval was defined when overlap of corresponding input pattern converged to a value higher than 0.9.

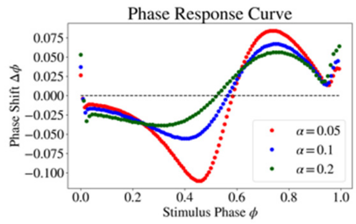


Fig. 4. Phase response curve of 3-variable DSSN model.

4.2. *Moderate neuronal interaction*

We observed chaotic transition within stored patterns when we set moderate neuronal interaction ($c = 0.05$). Figure 5a shows overlap dynamics with different α. Figure 5b is the staying period distribution without distinguishing at which pattern the retrieved pattern stayed. Maximum staying period was extended by decreasing α.

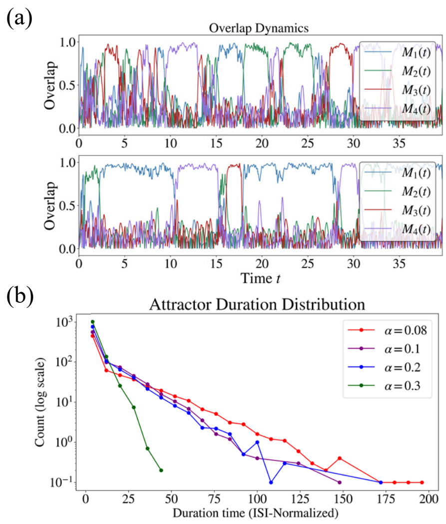


Fig. 5. (a) Overlap for $c = 0.05$. (Top) $\alpha = 0.1$, (Bottom): $\alpha = 0.07$. (b) Duration (normalized with ISI) distribution in retrieval processes. Averaged from 10 trials of $t = 1000$ simulation and 30 bins. Here, the retrieved pattern is defined to be staying at a specific pattern if the overlap exceeds 0.8.

We then introduced heterogeneity by grading each stored pattern with weight bias w_u. "Strong attractors" and "weak attractors" were created by the size of w_u. We set w_1 =1.03 and $w_{2,3,4}$=1.0 (pattern 1 corresponds to the strong attractor and the others weak). Then, we counted totally how often the retrieved pattern stayed in a specific pattern (overlap for the pattern exceeded 0.8) and observed the preference of attractor strength being modulated by α (Fig. 6). Strong preference of strong attractor is observed at $\alpha = 0.1$, whereas the preference is relaxed as α is decreased.

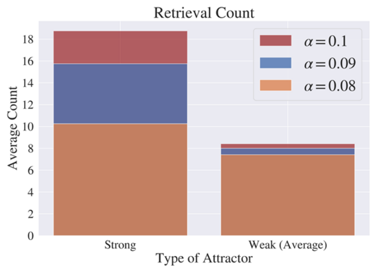


Fig. 6. Retrieval count of strong attractor and average count of weak attractors. Average computed from 40 trials.

4.3. *Transition of dynamical behavior*

The effects of c and α on the network's dynamical behavior is summarized in the following stability matrix (Fig. 7). Each value of the matrix is computed by calculating the gradient of the divergence between fiducial trajectories (vector of v, n, q, I_s per neuron) and perturbed trajectories to approximate the largest Lyapunov exponent of 1024 dimension dynamical system. We confirmed that the system is stable at small c, and is chaotic at larger c. Stability may also be altered by α where smaller value leads to more stable dynamics, although further analysis is needed.

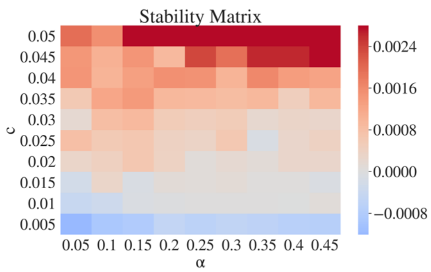


Fig. 7. Stability matrix computed from gradient of average divergence (log scale) between fiducial trajectories and 10 perturbed trajectories for each fiducial trajectory. Color bar shows gradient value.

5. Discussion

We constructed an all-to-all connected network composed of 3-variable DSSN model which exhibit SFA. We observed various network dynamics by modulating c and α, such as high and stable retrieval and chaotic transition between stored patterns. All-to-all connected networks are applicable also to solving optimization problems[7], and it is known that chaotic dynamics can be utilized to escape from local minima and obtain the global minima[8]. Similarly, our spiking neural network may be utilized to solve optimization problems by low power hardware in the future. All of the results in this work were obtained by floating-point operations on software simulation. One of the future works is to implement our models in fixed-point operations on FPGA devices.

Acknowledgements

This work was supported by JST SICORP Grant Number JPMJSC15H1, Japan.

References

1. T. Kohno and K. Aihara, "Digital Spiking Silicon Neuron: Concept and Behaviors in GJ-coupled Network," Proceedings of International Symposium on Artificial Life and Robotics (Beppu), 2007.
2. J. Li, Y. Katori, and T. Kohno, "An FPGA-Based Silicon Neuronal Network with Selectable Excitability Silicon Neurons," Frontiers in Neuroscience, vol. 6, no. December, pp. 1–13, 2012.
3. Kobayashi, W., Kohno, T., and Aihara, K. (2011). "3-variable digital spiking silicon neuron," in *Proceedings of the 24th Workshop on Circuit and Systems* (in Japanese) (Hyogo), 1–5.
4. J. P. Roach, E. Ben-Jacob, L. M. Sander, and M. R. Zochowski, "Formation and Dynamics of Waves in a Cortical Model of Cholinergic Modulation," PLoS Computational Biology, vol. 11, no. 8, pp. 1–16, 2015.
5. J. P. Roach, L. M. Sander, and M. R. Zochowski, "Memory recall and spike-frequency adaptation," Physical Review E, vol. 93, no. 5, pp. 1–6, 2016.
6. Ermentrout B, Type I membranes, phase resetting curves and synchrony (Neural Computation 8:979-1001, 1996).
7. J.J. Hopfield and D. W.Tank, "'Neural' computation of decisions in optimization problems", Biol. Cybern., vol.52, no. 1, pp.141-152, 1985.
8. Hiroshi Nozawa, "Solution of the optimization problem using the neural network model as a globally coupled map", *Physica D*, 75, pp179-189, 1994.

Appendix

Table 1. Parameter Set for 3-variable DSSN model.

Par.	Val.	Par.	Val.	Par.	Val.
a_{fn}	8.0	b_{fn}	-0.25	c_{fn}	-0.5
a_{fp}	-8.0	b_{fp}	0.25	c_{fp}	0.5
a_{gn}	4.0	b_{gn}	$-2^{-4}-2^{-5}$	c_{gn}	-0.77083333
a_{gp}	16.0	b_{gp}	$2^{-5}-2^{-2}$	c_{gp}	-0.6875
τ	2^{-9}	ϕ	0.625	r_g	-0.26041666
ϵ	0.03	v_0	-0.41	I_0	-0.09

Bioelectrical Signal Analysis of Mouse Cardiomyocyte Culture recorded on Thin-Film-Transistor Sensor Arrays

Anne-Claire Eiler
Research Center for Advanced Science and Technology, Graduate School of Engineering, The University of Tokyo, Japan

Junichi Sugita
Department of Cardiovascular Medicine, Graduate School of Medicine, The University of Tokyo, Japan

Satoshi Ihida
Institute of Industrial Science, Graduate School of Engineering, The University of Tokyo, Japan

Hiroshi Toshiyoshi
Institute of Industrial Science, Graduate School of Engineering, The University of Tokyo, Japan

Katsuhito Fujiu
Department of Cardiovascular Medicine, Graduate School of Medicine, The University of Tokyo, Japan

Timothée Levi
LIMMS/CNRS, Institute of Industrial Science, Graduate School of Engineering, The University of Tokyo, Japan

Agnès Tixier-Mita
Research Center for Advanced Science and Technology, Graduate School of Engineering, The University of Tokyo, Japan

E-mail: anne@iis.u-tokyo.ac.jp
www.u-tokyo.ac.jp

Abstract

The dynamical property of the heart bioelectrical system is closely associated with cardiac diseases. For this reason there is a growing interest in the development of system analysis for studying the cardiac signaling network. In this article, the electrical potentials of cardiac muscle cells have been measured on an array of microelectrodes using Thin-Film-Transistor (TFT) technology, and electrophysiological data were analyzed. This study shows the possibility of accurately analyzing extracellular signals measured on TFT arrays.

Keywords: Thin-Film-Transistor Arrays, Bioelectrical Signal Processing, Electrophysiology, Cardiomyocytes

1. Introduction

Cardiomyocytes are primary muscle cells derived from heart tissue that can generate and conduct bioelectrical signals for cardiac contraction and blood flow. A problem occurring in the cellular bioelectrical network can range from minor to fatal inconvenience.[1] The cells can retain their physiological functions and thus provide a useful *in vitro* model to look at the beating rate, the duration and the shape of the field potential. *In vitro* research of the general behavior of cardiomyocytes can help to understand arrhythmia, long Q-T syndrome, and cardiotoxicity. As a result, *in vitro* study of cardiomyocytes represents a valuable tool for drug discovery and cardiac research. However, the mechanism of the cardiac biosignaling network is still poorly understood.

In light of this problem, this paper proposes the analysis of bioelectrical signals of cardiomyocytes measured on an array of microelectrodes using a new Thin-Film-Transistor (TFT) sensor array. Measurements of the bioelectrical potentials of neurons were already demonstrated with TFT arrays.[2] The data flow generated by large arrays must be compressed to envision compact data acquisition systems. Hence, the electrical signals have been analyzed using a MATLAB program developed for bioelectrical processing of electrogenic cells. The recorded signals were filtered for the detection of spikes and grouped into clusters according to their similar features. Through this analysis, the experiments demonstrated the possibility of obtaining accurate spike sorting and analysis from extracellular recordings on TFT arrays.

2. Measurement Method

2.1. *Thin-Film-Transistor Arrays*

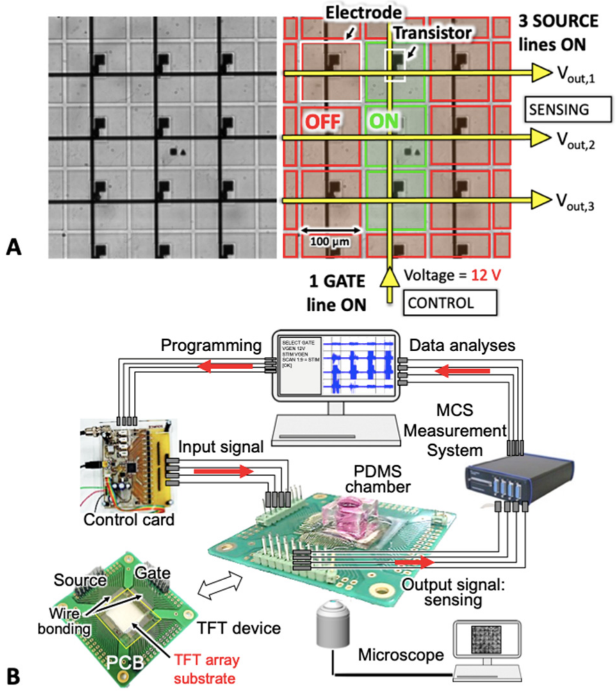
The TFT arrays were used for extracellular recordings *in vitro* of cardiomyocytes. TFT technology is well-known for Liquid Crystal Display in appliances including television sets, computer monitors, or mobile phones. Here, TFT technology is used for biological applications.[3]

The standard type of TFT array comes in a pattern of 150 x 150 transparent microelectrodes and is mounted on a printed circuit board (PCB). Microelectrodes are composed of indium tin oxide (ITO) with a size of 100 x 100 µm^2. The array of microelectrodes is controlled by an array of TFTs, which are used for switching ON/OFF the microelectrodes. The TFTs are controlled by means of gate and source/drain lines. The columns of the array control the gates of the TFTs, while the rows control the sources. When a 12V DC voltage is applied to one gate line, all the microelectrodes connected to that line are activated. Then one or more source lines are connected to a measurement system for sensing.

The bioelectrical signals of cardiomyocytes were measured using a Multi-Channel Measurement System (MCS USB-ME32-FAI-System) and optical observations were performed simultaneously with an inverted microscope. Fig. 1 describes the working principle of TFT arrays and the experimental setup.

Fig. 1. (A) Working principle of the transparent TFT substrate, with a close-up view of the array of 100 µm square microelectrodes; (B) Experimental setup.

2.2. *Culture of Cardiomyocytes*



Cardiomyocytes were dissociated into single-cell suspension from neonatal mice by combining mechanical dissociation with enzymatic degradation of the extracellular matrix, which maintains the structural integrity of tissues.

The neonatal hearts were enzymatically digested using a neonatal heart dissociation kit for mouse from Miltenyi Biotec and a dissociator was used for the mechanical dissociation steps. After dissociation, the sample was filtered to remove any remaining larger particles from the single-cell suspension, and red blood cell lysis was performed.

Cardiomyocytes were finally cultured for 3 days on the TFT array devices without surface treatment.

2.3. *Bioelectrical Analysis*

Embedded signal processing is an essential step in the development of recording instrumentation. Here, a spike sorting algorithm was used for data analysis. This data processing technique consists in identifying the cells that contribute to the signal recorded by each microelectrode, their number, and their spiking activities.[4] The identified basic functions are (1) bandwidth reduction for selective band amplification and noise reduction, (2) discrimination threshold computation, (3) extraction and alignment of biological spike signals, (4) data dimension reduction using principal component analysis (PCA) or spike shape features and finally (5) online spike clustering. Those functions are depicted in Fig. 2.

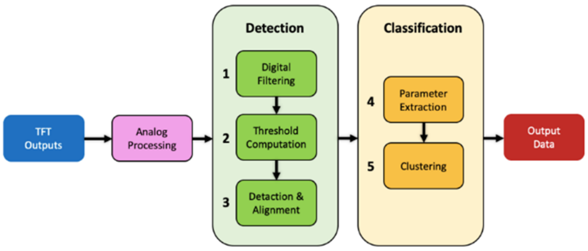


Fig. 2. Workflow of the functions performed for spike sorting.

3. Results

3.1. *Bioelectrical Signal Recording*

In this study, the extracellular electrical potentials of cardiomyocytes were first recorded on 28 x 28 microelectrodes. The measured noise level was approximately ±50 μV. The bioelectrical activity of the culture of cardiomyocytes was confirmed by optical visualization of cell contraction using an inverted microscope (Fig.3.). Here, a line of 4 microelectrodes was selected and data was extracted for the ensuing processing of the bioelectrical signals.

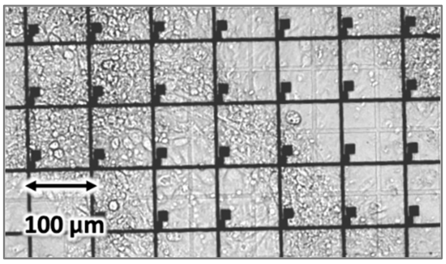


Fig. 3. Culture of cardiomyocytes on TFT array device, 3 days after cell seeding.

3.2. *Bioelectrical Signal Processing*

Bioelectrical signal processing of the data acquired from the measurements was performed. Each analysis provided valuable information about the bioelectrical signals of the cardiomyocytes. In this paper, the measurement data of a decrease in temperature has been analyzed.

3.2.1. Filtered Data

Raw data were first filtered to remove undesired signals according to their frequency. This low-pass filter passes signals with a frequency lower than the selected cutoff frequency of 200 Hz. The sampling rate of the recorded data was 10 kHz. Fig. 4. shows an example of data obtained before and after filtering.

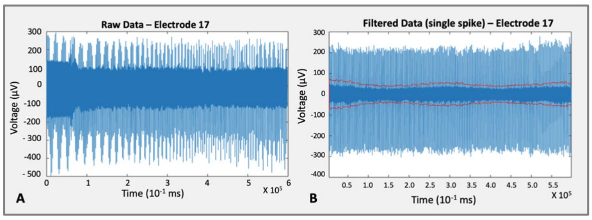


Fig. 4. (A) Raw data of the total measurement time (7 min) and (B) close-up view of a single spike on filtered data.

3.2.2. Spike Detection and Alignment

The action potentials of cardiomyocytes were distinguished from the noise according to a predefined threshold. The spikes were then aligned with respect to the maximum absolute value of the detected signal as shown in Fig. 5.

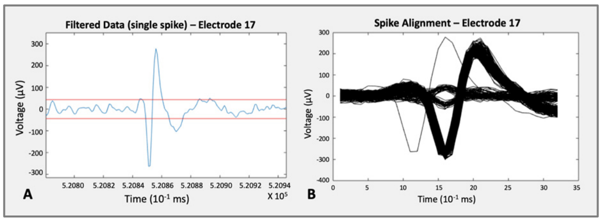


Fig. 5. (A) Filtered data and threshold lines in red for spike detection, and (B) spike alignment for 1 microelectrode.

3.2.3. Spike Intervals

For each microelectrode, the spike intervals were then classified. Fig. 6 shows the histograms that display the number of spikes according to time interval.

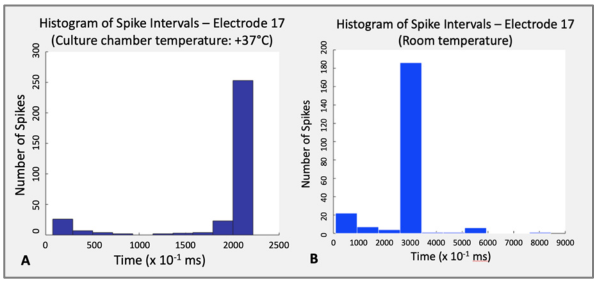


Fig. 6. (A) Histogram of the spike intervals detected on 1 microelectrode when the temperature of the culture chamber is at +37°C and (B) at room temperature (RT). In this study, the spike intervals went from 200 ms (5.0 Hz) at +37°C to 270 ms (3.7 Hz) at RT.

3.2.4. Clustering

Spikes were finally divided into clusters. To reduce sample space dimension, Principal Component Analysis (PCA) was used. This statistical technique provides features that are directions in the high-dimensional space. In this study, each spike became a point in a 3-dimension space. Fig.7 shows the mean spikes for each cluster.

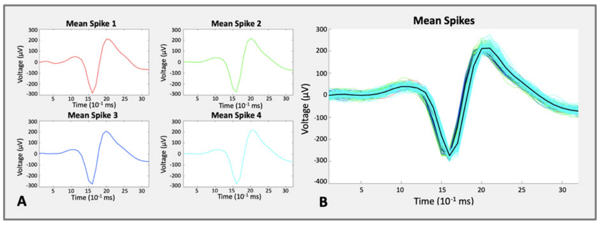


Fig. 7. (A) Mean spikes of each 4 clusters identified after analysis of 4 microelectrodes, and (B) alignment of the mean spikes of each cluster.

4. Discussion

The transparent TFT array device can provide a high spatial resolution of cell culture activity by surpassing currently available MEAs in terms of density of microelectrodes, and its cm-sized measurement surface area. With regard to the bioelectrical analysis, the spike sorting program allowed the extraction of useful information about the bioelectric signals generated by cardiomyocytes. Decrease of the beating rate with temperature (from +37°C to room temperature) was observed during the data acquisition with MCS measurement system. This observation is confirmed by the program with a decrease rate of around – 6.5 beats per minute/°C for each microelectrode. A possible explanation would be that the temperature drop depresses the speed of ion exchange, as it increases the permeability of the membrane to ions. The program also identified a strong irregularity of the peak-to-peak voltage amplitude with the decrease of temperature. Although one cluster of spikes is identified for each microelectrode, they share an equivalent waveform of the mean spikes. This waveform is highly similar to the typical cardiac action potential in electrocardiogram with the typical T wave associated with ventricular repolarization. A spike raster plot also revealed the spike synchronicity between each microelectrode. This confirms the synchronicity of the bioelectrical conduction among the cell culture, which was also observed in parallel with the microscope.

5. Conclusion

In this paper, the bioelectrical activity of cardiomyocytes was successfully measured on TFT sensor arrays and analyzed by using a spike sorting technique. This analysis confirms that TFT arrays can efficiently detect the bioelectrical signals generated by cardiomyocytes. Combining this technique with deep-learning algorithms[5] could allow the *in vitro* identification of abnormal cardiac cell conduction and aid for drug screening.

References

1. D. Farwell and M. H. Gollob, "Electrical heart disease: Genetic and molecular basis of cardiac arrhythmias in normal structural hearts," *Canadian Journal of Cardiology*, vol. 23, 2007.
2. R. F. A. Shaik, Y. Ikeuchi, G. Cathcart, S. Ihida, H. Toshiyoshi and A. Tixier-Mita, "Extracellular neural stimulation and recording with a Thin-Film-Transistor (TFT) array device," *2017 19th International Conference on Solid-State Sensors, Actuators and Microsystems (TRANSDUCERS)*, Kaohsiung, 2017, pp. 206-209.
3. Tixier-Mita, S. Ihida, B.-D. Ségard, G. A. Cathcart, T. Takahashi, H. Fujita, and H. Toshiyoshi, "Review on thin-film transistor technology, its applications, and possible new applications to biological cells," *Japanese Journal of Applied Physics,* vol. 55, no. 4S, 2016.
4. J.-F. Beche, S. Bonnet, T. Levi, R. Escola, A. Noca, G. Charvet, and R. Guillemaud, "Real-time adaptive discrimination threshold estimation for embedded neural signals detection," *2009 4th International IEEE/EMBS Conference on Neural Engineering*, 2009.
5. A. Mincholé and B. Rodriguez, "Artificial intelligence for the electrocardiogram," *Nature Medicine*, vol. 25, no. 1, pp. 22–23, 2019.

Sea trials for benthos sampling using autonomous underwater vehicle

Yuya Nishida
Khushu Institute of Technology, 2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196, Japan

Shinsuke Yasukawa
Khushu Institute of Technology, 2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196, Japan

Takashi Sonoda
Nishinippon Institute of Technology, 1-11 Aratsu, Kanda, Miyako, Fukuoka 800-0394, Japan

Keisuke Watanabe
School of Marine Science and Technology, Tokai Univ., 3-20-1 Orido Shimizu-ku, Shizuoka, 424-8610, Japan

Kazuo Ishii
Khushu Institute of Technology, 2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka 808-0196, Japan
E-mail: ynishida@lsse.kyutech.ac.jp, ishii@brain.kyutech.ac.jp

Abstract

For enhancement of oceanic engineering technology and researchers, underwater robot competition has been held since 2016. Seventh competition in this year consists AUV league that university's vehicles automatically cruise at field and junior league that underwater craft is made. The paper reports competition regulations of AUV and junior league and results of the competition held in October 2019.

Keywords: Autonomous Underwater Vehicle, Benthos sampling, suction device

1. Introduction

Autonomous underwater vehicles (AUVs) without umbilical cable cruise a wide area and is often used for bio-resource survey [1-2]. In the survey, AUVs take seafloor images for measurement of benthos biomass and distribution using a mounted camera and strobe. However, bio-resource survey by AUVs is difficult to obtain the sample useful for scientists. Previous AUVs can capture the specific benthos that its developer decided the target before resource survey, but scientists on the ship can't chose sampling target during survey. To realize an efficient bio-resources survey by AUVs, the authors developed the AUV Tuna-Sand2 with an epochal sampling method that previous problem is solved, as shown in Fig.1 [3-4]. The AUV succeeded to capture a shell on the seafloor of 100 m depth that the operator selected as sampling target during observation in March 2018 [5], but the AUV was able to only sample one time because its suction device may drop with suction more than twice. The paper explains the system of the AUV Tuna-Sand2 and new suction device developed for continuous sampling, results of experiments performed with the AUV in the ocean are shown.

Fig.1 The AUV Tuna-Sand2

Table 1 The specification of the AUV Tuna-Sand2

Category	Value
Dimension	(L)1.4 m x (W)1.2 m x (H)1.3 m
Weight	390 kg
Max. speed	1.2 kt (0.6 m/s)
Max. depth	2,000 m
Duration	8.0 h
Actuators	Horizontal thrusters x 4, Vertical thrusters x 2 Ballast releaser x 2 The arm with a suction device x 1
Battery	Li-Ion 5,000 Wh
Communications	Wireless LAN, Acoustic modem for command link Acoustic modem for image data
Sensors	INS, DVL, Depth sensor, Transponder for SSBL positioning
Payloads	LED array x 2 LED strobe x 2 Scanning sonar x 1 3D mapping device x 1 Camera for benthos sampling x 1

2. The hovering type AUV Tuna-Sand2

2.1. *System*

Hovering type AUV Tuna-Sand2 can dive up to 2,000 m depth and operate for up to 8 hours using a mounted Li-ion battery. Table 1 shows the specification of the AUV Tuna-Sand2. The AUV estimates self-position based on altitude and ground velocity from doppler velocity log (DVL), attitude and heading from inertial navigation system (INS) and depth sensor data. The transponder is installed on the top of the AUV, and the operator on the support ship can measure the global position of the AUV under the ocean using a positioning device based on super short base line (SSBL). The obstacle is detected by scanning sonar mounted on the front of the AUV, and the AUV increases altitude if the sonar detects the obstacle which 5 m. As observation device, the AUV has a mapping device for 3D reconstruction of the seafloor and a camera for benthos sampling.

2.2. *Survey procedure*

The AUV Tuna-Sand2 acts sequentially based on the state flow as shown in Fig.2. After deployment in the ocean, the AUV which has two ballasts dives by its

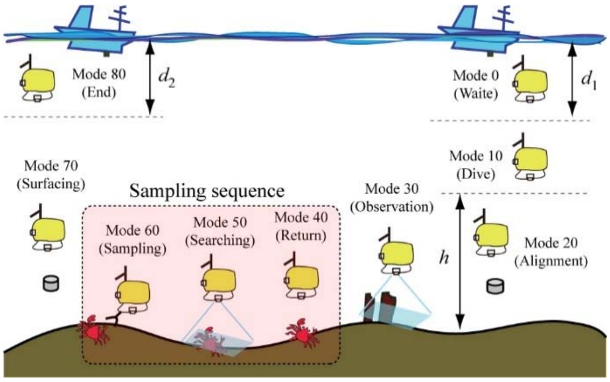


Fig.2 Survey procedure based on the state flow

weight to the seafloor (Mode 0). During the dive, mounted sensors and payloads is turned on for navigation and observation (Mode 10). The AUV becomes neutral buoyancy by releasing a ballast and camera parameters are adjusted while taking a seafloor image, when it closes to the ocean (Mode 20). Three minutes after reaching the seafloor, the AUV navigates at a constant velocity and altitude along preset waypoints and takes image of the seafloor using a camera and two LED arrays every five seconds (Mode 30). Then, the benthos candidate is searched by saliency map that is the same as human attention model from the seafloor image enhanced by the Retinex method. If the candidate is detected, reference image of the candidate is made for tracking, and the candidate and its surrounding images which data size reduced are transmitted to the support ship using a special acoustic modem. The operator on the ship selects sampling target from received images and instructs the image number to the AUV by acoustic modem for command link. After received operation instruction, the AUV returns to the points where the image of the instructed number was taken. Sampling target is searched and tracked by pattern marching method based on the reference image of the candidate. If the target is found, the AUV approaches and captures the target by mounted suction device. After the sampling sequence, the AUV waits a few minutes for next instruction from the operator while keeping the safe altitude. Then, if the instruction is received, the AUV runs the sampling sequence again. If there is no instruction or there is an instruction of mission abort, the AUV becomes positive buoyancy by releasing a ballast and goes up to the surface.

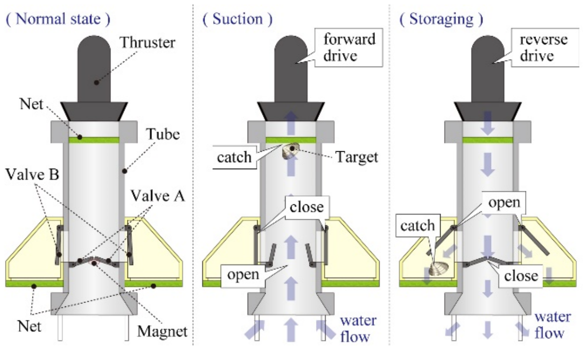


Fig. 3 Suction device for continuous sampling

2.3. *Suction device*

Previous suction device consists of a tube with check values and a thruster, and it may drop captured benthos with suction more than twice because there is no canister. To solve the problem, a suction device which has two canisters and two valves with different functions was developed as shown in Fig.3. If the benthos is wanted to catch, the device sucks the target with the water by a forward drive of the thruster. Then, the valve A opens by the water flow, and the net near the thruster catches the target. The valve A closes by its weight and the force of the magnet when the thruster is stopped, the target puts on the valve A. After that, the valve B opens by a revers drive of the thruster still with the valve A closed, and the target moves into the canister. Because the valve B closes during suction, the device continuously sucks the sampling target without dropping out.

3. Sampling experiments

To experiment for continuously sampling, the AUV Tuna-Sand2 with a developed suction device was deployed at 120 m depth in the Suruga Bay in December 2019. Figure 4 shows the trajectory of the AUV during cruising, and time series data of the AUV position during the sampling sequence is shown in Fig.5. The AUV cruised at 2.0 m/s velocity to 1st waypoint after reaching the seafloor and observed the seafloor of about 120 m^2 by the camera while keeping 1.0 m/s velocity and 1.5 m altitude. Then, 85 benthos candidates were detected by image processing from taken seafloor image and were transmitted to the support ship by using the acoustic modem. The operator on the ship received only half of the transmitted images, and three sampling targets were selected by the operator from the images. The AUV tried

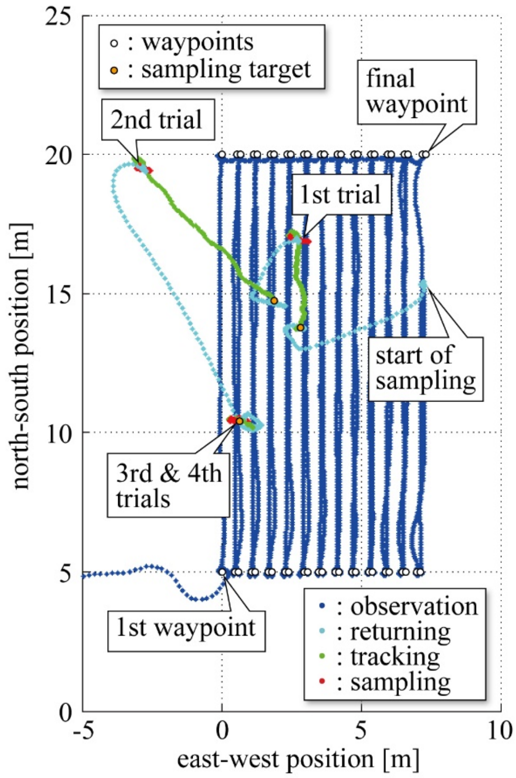


Fig.4 The trajectory of the AUV during cursing

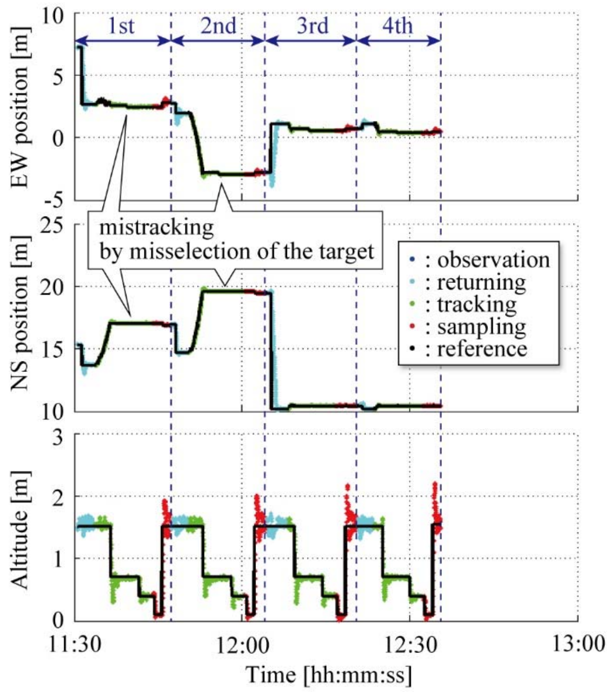


Fig.5 The position data during the sampling sequence

Fig. 6 Referenced images in the first and second trials

capturing different targets for the first and second trials, but the AUV couldn't search the targets and tracked the wrong targets. The cause is that the targets instructed by the operator were marine snows shown in Fig.6. Because the received images were the images with low resolution and color depth, the operator was difficult to search from the benthos candidates. In third and fourth trials, the AUV tried to capture the same benthos that stuck the seafloor like the sea anemone and didn't sample the benthos because the suction force of the device wasn't so high. The AUV controlled the position with high accuracy of ±0.05 m error from the references and tracked the sampling target in third and fourth trials.

4. Conclusions

To solve the problem of the previous method, the authors developed the benthos sampling method by the AUV with the suction device that is able to suck several times. In the sea trial using the method, the AUV could try to sample four times according the instruction of the operator on the ship, but the benthos couldn't be captured due to the mis-selection of the sampling target and weak force of the suction device. In future work, the authors will develop user-friendly image transmission method and the device which can suck the benthos sticking the seafloor

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 18K04585. The authors deeply grateful the crews of Tokai University's research vessel HOKUTO and the students of Watanabe laboratory who helped the sea trials.

References

1. Yuya Nishida, et al., Resource investigation for kichiji rockfish by autonomous underwater vehicle in Kita-Yamato bank off Northern Japan, ROBOMECH Journal, Vol.1, pp.1-6, 2014
2. Blair Thornton, et al., Biometric assessment of sea vent megabenthic communities using multi-resolution 3D image reconstructions, Deep Sea Research Part I: Oceanographic Research Papers, Vol.116, pp.200-2190, 2016
3. Yuya Nishida, et al., Underwater Platform for intelligent robotics and its application in two visual tracking systems, Journal of Robotics and Mechatronics, Vol.30, No.2 pp.238-247, 2018
4. Shinsuke Yasukawa, et al., Vision system for an autonomous underwater vehicle with a benthos sampling function, Journal of Robotics and Mechatronics, Vol.30, No.2 pp.248-256, 2018
5. Yuya Nishida, et al., Benthos sampling by autonomous underwater vehicle equipped a manipulator with suction device, Proc. of IEEE Underwater Technology, DOI: 10.1109/UT.2019.8734330, Taiwan, 2019

Field Experiments of Underwater Image Transmission for AUV

Shinsuke Yasukawa
Dept. of Intelligence Systems, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka, 808-0196, Japan
E-mail: s-yasukawa@brain.kyutech.ac.jp
Yuya Nishida
Dept. of Intelligence Systems, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka, 808-0196, Japan
E-mail: ynishida@lsse.kyutech.ac.jp
Jonghyun Ahn
Department of Intelligent Mechanical Engineering, Hiroshima Institute of Technology,
2-1-1 Miyake,Saeki-ku, Hiroshima, 731-5193, Japan
E-mail: j.ahn.h2@it-hiroshima.ac.jp
Takashi Sonoda
Dept. of Integrated Systems Engineering, Nishinippon Institute of Technology,
1-11 Aratsu, Kanda, Miyako, Fukuoka, 800-0344, Japan
E-mail: sonoda@nishitech.ac.jp
Kentaro Yanagise
Dept. of Intelligence Systems, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka, 808-0196, Japan
E-mail: yanagise.kentaro953@mail.kyutech.jp
Keisuke Watanabe
Dept. of Navigation and Ocean Engineering, Tokai University,
3-20-1 Orido Shimizuku, Shizuoka, Shizuoka, 424-0902, Japan
E-mail: keisukejapan@gmail.com
Kazuo Ishii
Dept. of Intelligence Systems, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu, Kitakyushu, Fukuoka, 808-0196, Japan
E-mail: ishii@brain.kyutech.ac.jp

Abstract

In order to improve the efficiency of the seafloor survey, it is necessary for AUV to report the state of the seafloor to the operators on board reasonably. We have been developing seafloor image selection and image compression technology for the seafloor image transmission using the underwater acoustic communication device. In this paper, we report the results of underwater image selection and transmission in biological sampling experiments conducted in November 2019 off the coast of Suruga-bay, Shizuoka, Japan.

Keywords: underwater image processing, sampling-AUV, image selection, field robot, seafloor survey.

1. Introduction

m the viewpoint of marine biodiversity conservation, it is important to improve marine survey techniques. Autonomous Underwater Vehicle (AUV) has enough advantage in the task of observing a wide range area [1-3], however, it is difficult to achieve a task that requires high adaptability and recognition such as biological sampling. In order to improve the efficiency of the survey, it is necessary for AUV to report the state of the seafloor to the operators on board reasonably and in real time. In order to realize such new seafloor observation, the seafloor image transmission using the underwater acoustic communication device is the one of the solutions. We have been developing seafloor image selection and image compression technology for acoustic transmission [4, 5]. In this paper, we report the results of underwater image selection and transmission in biological sampling experiments conducted in November 2019 off the coast of Suruga-bay, Shizuoka, Japan.

2. Mission and underwater image transmission

An outline of the mission for sampling Benthos are described. Missions of sampling-AUV are mainly divided into the observation phase, return phase, tracking phase, and sampling phase. Sampling-AUV operates in different modes in each phase. Sampling-AUV comprised of a suction device, a vision system, an acoustic communication system for image transmission, and a hovering-type AUV "Tuna-sand2"[6]. In observation mode, the sampling-AUV moves while maintaining a constant altitude. It photographs the seafloor with a downward camera. The vision system selects an image with an interesting object and compresses the image. The compressed image is transmitted to the ship using the acoustic communication system (see [4, 5] for details). In the return mode, the operator on the ship specifies the selected image including the sampling target. The robot moves to the location where the selected image was taken. In the sampling mode, the robot performs image-based object tracking and sampling of an interesting object.

3. Field Experiments and results

We conducted biological sampling experiments in November 2019 off the coast of Suruga-bay, Shizuoka,

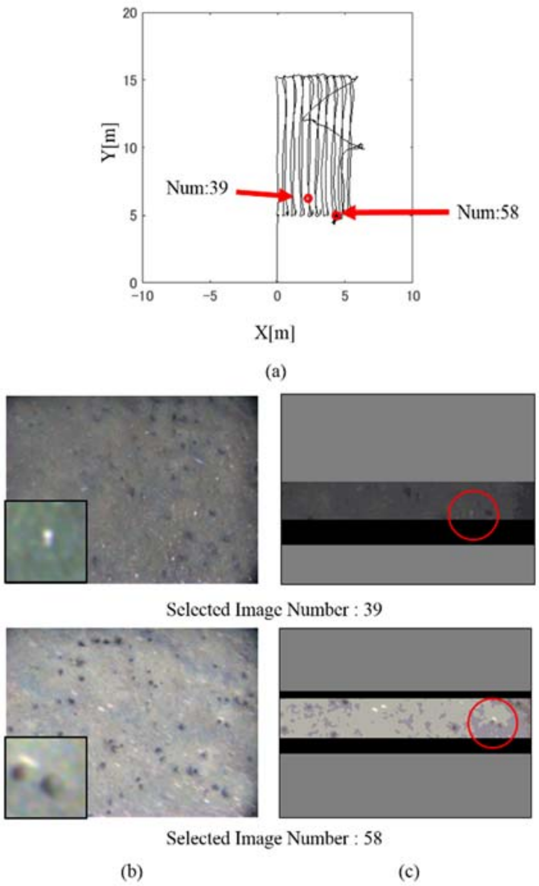


Fig. 1 Results of the first dive (a) Robot trajectory and image capturing position (b) selected images (c) received images.

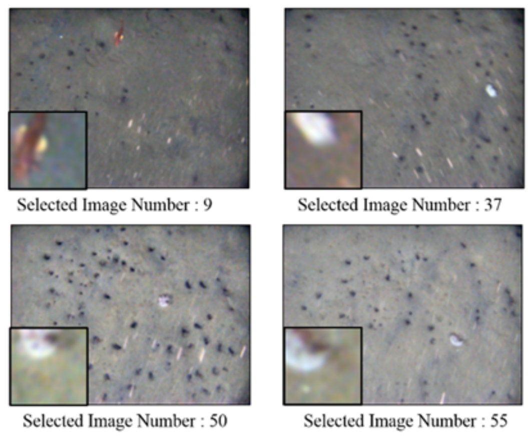


Fig. 2 Selected images not received on board (first dive).

Japan. In this paper, we report the results of image selection and acoustic communication of three trials. Fig. 1 shows the result of the first dive. A total of 62 images of interest were acquired. At the first dive point, the robot captured various underwater objects. However, few selected images were received on board the ship (Fig. 2). At the second dive point, the robot shot the creatures on the seafloor as shown in Fig. 3. Many of the received images selected marine snow (the middle image in Fig. 3), but received images of marine creature (The bottom image in Fig. 3). In this dive, the robot was able to attempt biological sampling. In the third dive, many images selected marine snow more (Fig. 4). The robot also selected creatures (Fig. 5), but no captured images of the creatures were received on board.

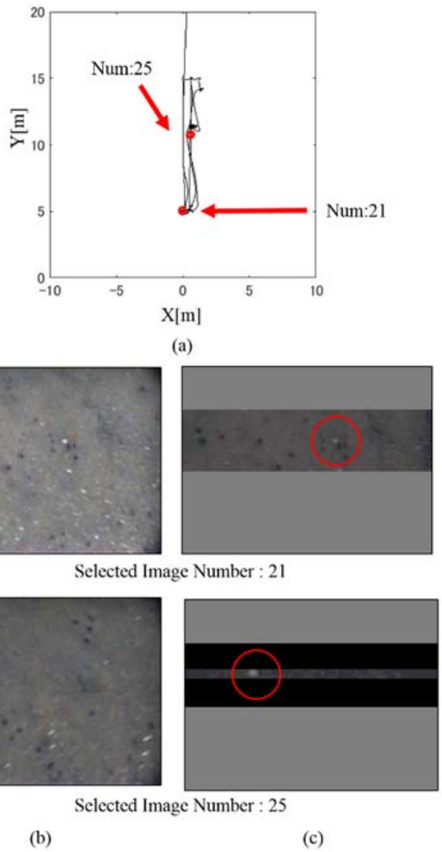


Fig. 3. Results of the second dive (a) Robot trajectory and image capturing position (b) selected images (c) received images.

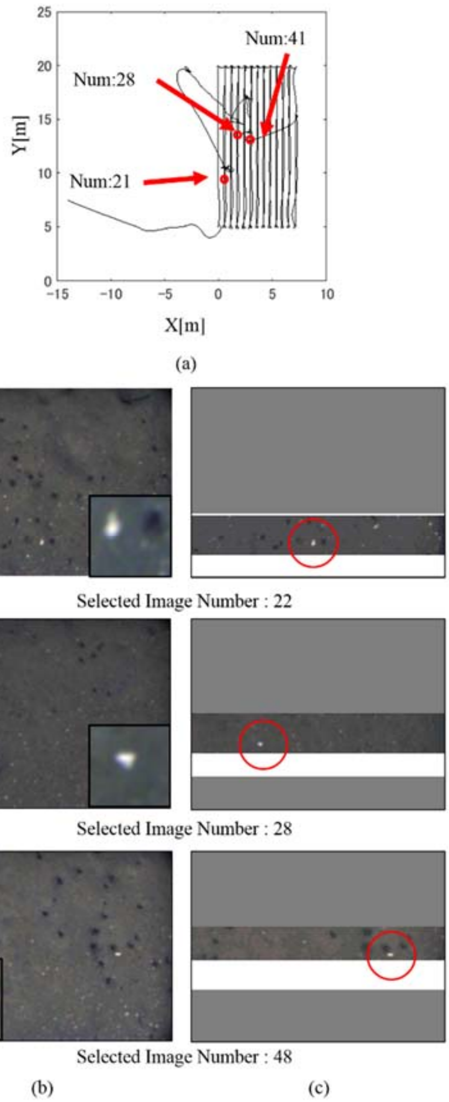


Fig. 4. Results of the third dive (a) Robot trajectory and image capturing position (b) selected images (c) received images.

Fig. 5. Selected images not received on board (third dive).

4. Conclusion

We reported on three field experiments in Shizuoka in November 2019. Underwater creatures were selected by the vision system, but marine snow was often selected. It is difficult to distinguish between marine snow and living things by the image selection algorithm using visual attention based on the conventional bottom-up approach. Therefore, as a future work, we will develop an image selection algorithm that adds a top-down element.

References

1. Whitcomb, L. L. "Underwater robotics: Out of the research laboratory and into the field," In Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation, Vol. 1, pp. 709-716, 2000.
2. Johnson‐Roberson, M., Pizarro, O., Williams, S. B., & Mahon, I.,"Generation and visualization of large‐scale three‐dimensional reconstructions from underwater robotic surveys", Journal of Field Robotics, Vol. 27, No. 1, pp. 21-51, 2010.
3. Johnson‐Roberson, M., Bryson, M., Friedman, A., Pizarro, O., Troni, G., Ozog, P., & Henderson, J. C., "High‐resolution underwater robotic vision‐based mapping and three‐dimensional reconstruction for archaeology", Journal of Field Robotics, Vol. 34, No. 4, pp. 625-643, 2017.
4. Ahn, J., Yasukawa, S., Sonoda, T., Nishida, Y., Ishii, K., & Ura, T., "An Optical Image Transmission System for Deep Sea Creature Sampling Missions Using Autonomous Underwater Vehicle", IEEE Journal of Oceanic Engineering, 2018. DOI: 10.1109/JOE.2018.2872500
5. Yasukawa, S., Ahn, J., Nishida, Y., Sonoda, T., Ishii, K., & Ura, T., "Vision system for an autonomous underwater vehicle with a benthos sampling function", Journal of Robotics and Mechatronics, Vol. 30, No. 2, pp. 248-256, 2018.
6. Nishida, Y., Sonoda, T., Yasukawa, S., Ahn, J., Nagano, K., Ishii, K., & Ura, T., "Development of an autonomous underwater vehicle with human-aware robot navigation", In OCEANS 2016 MTS/IEEE Monterey, pp. 1-4, 2016.

Development of Subsea Creature Monitoring Station for AUV Exploration Assistance

Keisuke Watanabe
School of Marine Science and Technology, Tokai Univ., 3-20-1 Orido Shimizu-ku Shizuoka, 424-8610, Japan

Koshi Utsunomiya
School of Marine Science and Technology, Tokai Univ., 3-20-1 Orido Shimizu-ku Shizuoka, 424-8610, Japan

Kazumasa Harada
School of Marine Science and Technology, Tokai Univ., 3-20-1 Orido Shimizu-ku Shizuoka, 424-8610, Japan

Amir Sadiq
School of Marine Science and Technology, Tokai Univ., 3-20-1 Orido Shimizu-ku Shizuoka, 424-8610, Japan

Kazuo Ishii
Center for Socio-Robotic Synthesis, Kyusyu Institute of Technology, 2-4 Hibikino Wakamatsu-ku Kita-Kyusyu, 808-0196, Japan
E-mail: keisuke@tsc.u-tokai.ac.jp, ishii@brain.kyutech.ac.jp
www.scc.u-tokai.ac.jp/~keisuke/index.html

Abstract

In this paper, we introduce our recent result on developing a low cost monitoring station. This station is intended to monitor underwater images of subsea creatures wirelessly from a boat in advance before an exploration AUV is put into water. We describe the concept, the design and system integration of the station in the paper. We also present an experimental result that was carried out to test the implemented functions at Suruga bay at 65m water depth.

Keywords: Subsea creature monitoring, Independent device, Real time monitoring, AUV assistance, Field experiment.

1. Introduction

As the global climate change is ongoing, ocean environment seems also changing rapidly[1] and it is often said that this change damages ocean eco system, which impacts variety of creatures. For example, coral bleaching is progressing worldwide and it becomes a big issue for people in those area because their food supply or economic activity like tourism is being damaged. In Suruga bay in Shizuoka prefecture, a unique shrimp called “Cherry shrimp” was very famous and very important income resource for fishermen. However, recently the amount of this shrimp catch decreases rapidly, whose cause has not been found clearly because no data accumulation underwater has been tried.

Considering these background, we are developing a set of sea creature investigation systems which consists of

AUVs, a support vessel and subsea monitoring stations[2]. In this paper, we introduce our recent result on developing a low cost monitoring station. The similar real time monitoring system using underwater camera has been developed and used effectively by biologists, for example, as shown in the Ref 3. Our station is cheaper and intended to monitor underwater images of subsea creatures wirelessly from a boat in advance before an exploration AUV is put into water. As AUV images cannot be checked until it returns, we need information where we should put an AUV. The station consists of a cage on the seabed, a float which relays underwater images to a support vessel and an umbilical cable from the cage to the float. We describe the concept, the design and system integration of the station in the paper. We also present an experimental result that was carried out to test the implemented functions at Suruga bay at 65m water depth.

2. The system concept

We have carried out many sea experiments to try to catch sea creatures for the purpose of investigation of underwater environment using the AUV "Tuna Sand 2" at Suruga bay for these years[4,5]. "Tuna Sand 2 (TS-2)" can send compressed subsea pictures from underwater to the surface support vessel through ultrasonic acoustic communication link. However, through our experience, we found lower cost underwater survey system in advance of diving TS-2 is effective. As the weight of TS-2 is around 400kgf, so its deployment from our support vessel on the sea is not an easy task, whereas sometimes the diving area found not abundant after TS-2 came back. It is a kind of AUV drawback that we cannot monitor its surrounding situation until it comes back and we transfer images from it. To make AUV performance more efficient, we need to know underwater image in advance whether the area is worth diving with the AUV.

The schematic view of our concept is shown in Fig.1. As shown in Fig.1, the device consists of two main parts of a float and a subsea camera module. On the float, a wireless LAN module is put so that a researcher on the support vessel can monitor the real time subsea images. We decided to put a cage on the seabed to collect sea creatures using bait. This cage is also an anchor weight and a structure where cameras and lights are attached. The cameras are wired to the float so the wire must be strong enough to bear its tension as well as connectors

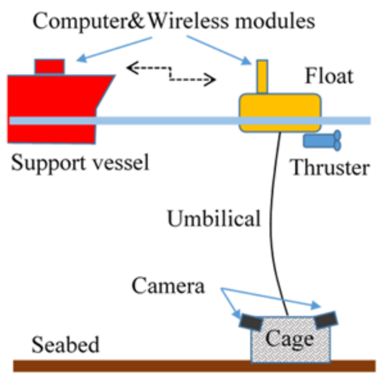


Fig.1. Schematic view of the monitoring system concept.

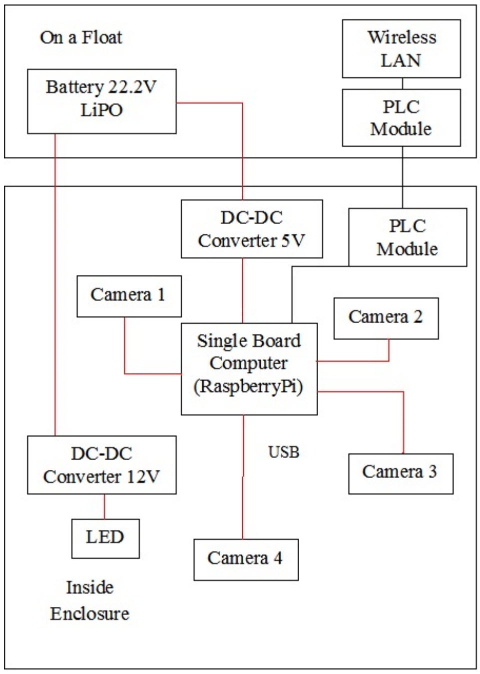


Fig.2. System architecture and devices.

must be bearable to the tension by the wire. The mechanical design which is applicable to sea experiments is very important considering its handling such as throwing or collecting on the rolling vessel as well as ocean current which drags the cable.

3. The system architecture

The system architecture is shown in Fig.2. As the size of the pressure hull is limited, we designed the battery is arranged on the float. The length of the power and signal

cable is around 100m and its resistance becomes around 3Ω. The maximum current is more than 2A only when the board computer starts but after that the current becomes less than 1A. Considering the voltage drop at the end of the cable and current fluctuation, the DC-DC was selected to allow wide range input voltage.
Four cameras were installed. One camera is facing to the sea bottom and the others are facing horizontally at each vertex of equilateral triangle as shown in Fig.3 and Fig.4. Each camera image can be monitored through LAN signal from the board computer through PLC modules.
LED module is shown in Fig.3. LED module is arranged near each camera respectively. It can be turned on/off from the PC on board.

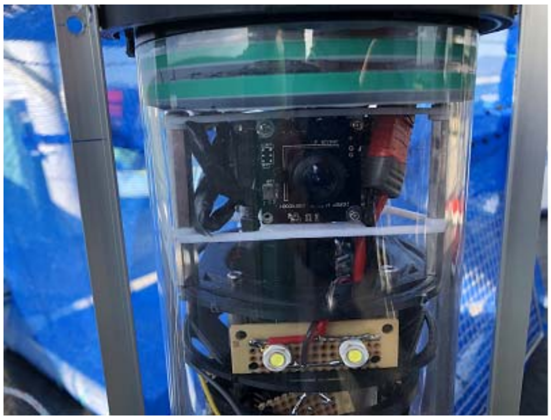


Fig.3. Horizontal camera

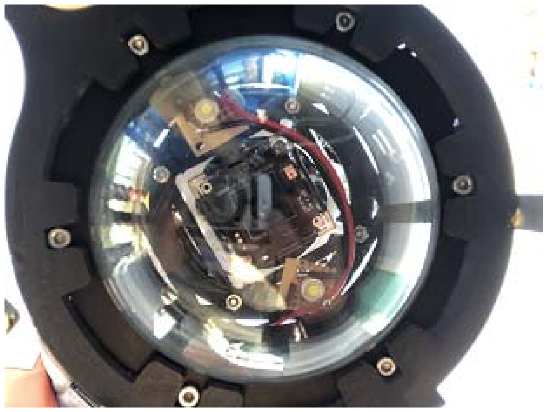


Fig.4. Horizontal camera

4. Sea experiment

Sea experiment was carried out on 16th December 2019. The experimental site is shown in Fig.5. Tokai Univ.'s research vessel "Hokuto" is deployed for this experiment. Fig.6 shows the cage, the cable and the float. The cage is suspended from the vessel using a rope to support tension. While the cage is descending, we pay out the cable with handling not to exert tension to the cable. The float is released finally from the vessel, then the vessel can be away from the monitoring device with no fear of winding the cable or wire into its thruster, which makes the operation safer.
Fig.7 shows the structure of the cage. The cage has opening to lure creatures. The bait is put on its lower floor and the cage has a structure that once a creature is coming into the bait floor it is not easy for the creature to go back to its entrance.
Fig.8 shows LED in the pressure hull attached to the cage. This picture was taken by an action camera which is set inside the cage.
Fig.9 shows the captured images of four cameras' video after the cage was reached on the seafloor. We confirmed some planktons like krill near the seabed real time on board as shown in Fig.9.

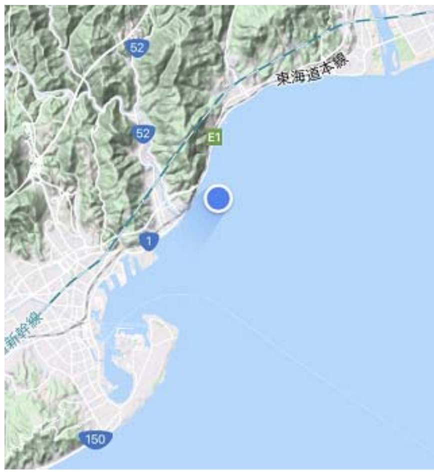


Fig.5. Site of sea experiment

5. Summary

We designed and developed a low cost monitoring station as a prototype. We succeeded to monitor real time underwater image on board. In this experiment, we found the drag force of the cable is very strong as is considered before the experiment. The cable length was fixed as

Fig.6. The cable, cage and float.

Fig.7 The cage includes bait.

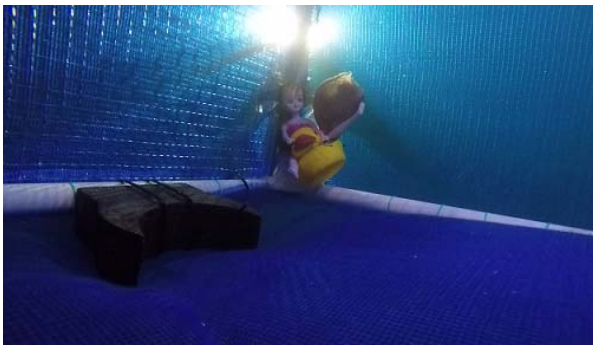


Fig.8. Bottom camera

around 100m and considering its curved shape, the water depth 65m was selected. We need to improve to be able to change the cable length according to the water depth. This time the cable was pulled out due to strong current.

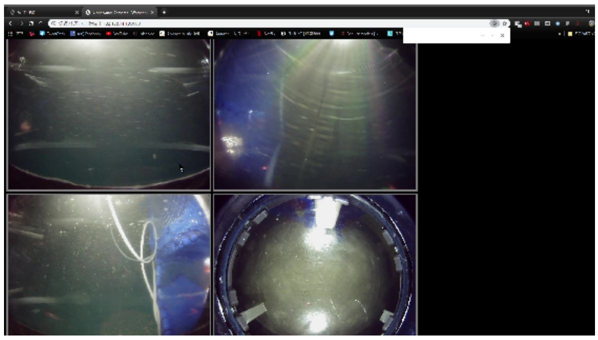


Fig.9. Bottom camera

So the experiment time was not enough to catch creatures. In our next work will be the float should be modified to an autonomous surface vehicle and the cage should be modified to a subsea crawler, which enables longer investigation and monitoring at deeper site.

Acknowledgements

The authors deeply grateful to Captain I.Hashigaya and Chief engineer H. Sakata of Tokai Univ's research vessel HOKUTO as well as those students of our laboratory who helped the experiment, K. Kitazawa, D. Hiramaki. Part of this work was supported by JSPS KAKENHI Grant Number 18K04585.

References

1. W.W.L.Cheung, R. Watson and D. Pauly, Signature of ocean warming in global fisheries catch, *Nature,* Vol. 497, 2013, pp.364-369, doi:10.1038/nature12156
2. K. Watanabe, K. Mitsumura, K. Utsunomiya, K. Harada, An Experiment of Multi-copter and Small ROV Combined Monitoring System for Coastal Area Environment, Proc.of *the 28th International Ocean and Polar Engineering Conference,* 2018, pp.373-378
3. D. J. Mills , G. Verdouw and S. D. Frusher, Remote multi camera system for in situ observations of behavior and predator/prey interaction of marine benthic macrofauna, *New Zealand Journal of Marine and Freshwater Research*, 2005, Vol.39(2), pp.347-352, DOI:10.1080/00288330
4. Y. Nishida, T. Sonoda, S. Yasukawa, J. Ahn, K. Watanabe and K. Ishii, Benthos Sampling by Autonomous Underwater Vehicle Equipped a Manipulator with Suction Device, *IEEE Xplore Underwater Technology*, 2019, DOI10.110/UT.2019.8734330
5. S. Yasukawa, Y. Nishida, J. Ahn, T. Sonoda, K. Watanabe and K. Ishii, Development and Sea Trials of Vision-Based Control for Sampling-AUV, *IEEE Xplore Underwater Technology*, 2019, DOI10.110/UT.2019.8734462

Environment map generation in forest using field robot

Noboru Takegami
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

Ryusuke Fujisawa
Department of Mechanical Information Science and Technology, Kyushu Institute of Technology
680-4, Kawazu, Iizuka-City, Fukuoka, 820-8502, Japan
E-mail: takegami@mmcs.mse.kyutech.ac.jp , haya@.mse.kyutech.ac.jp, fujisawa@.ces.kyutech.ac.jp
http://www.kyutech.ac.jp/

Abstract

The employment rate of young people in Japanese forestry tends to decline, and the unmanaged forest area is expected to increase in the future. Therefore, in our laboratory we propose an autonomous field robot with all terrain vehicles that focuses on the automation of work. The robot automates weeding and observation in the forest. In this research, we observed trees by generating an environmental map in the forest using Simultaneous Localization and Mapping (SLAM). The error of the generated environmental map was about 1[m].

Keywords: Field Robot, ROS, SLAM, Mapping, Forestry

1. Introduction

In Japan's forestry industry, the decline in timber prices and labor shortages are causing the decline of forestry. However, much labor and time have to be spent on forest management and conservation. Especially, In-forest investigations and weeding work are burdensome for forestry workers. The in-forest investigations measures and manages the amount of resources, the amount of growth and quality of trees. The weeding work removes weed that impede the growth of seedlings after afforestation. In order to reduce the work burden, mechanization and automation of forestry work are required. In this research, we developed an autonomous moving robot with platform of all-terrain vehicle with high moving ability in rough terrain. With this robot, we aim to automate the weeding work and resource management within the forest. We introduced the Simultaneous Localization and Mapping (SLAM) system that performs self-position estimation and environmental map generation based on Robot Operating System (ROS) to the robot.[1]

2. The Robot Platform

Fig.1 shows the robot developed in this study. We will operate the robot on a steep slope in the actual mountain hillside, and the road is rough. Thus, the robot is based on an ATV (Kawasaki, Inc.) to handle on the rough terrain. The size of robot is following; length: 1400 [mm], width: 900 [mm], height: 1200 [mm]. In the Japanese forestry work, the planting interval is approximately 2.5 [m], which is narrow compared to forestry regulations in other countries. The size of our developed robot is sufficiently small compared to the planting interval that it can easily travel into the native forest. The robot has four wheels, the front two wheels are turning and the rear two wheels are driving. The external environment is recognized by RealSense R200 (Intel, Inc.) which is constructed with a depth sensor and a RGB-CCD camera, and equipped on the ATV. The depth sensor can acquire an object's position as three-dimensional (3D) point cloud data via an infrared laser measurement. The maximum measurable depth direction of the sensor is 15 [m]. In addition, the robot has an inertial measurement unit (IMU) for pose detection and RTK-GPS for global localization. The power source is gasoline engine. The brake lever, accelerator lever, and steering are all controlled by DC motors. A weeding task is performed by a weeding mechanism that has multi blades and attached on the front of the ATV.

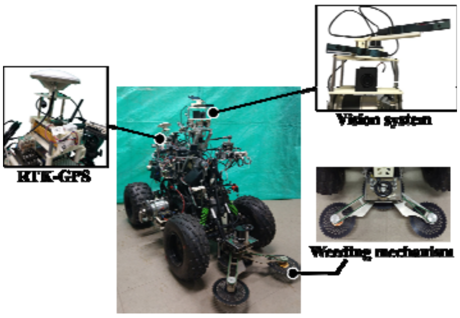


Fig. 1. Outline of robot appearance

3. Map Generation System

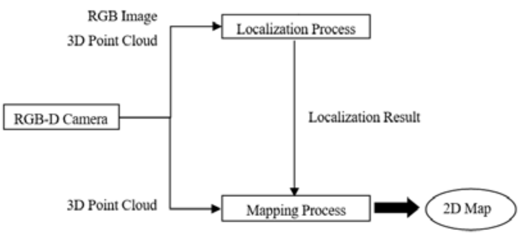


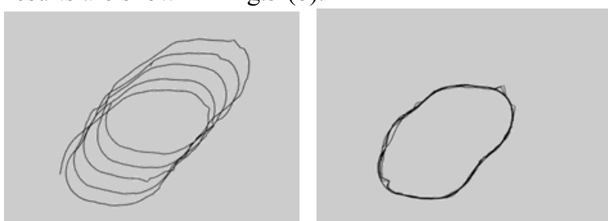
Fig. 2. Overview of the system

Fig.2 shows the overview of the environmental map generation system introduced on this robot. This system configures SLAM using ROS. This system is divided into the two processes, localization process and mapping process.

3.1. *Localization process*

In localization process, RGB image and 3D point cloud data obtained from RealSense R200 are used as input. We used ORB-SLAM2 for localization process.[2] In ORB-SLAM2, the ORB feature point is obtained from the RGB image, and the camera pose is estimated from the matching result of the ORB feature points at discrete times t-1 and t. ORB feature points are robust against violent motion and interference of light. The camera pose is converted to metric units by using a 3D point cloud data.

As shown in Fig.3 (a), even when the robot moves along the same path, an error occurs in the estimated self-position. This error causes a problem that environmental map generation does not work well. Therefore, we calculate the similarity of ORB feature points between current and previous RGB images. Then, if there is an RGB image whose similarity is higher than the threshold value, the self-position estimation result between that RGB image and the current RGB image is corrected. The results are shown in Fig.3 (b).



(a) Before correction (b) After correction

Fig. 3. Localization result

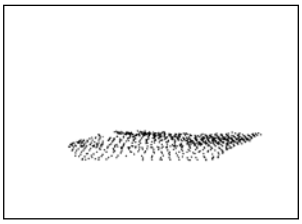
3.2. *Mapping process*

In mapping process, 3D environmental map and 2D environmental map are generated using localization process results and 3D point cloud.

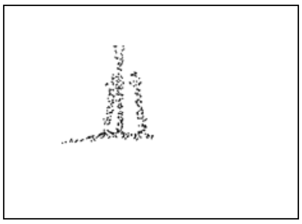
The 3D environmental map is generated by arranging 3D point cloud in space.[3] If the movement path of the robot is corrected by localization process, the position of the 3D point cloud is corrected.

When generating a 2D environmental map, identification of the ground and obstacles of the 3D point cloud is performed. The identification is divided into the two steps. First, Estimate the ground using the least squares method for any 20 point clouds and calculate the normal vector. Second, the vertical component of the normal vector is compared with the threshold, and if it is larger than the threshold, the point cloud is identified as the ground. The point clouds other than the ground is obstacle. This identification is performed on all point clouds, and the point cloud are projected on a 2D plane to generate a 2D environmental map. All point clouds, ground and obstacle are shown in Fig.4.

(a) All point clouds



(b) Ground point clouds



(c) Obstacles point clouds

Fig. 4. Point clouds

4. Experiment

We evaluated the constructed system using the robot movement results at the experimental site in our university. The experimental site consists of 39 trees, and there are no objects other than trees. The coordinates of the center of a standing tree are measured in advance using RTK-GPS, and this coordinate is used as a true value. The generated 3D environmental map is shown in Fig.5, and the 2D environmental map is shown in Fig.6. The true value of the tree coordinate, the estimation value of the tree coordinate, and the error is shown in Table 1. A unique number is set for tree. The estimated value is the central coordinates of the clustering result by Euclidean distance. The minimum value of the x-coordinate error is 0.04[m] at No.30, and the maximum value is -1.8[m] at No.35. The minimum value of the y-coordinate error is 0.15[m] at No.8 and the maximum value is 1.95[m] at No.44. The error of more than ± 1.5[m] occurred at No.35, 39 and 44.

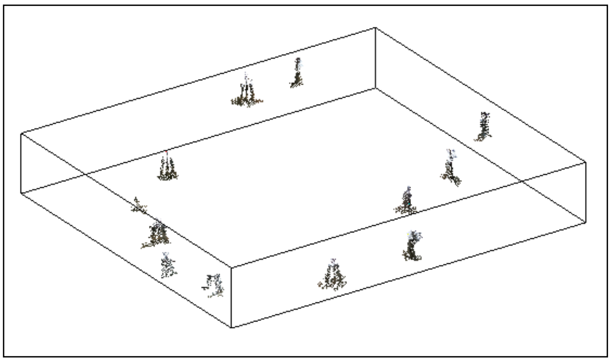


Fig. 5. 3D environmental map

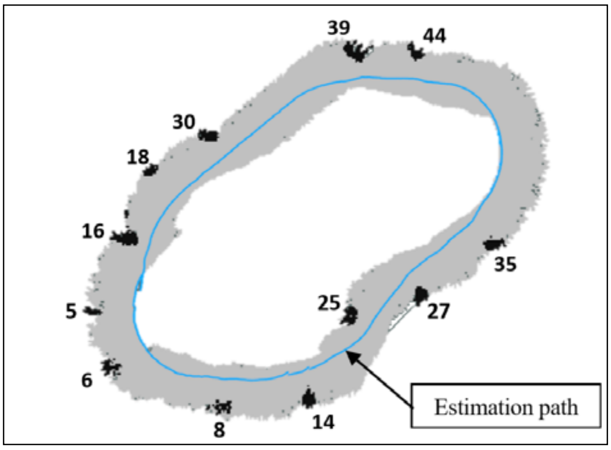


Fig. 6. 2D environmental map

Table 1. True/estimation value and error

	x-coordinate			y-coordinate		
No.	True value[m]	estimation value[m]	error[m]	True value[m]	estimation value[m]	error[m]
5	0.91	0.52	0.39	2.11	3.05	-0.94
16	0.66	0.30	0.36	6.52	5.48	1.04
18	0.34	0.27	0.07	10.12	9.35	0.77
30	2.44	2.40	0.04	12.22	12.60	-0.38
39	7.35	8.14	-0.79	18.32	20.02	**-1.70**
44	10.15	11.06	-0.91	19.52	21.47	**-1.95**
35	17.85	19.72	**-1.87**	13.02	13.97	-0.95
27	16.25	17.26	-1.01	8.82	9.83	-1.01
25	13.35	14.23	-0.88	6.12	7.17	-1.05
14	12.45	13.47	-1.02	2.00	2.17	-0.17
8	8.55	8.90	-0.35	0.13	0.28	-0.15

5. Consideration

The error of the coordinates of No.35, 39 and 44 is larger than that of other trees. The reason is that the error of self-position estimation is large because the movement distance from No.30 to No.39 and No.44 to No.35 are longer than the movement distance between other trees. In the RGB image when moving between No.44 and No.

35, although ORB feature points can be obtained from the ground, but there are few trees within the range recognized by RealSense R200, so almost ORB feature points are not obtained from trees. The ORB feature points that can be obtained from the ground are unstable because it is difficult to obtain ORB feature points from the same place at discrete times t-1 and t. Therefore, it is considered that the matching of ORB feature points can't be performed sufficiently, and the accuracy of self-position estimation is degraded. When matching ORB feature points, it is considered necessary to track ORB feature points obtained from unstable ground.

6. Conclusion

In this study, we constructed SLAM by ROS and generated the environmental map in the forest using robot. It was possible to identify trees and ground even in a site simulating a forest with complex terrain. In addition, it was found that self-position estimation is possible when there is an obstacle. However, it was found that an error occurs when there is no obstacle.

In the future, when matching ORB feature points, not only RGB images at discrete time t-1 and t are used, but also matching with RGB images before t-1 is used. By maintaining unstable ORB feature points obtained from the ground, we aim to develop a system that reduces the error of self-location estimation. And we aim at the further accuracy improvement of self-position estimation by mutually compensating the self-position obtained from this system and RTK-GPS.

References

1. M.Quigley, B.Gerkey, K.Conley, J.Faust, T.Foote, J.Leibs, E.Berger, R.Wheeler, A.Y.Ng. "ROS: an open-source Robot Operating System," In Proceedings of the IEEE International Conference on Robotics and Automation, Vol.3, p.5 12–17 May 2009.
2. R. M. Artal and J. D. Tardós. "ORB-SLAM2: an Open-Source SLAM System for Monocular, Stereo and RGB-D Cameras", IEEE Transactions onRobotics, Vol.33, No.5, pp.1255-1262, 2017.
3. M. Labbé and F. Michaud. "RTAB-Map as an Open-Source Lidar and Visual SLAM Library for Large-Scale and Long-Term Online Operation," Journal of Field Robotics, 2018.

Graph-Based Path Generation for Area Coverage

Ayumu Tominaga
Graduate School of Computer Science and Systems Engineering, Kyushu Institute of Technology, 680-4 Kawazu Iizuka, Fukuoka, Japan

Eiji Hayashi
Graduate School of Computer Science and Systems Engineering, Kyushu Institute of Technology, 680-4 Kawazu Iizuka, Fukuoka, Japan

Ryusuke Fujisawa
Graduate School of Computer Science and Systems Engineering, Kyushu Institute of Technology, 680-4 Kawazu Iizuka, Fukuoka, Japan

E-mail: tominaga@mmcs.mse.kyutech.ac.jp, haya@mse.kyutech.ac.jp, fujisawa@ces.kyutech.ac.jp

Abstract

This research addresses the trajectory generation in 2D Euclidean space for navigation of an autonomous land vehicle (ALV). The main aim of the approach presented is to give the trajectory in order to cover the given work space to the ALV for maximize effect of the ALV. In this work, we propose graph-based offline trajectory generation method. Here, a Hamiltonian path would be find using 2D coordinates of environmental landmarks as nodes in the graph. The Hamiltonian path contains nodes that will express a midpoint between a pair of landmarks, it could be treated as global trajectory. We applied the method to an actual artificial forest with treating cultivated trees as nodes in the graph.

Keywords: Path Generation, Area Coverage, Field Robot, Navigation, Graph, Forestry,

1. Introduction

An autonomous mobile robots are imposed with various tasks for social implementation. In previous researches, many robots and robotic applications had been developed and proposed with required applications and environments given to these. This research is focusing on developing an Autonomous Grand Vehicle (AGV). An important factor in AGV will be path planning. As an example of previous research of path planning, the problem of searching for the shortest path which connecting any two points has been treated [1-3]. On the other hand, some robot applications require to generate the path which can cover a whole work space given to the robot, for example, cleaning robot, painting robot and security robot and so on [4]. This research addresses the path generation problem aiming at work area coverage as mention above. Previous researches has proposed a method that divides the work space by several grids and find the path connecting every grids. In addition, as Travelling Salesman Problem, some studies had been proposed to find the minimum cost route which contains every way points [4]. In any case, free-spaces (a region that the robot can pass through without collision with obstacles) in the given work space would be connected with covering entire of work area based on graph structure and graph algorithms.

In this research, we propose a new cyclic path generation based on a line graph in graph theory, which has not attention until now. The proposed method does not require accurate size information of obstacles. Only a representative point (for example, center coordinate of the obstacle) for indicating the existence of the obstacle is input, and the cycle is obtained by only some geometric computation and graph algorithms. This method was applied to actual artificial forest to implement an Autonomous Grand Vehicle (AGV) to forest industry.

2. Graph and line graph

In this paper, a graph is denote $G = (V, E)$ constructed by a set of nodes V and a set of edges E. An individual node $v_i \in V$ represents a position on a 2 dimensional work space, and an edge $e_j \in E$ represents a line which join two nodes. Here, $e_j = \{v_s, v_t\}$ represents e_j is joining v_s and v_t. Our proposed path generation method utilize "Line graph" in graph theory. A line graph in graph theory is one of the graph transformation in graph theory, defined as below. The line graph of a graph G will denote $L(G)$ in this paper.

- $L(G)$ has a set of nodes whose nodes corresponding the edge in G.
- Two nodes $v_s, v_t \in V(L(G))$ are adjacent if the corresponding edges in G are incident.

As shown the example of line graph (Figure 1), $L(G)$ is the graph with edges in G replaced to by nodes and represents the condition of edge joining of the original graph G [5]. It is known that the line graph has the following characteristic; If the original graph G is Eulerian graph then its line graph $L(G)$ will be Hamiltonian. For Eulerian graph, the graph has a cycle which can traverse every edges with using each edge exactly once in the graph. For Hamiltonian graph, the graph has a cycle which can visit every nodes with using

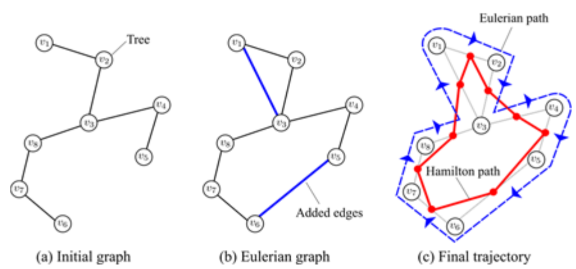


Fig.2. Procedure of graph-based path generation

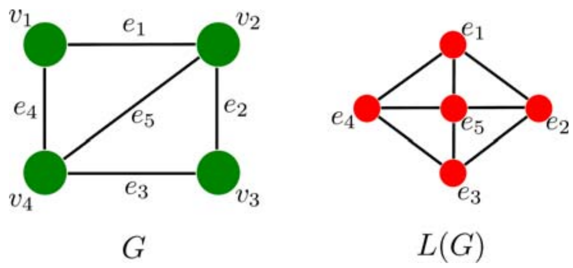


Fig.1. Example of the line graph

each nodes exactly once. Well-known Travel Salesman Problem, to find Hamiltonian cycle in the graph will acquire the shortest distance path visited every positions [6]. In this proposed method, we can find the Hamiltonian cycle in a work space given to an autonomous moving robot based on the line graph characteristic as mention above, and short time.

2.1. *Procedure of path generation*

Figure 2 shows the procedure of path generation. In this method, the information of work area would be obtained in advance, for example, the size of work area and the location of landmarks in the environment. In this paper, at first, to refer the location of landmarks, the initial graph G is generated. Then, edit G to be Eulerian graph by removing and adding edges appropriately. Finally, find the Eulerian path, and the Hamiltonian cycle could be computed from it. Its Hamiltonian cycle will represent the trajectory that can traverse entire of work area. Details of each step are mention below.

- Generation initial graph

Initial graph G has the set of nodes V, here each node was given two-dimensional coordinates of landmarks in the environment. Each landmark would be obstacles for the robot, and also, might be an object that should not cause damage. Based on the neighborhood situation of the nodes, the set of edges is constructed appropriately. Every edges will represent the segment that the robot should cross these. Therefore, all edges should be connected between pairs of nodes that are on the line of sight condition. As most easiest way, we used Delaunay Triangulation method, and its every Delaunay edges add to G.

- Editing to be Eulerian graph

To ensure $L(G)$ contains the Hamiltonian cycle, G should be Eulerian graph, therefore, every nodes in G should be indented even number edges. The Euler graph

G^* will be obtained by removing or appending the edges subject to the following conditions.

1. The pair of nodes $\{v_s, v_t\}$ is joined by e_j. If both nodes are incident odd number edges, than e_j should be removed from $E(G)$.
2. The pair of nodes $\{v_s, v_t\}$ is not joined by single edge. If the path between these nodes is found then the edges construct the path should append to G.

- Computation Hamiltonian cycle
  The Eulerian path in G^* will be found by Fleury's way [5], and denoted $B = \{e_1, e_2, \dots, e_m\}$. Finally, new nodes $v_m^l \in V(L(G))$ defined by $v_m^l = h(e_m)$, and the order of visiting of $v_m^l \in V(L(G))$ should be obey the B, we could get the Hamiltonian path in the environment. Here$h(e_m)$, we imposed the natural geometry condition as shown in Eq.(1). Here, $\{x_s, y_s\}, \{x_t, y_t\}$ is the positions given to the nodes $v_s, v_t \in V(G)$. In the other words, the nodes of $L(G)$ are represent the midpoints of every edges in G^*.

$$h(e_m) = \left\{\frac{x_s + x_t}{2}, \frac{y_s + y_t}{2}\right\} \qquad (1)$$

3. Experiment of the Hamiltonian cycle generation

The proposed method was applied to an actual work space. In our research group, we have been developed the autonomous ground vehicle (AGV) for labor-saving of forest industry. So, the AGVs work space is an actual artificial forest, the environmental landmarks are the trees constructing it. One of the task of the forestry AGV is eliminating of weed plants. For such tasks, to maximized the effective of the AGV, the navigation for given work area coverage might be important and challenging process.

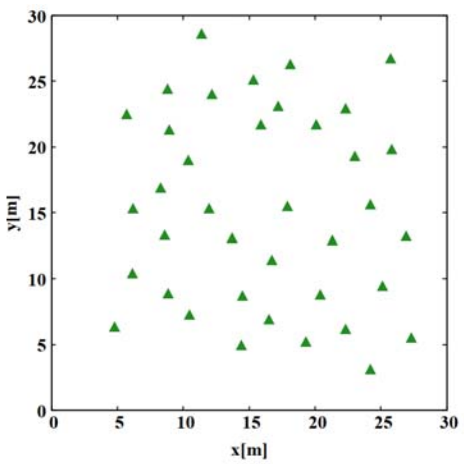
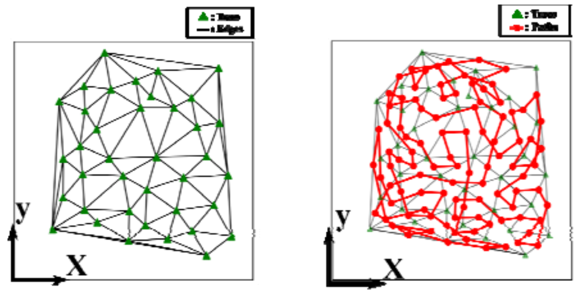


Fig.3. Tree locations of artificial forest using experiment



(a) Initial graph **(b) Hamiltonian cycle**

Fig. 4. Result of Hamiltonian cycle generation

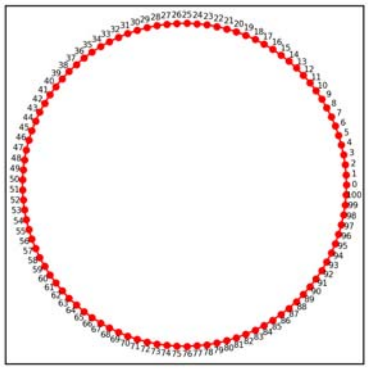


Fig.5. Circular layout of generated Hamiltonian cycle

Figure 3 shows the location of the trees of actual artificial forest using the experiment. The forest is constructed by 38 trees, each tree position were measured by RTK-GPS. Figure 4 shows the initial graph and Hamiltonian cycle that computed by the proposed method. Figure 5 shows the Hamiltonian cycle displayed in a difference layout. As shown figure 4, every nodes were generated at midpoints of the tree pairs and connected by the edges. In addition, as shown in figure 5, every generated nodes were visited exactly once. This result suggests that line graph based method could compute Hamiltonian cycle in given work space with using only the center points of environmental landmarks.

4. Conclusion

This research proposed the graph-based path generation method for the AGV. A final trajectory in the given work space con be computed systematically from an initial geographic map contains the positions of landmarks. And also, the proposed method would provide the Hamiltonian cycle which can traverse entire of work space by simple geometrical computation and graph algorithms. However, to improvement the safety for the AGV, the collision of the segments in Hamiltonian cycle and the obstacles in the environment should be considered in offline and online. In addition, in order to be applicable to the placement of various landmarks, verification in urban environment, underwater environment and so on are necessary in the future.

References

1. S. M. Lavalle and J. J. Kuffner, “Randomized kinodynamic planning”, In Proc. IEEE Int’l Conf. on Robotics and Automation, 1999.
2. E. W. Dijkstra, “A note on two problems in connexion with graphs,” Numerische mathematic, vol. 1, pp. 269-271, 1959.
3. A. V. Goldberg and C. Harrelson, “Computing the shortest path: A search meets graph theory,” in Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms, pp. 156-165, Society for Industrial and Applied Mathematics, 2005.
4. E. Galceran and M. Carreras, “A suvey on coverage path planning for robotics,” Robotics and Autonomous System, vol. 61, no. 12, pp.1258-1276, 2013.
5. F. Fujie and P. Zhang, “Line Graphs and Power of Graphs,” in Covering Walks in Graphs, Springer, 2014.
6. F. Harary, “Graph Theory,” Reading MA, Addison-Wesley, 1969.

A Hardware-Oriented Echo State Network for FPGA Implementation

Kentaro Honda
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka, 808-0196, Japan
Hakaru Tamukoh
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka, 808-0196, Japan
E-mail:honda.kentaro461@mail.kyutech.jp,tamukoh@brain.kyutech.ac.jp
http://www.lsse.kyutech.ac.jp/english/

Abstract

This paper proposes implementation of an echo state network (ESN) to field programmable gate array FPGA). The proposed method is able to reduce hardware resources by using fixed-point operation, quantization of weights, which includes accumulate operations and efficient dataflow modules. The performance of the designed circuit is verified via experiments including prediction of sine and cosine waves. Experimental result shows that the proposed circuit supports to 200[MHz] of operation frequency and facilitates faster computing of the ESN algorithm compared with a central processing unit.
Keywords: Reservoir Computing, Echo State Network, Field Programmable Gate Array

1. Introduction

Neural networks are highly expected to be applied into embedded systems such as robots and automobiles. However, Deep neural networks (DNNs) [1] require high computational power because a lot of accumulate operations are being processed using them. Generally, graphics processing units are used to accelerate these computations, however as their power consumption is high, implementing embedded systems using them is difficult due to a power limit. To mitigate this problem, we have implemented DNNs into hardware such as field programmable gate arrays (FPGAs), realizing high-speed calculation with low power consumption.

In this paper, we have implemented an echo state network (ESN) [2], a kind of reservoir computing (RC) into an FPGA. An RC is a recurrent neural network (RNN) model in which only the weights of an output layer are defined in the training step. ESNs are able to learn time-series data faster than general RNNs such as long short term memory (LSTM). In ESNs, a lot of accumulate operations of input data and weights are executed, however, there are limitations of FPGA resources such as loot up table (LUT), flip flop (FF) and digital signal processor (DSP). As a result, we have modified the algorithms and architectures of ESNs. Furthermore, we implement the proposed hardware-oriented algorithms into the FPGA and show the effectiveness of the proposed methods by comparing the proposed circuit with other.

2. Echo State Network

The ESN is a type of RC which consists of three layers: an input layer, a reservoir layer and an output layer, shown in Fig1, in which the neurons of the reservoir layer are randomly connected to each other.

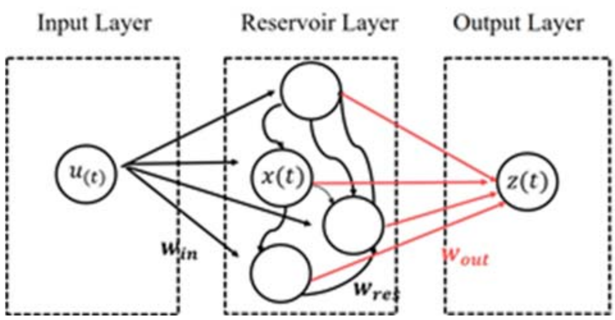


Fig 1: Echo Sate Network

ESN is described by the Equations 1 and 2,

$$x(t) = f\left((1-\delta)x(t-1) + \delta\left(w_{in}u(t) + w_{res}x(t-1)\right)\right) \quad (1)$$

$$z(t) = w_{out} \cdot x(t) \quad (2)$$

where $x(t)$ and $z(t)$ are output of the reservoir and output layer, respectively, time t, $u(t)$ is an input data, δ is the leak rate, which is the rate of the term $x(t-1)$ that affects $x(t)$. w_{in}, w_{res} ,and w_{out} are the weights of the input, reservoir and output layer ,respectively. The activation function f is defined as the hyperbolic tangent function. The reservoir layer follows the echo state property (ESP) [3] and its weights are initialized by the following steps:

1. All weights of the reservoir layer are generated from a normal distribution.
2. A spectral radius (the maximum eigenvalue of the weights) is calculated and all the generated weights are divided by it.
3. All weights are multiplied by a constant value.

In standard RNNs, all weights are updated following the backpropagation through time (BPTT) algorithm [4]. On the other hand, in the ESN, only the weights of the output layer are updated in one-shot learning through ridge regression as follows;

$$w_{out} = (X^T X + \lambda I)^{-1} X^T Y \quad (3)$$

where X is the matrix of $x(t)$ for all time-series, Y is the matrix of the supervised signal for all time-series, and λ is the regularization term.

3. Hardware-Oriented ESN

There are certain limitations of FPGA's resources, therefore, we have to modify the algorithms to suitable for FPGA implementation. In this paper, we design a circuit for the ESN by following three methods.

3.1. *Fixed Point*

Generally, computations are conducted with floating point numbers, which are an exponential representation, and can represent a wide range of numbers. A circuit using floating point numbers is more complex as it requires many FPGA resources. In contrast, although the fixed-point representation can only represent a narrow range of numbers, the circuit resources is less complex compared with that using the floating point.

3.2. *Quantization*

One way to reduce the complexity of a circuit is using quantized values that are able to simplify the computation while maintaining its accuracy [5].

Therefore, we calculated the outputs of the reservoir layer (Equation 1) using quantized weights. Generally, the weights of input and reservoir layers are real numbers resulting in several DSPs to compute real number multiplications. Therefore, we transformed the real valued weights to ternary values: 0 or ± 1. Furthermore, the accuracy by using this quantization for both training and prediction mode are maintained.

The circuit of the neuron is shown in Fig2. Where n is the number of reservoir's neurons. u_n and w_n are inputs and weights of input and reservoir layers, respectively, and m is the bit width of the input data. Furthermore, the circuit is able to calculate accumulate operations using only AND and XOR operations.

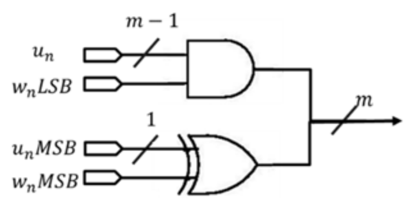


Fig 2 : circuit of neuron

3.3. *Sequence product-sum of output layer*

As shown in Fig 3, general product-sum operations can be represented by a tree structure. Using this representation, the number of adders and multipliers increases with the number of neurons. Therefore, in this research, we sequentially calculated the product-sum of the output layer. Fig 4 illustrates the product-sum operation by the proposed method, where A_i is an intermediate variable that temporarily stores the accumulate value. As this method consists of only a single adder, multiplier, and register per neuron in the output layer, the complexity of the circuit is reduced.

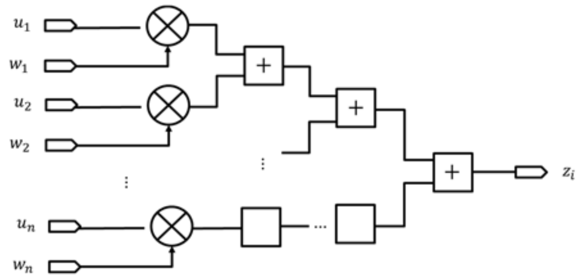


Fig 3: Tree structure of the product-sum circuit

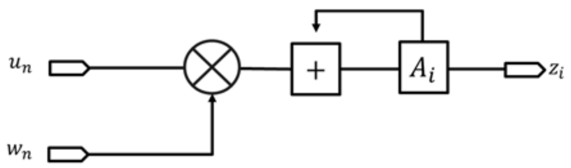


Fig 4: Sequence structure of the product-sum circuit

4. Experiment

In the experiment, we have created two types of circuits in order to verify the effectiveness of the proposed circuit and compared its calculation speed with those of the other devices. The task to evaluate their performance is the prediction problem of sine and cos waves. The number of neurons in the input, reservoir and the output layers were 2-,100-, and 2, respectively, and the prediction was computed in an FPGA. The target device is a Zynq UltraScale+ MPSoC ZCU102, Furthermore, the experiment was conducted with an operating frequency of 200 MHz and a data width of 32bits operations [6]. Table 1 shows experimental conditions and Table 2 shows the comparison between the conventional and the proposed circuit.

Table 1: Experimental conditions

Tool	SDSoC 2018.3
Target device	Zynq UltraScale+ MPSoC ZCU102
Clock frequency	200[MHz]

Table 2: Details of each circuit

	Operation	Weights of input and reservoir layers	Module of output layer
Conventional	Floating point	Real values	Tree structure
Proposed	Fixed point	Ternary values	Sequence structure

5. Result

Fig 5 shows the prediction of the conventional and proposed circuits. The black, blue, and red lines represent the supervised signal, prediction of the conventional circuit and prediction of the proposed circuit, respectively. Both circuits were able to reproduce sine and cosine wave. Tables 3 and 4 shows the utilization of resources for the conventional and proposed circuit, respectively. The proposed method was able to reduce the overall use of resources approximately 50%. Table 5 shows the comparison between electric energy of conventional circuit and proposed circuit. The proposed method reduced the electric energy consumption by approximately 84% compared with the conventional one. Table 6 shows a comparison between the computation speed of the FPGA and other devices The proposed circuit was approximately 25 times and 340 times faster than a desktop CPU and embedded CPU, respectively.

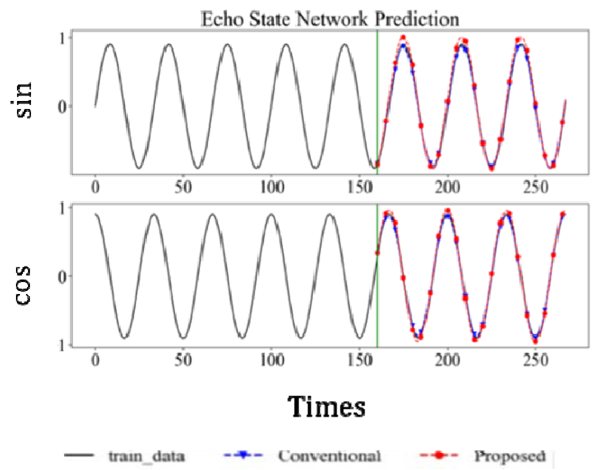


Fig 5: Output of circuits

Table 3: Utilization of resources for the conventional circuit

	Used	Total	Utilization
BRAM_18k	106	912	11.57
DSP_48E	519	2520	20.60
LUT	60557	274080	22.09
FF	96556	548160	17.61

Table 4: Utilization of resources for the proposed circuit

	Used	Total	Utilization
BRAM_18k	48	912	5.26
DSP_48E	20	2520	0.79
LUT	28933	274080	10.56
FF	44021	548160	8.03

Table 5: Electric energy consumption of each circuits

	Latency [ms]	Power [w]	Electric energy [w*ms]
Conventional	0.43	2.17	0.93
Proposed	0.20	0.69	0.14

Table 6: Computation speed of devices

Platform	Latency[ms]
CPU 3.2GHz (i7-8700)	5.215
Embedded CPU 1.2GHz (Quad Arm Cortex-A53)	68.123
FPGA 200MHz (XCZU9EG-2FFVBG1156E)	0.200

6. Conclusion

We were able to successfully adapt the circuit to enhance ESN computation in the FPGA. As a result, high-speed computation was possible while the circuit resources were reduced. To achieve this, fixed-point computation, quantification of weights, and sequential product-sum computation techniques were used.

In the future, it is expected to apply the proposed circuit and methods to embedded systems such as automobiles and robots.

Acknowledgements

This research is supported by JSPS KAKENHI grant number 17H01798.

References

1. G. E. Hinton, S. Osindero, and Y.W. Teh. “A fast learning algorithm for deep belief nets.” Neural Computation, vol. 18, no.7, pp. 1527–1554, 2006.
2. H. Jaeger, “The echo state approach to analyzing and training recurrent neural networks with an erratum note,” Bonn, Germany: German National Research Center for Information Technology GMD Technical Report, vol. 148, no. 34, pp. 13, 2001.
3. I. B. Yildiz, H. Jaeger, and S. J. Kiebel, “Re-visiting the echo state property,” Neural networks, vol. 35, pp. 1–9, 2012.
4. P. J. Werbos, “Backpropagation through time: What it is and how to do it,” Proc. IEEE, vol. 78, no. 10, pp. 1550–1560, 1990
5. Y. Aratani, Y. Y. Jye, A. Suzuki, D. Shuto, T. Morie, and H. Tamukoh, “Multi-Valued Quantization Neural Networks toward Hardware Implementation,” IEEE International Conference on Artificial Life and Robotics (ICAROB), pp.58, 2017.
6. Xilinx Zynq UltraScale+ MPSoC ZCU102 Evaluation Kit, https://www.xilinx.com/products/boards-and-kits/ek-u1-zcu102-g.html, 2019/12/1 access

Network with Sub-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
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
http://www.lsse.kyutech.ac.jp/english/

Abstract

We introduce network with sub-networks, a neural network which it's weight layers can be detached into sub-neural networks during inference phase. To develop trainable parameters which can be inserted into both base- and sub-models, firstly, the parameters of sub-models are duplicated to base-model. Each model is forward-propagated separately. All models are grouped into pairs. Gradients from selected pairs of networks are averaged and updated both networks. With MNIST dataset, our base-model achieves the identical test-accuracy to the regularly trained models. In other hand, the sub-models are suffered an extend of loss in test-accuracy, nevertheless the sub-models provide alternative approaches to be deployed with less parameters compare to the regular model.

Keywords: Model Compression, Neural Networks, Multilayer Perceptron, Supervised Learning.

1. Introduction

Deep neural networks (DNNs) have been gained the attraction in the most recent years from their ability to provide the state-of-the-art performance in varied applications. However, to deploy those DNNs into the mobile devices is proved to be problematic from the mobile devices are diverse in the specification. This raises the question: how to effectively design DNNs by given the specification of the mobile phone? To answer this question, two main factors within DNNs could be optimized.

The first factor is the performance of DNNs. In general, DNNs are provided an assumption by stacking the number of weight layers of DNNs, the better the performance of the model will be. One of the widely used example is the growing trend in the number of weight layers in ImageNet Large Scale Visual Recognition Competition (ILSVRC). AlexNet[1], the model which won ILSVRC-2012 consists of 8-weight layers. ResNet[2], the winner of ILSVRC-2015, contains of 152-weight layers. From AlexNet, ResNet reduces top-5 test error from 15.3 to 3.57. Even though, the growth in the number of weight layers might reduce the test-error rate of the model, it comes with the trade-off of the second factor, latency. More layers of DNNs means the higher number of parameters to compute. This also increases in the memory footprint which is crucial for the mobile device.

To solve this optimization problem, we might select the model which achieves the real-time performance given a mobile device specification. However, if the user diversely prefers the performance over the latency, this method does not satisfy the demand. Another method is to let the user select the preference and subsequently match the preference to the most suitable model. This method consumes the memory footprint for keeping various models into the mobile device. To satisfy user's preference in selectivity in both performance and latency without highly consuming memory footprint, we propose *network with sub-networks* (NSNs), DNNs which could be removed weight layers without dramatically decrease in the performance.

Generally, if one of the weight layers of DNNs is detached during the inference time, the performance of that model will be diminished. To explain our hypothesis, one of the widely used examples to explain how DNNs operate is to compare it as a feature extraction model. From the first weight layer, extracts the low-level features to the last layers extract the high level features. This process creates a dependent relationship between each weight layer.

To challenge this concept, we propose the training method that allows NSNs to dynamically adapt to the removing of weight layers. We call this method, *copying learn-able parameters* and *sharing gradient*. Both methods are designed to optimize the learn-able parameters for both models, the model with or without the weight layer to detached.

2. Related Works

2.1. *Slimmable Neural Networks*

Slimmable Neural Networks[3] (SNNs) is the main inspired of this research. If our purposed method adds or remove weights in depth-wise direction, SNNs append or detach weights in width-wise direction. The range of possible width of networks requires to be pre-defined as the switch. The main research problem is the mean and variance of activations which come out from different-width weight layers are generally diverse. SNNs proposed *switchable batch normalization* to correct the mean and variance of SNNs.

3. Network with Sub-Networks

There are two types of models in NSNs: the base and sub-model. We define the base-model as DNNs with *n* hidden layers. Where *n* is a positive integer more than zero. From base-model, we could create n number of sub-models. Each of sub-model is mapped with *0,…,n-1* hidden layers. From this concept, the biggest sub-model takes all of the weight layers of the base-model except the input layer. The second biggest sub-model takes all of the weight layers of the biggest sub-model except the input layer of the biggest sub-model. This could be done repeatedly until we get the sub-model that has not any hidden layer.

In the next section, we will describe two processes in our purposed method: *copying learn-able parameters* and *sharing gradient*. Those processes are designed to be applied repeatedly in every mini-batch training.

3.1. *Copying Learn-able Parameters*

The goal of *copying learn-able parameters* is to combine each sub-model into the base-model. To enforce the similarity between weight and bias parameters between each model, the weights and biases are copied from the lesser sub-model to bigger sub-model and repeat until the base-model. The process is shown in Eq. (1) and Fig. 1. Where $W_{o,m}$ is a weight variable, o is an integer indicating the order of weight layer and m is an integer indicating the model number.

$$W_{o+1,m+1} = W_{o,m} \quad (1)$$

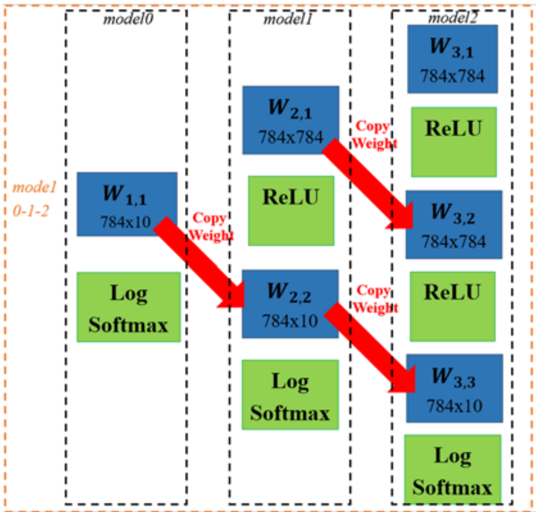


Fig. 1. Illustration of both *network with sub-networks* and *copying learn-able parameters* process. Where the base-model is two hidden layers DNNs and the sub-model as one hidden layer DNNs and a softmax-regression model. The name of the variable of weight, $W_{o,m}$ following with the size of weight array. *copying learn-able parameters* makes $W_{1,1}$, $W_{2,2}$ and $W_{3,3}$ to have exactly the same weight.

After we apply this process, if we remove the input weight layer of base-model with the non-linear activation function, it will become the sub-model.

3.2. *Sharing Gradient*

sharing gradient is designed to constraint the learnable-variables to able to perform in two or more networks. Firstly, we forward propagate all of the models. During back propagation, the gradients from each model are collected separately. Each model is paired from the sub-model without the hidden layer to sub-model with a hidden layer until, the sub-model with *n-1* hidden layers to base-model. The gradients from each model's pair are averaged and updated the weights and bias. *sharing gradient* process is expressed in Eq. (2) and Fig. 2 where *lr* is the learning rate and *L* is the loss function.

$$W_{m,o} = W_{m,o} - \frac{lr}{2}\left(\frac{\partial L_m}{\partial W_{m,o}} + \frac{\partial L_{m+1}}{\partial W_{m+1,o+1}}\right) \quad (2)$$

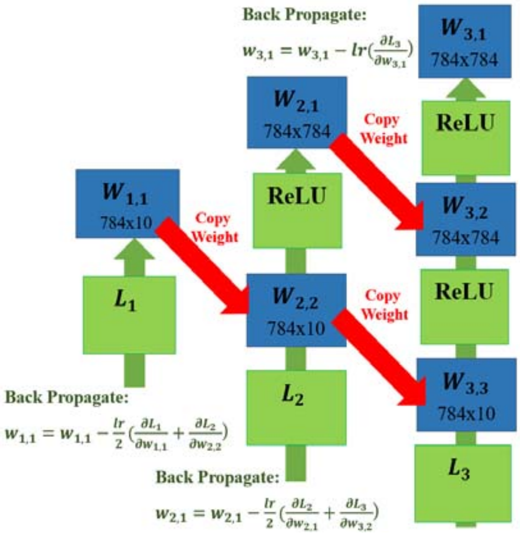


Fig. 2. *sharing gradient* in *model0-1-2* section. The gradients are shared from the sub-model to base-mode, pair by pair. Only the input weight layer of $W_{3,1}$ is regularly updated without sharing. Where L_m is the loss function at the *m* model.

The reason behinds *sharing gradient* for only a pair of models is from when sharing more than a pair gradient, the optimization becomes too complicated. In this case, the performance of NSNs hardly reaches the optimal point. Nonetheless, only an input layer of base-model, which has not a pair, is updated with the regular back propagation.

4. Experiments

The experiment was conducted using a hand-written digit image dataset, MNIST[4]. MNIST dataset consists of 60,000 training images and 10,000 test images. Each image in the dataset is the gray-scale image and, composed of 28x28 pixel. Each image pixel of MNIST mage was pre-processed into the range of [0, 1] by dividing all pixel value with 255.

Multi-layer perceptron (MLP) was applied with rectified linear unit (ReLU) as the non-linear activation function. The last layer was applied with log-softmax with the cost function as cross-entropy loss. The input layer of MLP was applied with Dropout[5], *p=0.8*. The hidden layers were put with dropout rate, *p=0.5*. In the case of softmax-regression, we did not apply Dropout into the model since it was already under-fitting. The base-models were further regulated by using L2-weight penalty.

We applied stochastic gradient descent (SGD) with momentum, $\alpha = 0.9$. Although, we applied with slightly different format of SGD with momentum. From Tensorflow[6], neural networks framework, Tensorflow format of SGD with momentum was shown in Eq. (3). Where *V* is the gradient accumulation term, *t* is the batch-wise iteration step and *G* is the gradient at *t+1*. Our format of SGD with momentum is shown in Eq. (4). After *V* was found, both of format was used the same Eq. (5). to update the weight, *W*.

$$V_{t+1} = \alpha V_t + G \tag{3}$$
$$V_{t+1} = \alpha V_t + (1-\alpha)G \tag{4}$$
$$W_{t+1} = W_{t+1} - lr(V_{t+1}) \tag{5}$$

NSNs performed better with our format of SGD with momentum comparing the regular or Tensorflow format at $\alpha = 0.9$. We speculated that NSNs required the higher proportion of the gradient accumulation, V comparing with the current gradient, G to converge. In the other hand, with the regularly trained DNNs, our format of SGD with momentum performed slightly worse in term of test accuracy. Hence, to perform a fair comparison between both type of models, the regularly-trained models were trained with Eq. (3). Our purposed method models were trained Eq. (4). We set the training batch as 128. Each model had been training for 600 epoch. We reported the best test accuracy which might occur during the training. The initial learning rate, $lr = 0.3$ and step down by one third every 200 epoch.

The experimental result consists of two sections. First section is *model0-1* or the base-model as MLP with a hidden layer, *model1*, with a sub-model as the soft-max regression, *model0*. Second section is *model0-1-2* or the base-model as two hidden layers MLP, *model2*. The sub-models are MLP with a hidden layer, *model1*, and the soft-max regression, *model0*. The graphical of *model0-1-2* is shown in Fig. 2. The base-line models which are regularly trained are referred as *ref-model* and following with number hidden layers. For example, *ref-model1* is the base-line MLP with a hidden layer. The results of base-line model are shown in Table. 1.

Table 1. Results of MNIST classification of base-line.

	Test Accuracy	Number Parameters	Regularization Parameter
ref-model2	0.9886	1.24M	1×10^{-5}
ref-model1	0.9882	0.62M	5×10^{-6}
ref-model0	0.9241	7.85k	9×10^{-5}

4.1. *Model0-1*

MLP with a hidden layer was used as the base-model. The sub-model was the softmax-regression. In all of the following experiment, we prioritized the base-model performance. We reported all of the test accuracy of models in epoch that contains the best test accuracy of the base-model. *model0-1* results are displayed in Table 2. We applied the regularization parameter as 9×10^{-6} at the base-model.

Table 2. Results of MNIST classification of *model0-1*.

	Test Accuracy	Number Parameters
model1	0.9857	0.62M
model0	0.9253	7.85k

Comparing with the *ref-model1* and *model1*, the test accuracy of *model1* were dropped for an extent. This indicated that our purpose methods negatively affected the performance of the model for ability to removing the weight layers.

4.2. *Model0-1-2*

MLP with two hidden layer was used as the base-model. The sub-models were MLP with a hidden layer and the softmax-regression. *model0-1-2* results were displayed in Table. 3. We applied the regularization parameter as 9×10^{-5} at the base-model.

Table 3. Results of MNIST classification of *model0-1-2*.

	Test Accuracy	Number Parameters
model2	0.989	1.24M
model1	0.9843	0.62M
model0	0.926	7.85k

The difference in test accuracy between *model1* and *model2* indicated the bias of our purposed method towards the base-model. We speculated this bias might come from the *sharing gradient* process. All of the gradients that sub-models received, were averaged from multi-models. However, our base-model had an input layer that updated from gradient from the model itself as shown in Fig. 2.

Comparing with the result of *model1* in model*0-1* and *ref-model1*, our *model2* in *model0-1-2* contrastingly out-performed with *ref-model2* for a tiny margin. We hypothesized that the constraints of our purposed method might cause some type of regularization into the models. In case of *model1* in *model0-1*, this regularization effect might excessively strong and negatively affected the performance. Nevertheless, in case of *model2* in *model0-1-2*, the regularization effect seems to be adequate and positively affected the accuracy.

5. Conclusion

We purpose NSNs, DNNs that could be removed weight layers on fly. NSNs consists of a base-model and, sub-models. To assemble sub-models into the base-model, *copying learn-able parameters* is introduced. *sharing gradient* is applied for learn-able parameters could be used in two or more models. Our purposed method was conducted in the small scale experiment with a few hidden layers DNNs with MNIST dataset. The bigger-scale models with the dataset will be focused on future works.

Acknowledgements
This research was supported by JSPS KAKENHI Grant Numbers 17K20010.

References

1. A. Krizhevsky, I. Sutskever, and G. E. Hinton, Imagenet classification with deep convolutional neural networks, in *Advances in neural information processing systems*, pp. 1097–1105, 2012.
2. K. He, X. Zhang, S. Ren, and J. Sun, Deep residual learning for image recognition, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.
3. J. Yu, L. Yang, N. Xu, J. Yang, and T. Huang, “Slimmable neural networks,” *arXiv preprint arXiv:1812.08928,* 2018.
4. Y. LeCun, C. Cortes, and C. Burges, “Mnist hand-written digit database,” *AT&T Labs [Online]. Available: http://yann.lecun.com/exdb/mnist,* vol. 2, p. 18, 2010.
5. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: a simple way to prevent neural networks from overfitting”, *The Journal of Machine Learning Research,* vol. 15, no. 1, pp. 1929–1958, 2014.
6. M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, et al., “Tensorflow: A system for large-scale machine learning,” in *12th {USENIX} Symposium on Operating Systems Design and Implementation* ({OSDI} 16), pp. 265–283, 2016.

A Fast Pick-and-Place Method for Home Service Robots using 3D point clouds

Tomohiro Ono
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka, 808-0196, Japan
Hakaru Tamukoh
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka, 808-0196, Japan
E-mail: ono.tomohiro342@mail.kyutech.jp, tamukoh@brain.kyutech.jp
http://www.lsse.kyutech.ac.jp/english/

Abstract

Home service robots have begun attracting attention due to decreasing birthrate and increasing aging population. The basic functions of home service robots are object recognition, picking and placing, recognition of people and environment, and interaction with peoples. In this paper, we focus on object picking and placing in the domestic environment. The pick-and-place task is a very important technique used for arrangement shelves and tidying up rooms. In order for the robot to operate smoothly, the movement to pick-and-place the object must be fast. Therefore, we develop a fast pick-and-place method using 3D point clouds. Regarding the picking, we describe a grasping-point estimation method. Regarding the placing, we describe a placeable position estimation method. These methods are used at RoboCup@Home, an international competition aimed at the practical application of home service robots, and their effectiveness and validity are verified.

Keywords: Home Service Robot, Pick-and-Place, 3D point cloud, RoboCup@Home

1. Introduction

Home service robots have begun attracting attention due to decreasing birthrate and increasing aging population [1-3]. The basic functions of home service robots are object recognition, picking and placing, recognition of people and environment, and interaction with peoples. In addition, using these technologies, home service robots perform a variety of tasks, such as shelf placement, room cleaning, and restaurant clerk. These tasks always use Pick-and-Place.

In this paper, we focus on object picking and placing in the domestic environment. In order for the robot to operate smoothly, the movement to pick-and-place the object must be fast. Therefore, we develop a fast pick-and-place method using 3D point clouds. Regarding the picking, we describe a grasping-point estimation method. Regarding the placing, we describe a placeable position estimation method. These methods are implemented on TOYOTA HSR[1] and used at RoboCup@Home[4], an international competition aimed at the practical application of home service robots, and their effectiveness and validity are verified.

2. Pick-and-Place method

2.1. *Picking method*

Home service robots need to grasp objects of various shapes in the domestic environment. Sometimes the robot grasps the object it sees for the first time. Therefore, we propose a method that can grasp unknown objects at high speed based on three-dimensional (3D) point clouds.

Figure 1 shows the overview of our picking method. The procedural step are as follows:

(a). Generate 3D point clouds.
(b). Remove planes using segmentation.
(c). Extract information on each object using clustering.
(d). Calculate the size and direction of each object and estimate the grasping direction.

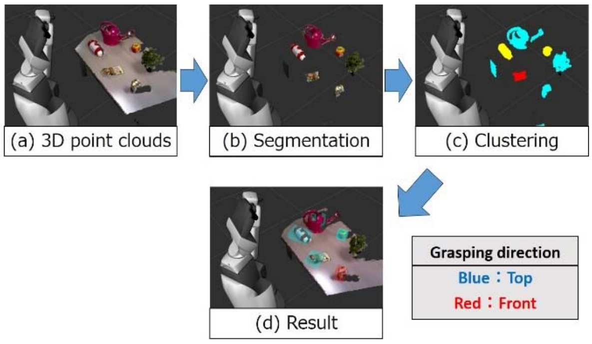


Figure 1. Overview of the proposed picking method.

This method is very fast because it works with simple calculations, and it also can be applied to unknown objects, enabling home service robots to have a stable grasp.

2.2. *Placing method*

After the picking, the home service robot needs to move the object to a predetermined position and place it. When the robot places an object, it is necessary to estimate where it can be placed on a desk, a shelf, and so on.

Figure 2 shows the overview of our placing method. The procedural step are as follows:

(a). Generate 3D point clouds.
(b). Extract and divide planes.
(c). Convert from 3D point clouds to 2D images.
(d). Perform morphology processing.
(e). Calculate plane areas.
(f). Estimate a placeable position.

This method is also very fast because it works with simple calculations and can be applied to various places with flat surfaces, so that home service robots can be placed stably.

3. Experimental

These proposed methods were used at RoboCup@Home, an international competition aimed at the practical application of home service robots. In this competition, pick and place tasks are used in various situations. For example, robots organize shelves and clean rooms in a domestic environment.

Regarding the picking, it was possible to realize stable grasping regardless of whether the object was known or unknown. It also worked in real time and was able to perform smooth picking. It takes about 4.0 seconds for the robot to recognize and grasp the object.

Regarding the Placing, it was possible to achieve an arrangement that avoided other objects for furniture such as desks and shelves. It also worked in real time and was able to perform smooth Placing. The estimated time where the robot can be placed is about 0.205 seconds. Using these proposed methods, we got third prize at RoboCup@Home 2019 Sydney domestic standard platform league.

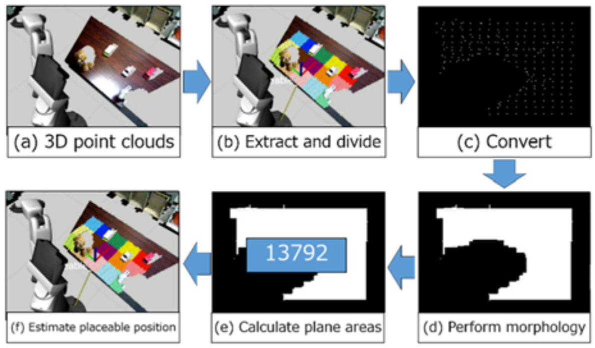


Figure 2. Overview of the proposed placing method.

4. Conclusion

In this paper, we proposed a fast method using 3D point clouds in Pick-and-Place for home service robot. In the picking, high-speed and stable gripping was achieved, and in the placing, high-speed and stable object placement was realized. The effectiveness and validity of the proposed method was verified in experiments in various situations at RoboCup@Home.

Acknowledgements

This work was supported by JSPS KAKENHI grant number 17H01798, and the New Energy and Industrial Technology Development Organization (NEDO).

References

[1] T. Yamamoto, K. Terada, A. Ochiai, F. Saito, Y. Asahara, and K. Murase, "Development of Human Support Robot as the research platform of a domestic mobile manipulator," ROBOMECH Journal, 2019.
[2] L. Iocchi, D. Holz, J. Ruiz-del-Solar, K. Sugiura, T. van der Zant, "Analysis and results of evolving competitions for domestic and service robots," Artificial Intelligence, pp. 258-281, 2015.
[3] H. Okada, T. Inamura, and K.Wada, "What competitions were conducted in the service categories of the World Robot Summit?," Advanced Robotics, 2019.
[4] RoboCup@Home. http://www.robocupathome.org/, (Accessed 2019-12-19).

Acceleration of training dataset generation by 3D scanning of objects

Yushi Abe, Yutaro Ishida, Tomohiro Ono, Hakaru Tamukoh
Life Science and Systems Engineering, Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Fukuoka-ken, Japan
E-mail: abe.yushi109@mail.kyutech.jp, ishida.yutaro954@mail.kyutech.jp, ono.tomohiro342@mail.kyutech.jp, tamukoh@brain.kyutech.jp
http://www.lsse.kyutech.ac.jp/english/

Abstract

A semi-automatic dataset generation system is effective to prepare a training dataset for object recognition in a personal residence. However, a semi-automatic that method requires significant manual processing to capture images of household objects. Therefore, we apply three-dimensional object scanning to eliminate manual processing and speedup dataset generation. Experimental results demonstrate that the proposed method can generate the dataset 40 minutes faster than a comparable previous method that did not require manual processing.

Keywords: dataset generation, CNN, Home Service Robot, Image recognition

1. Introduction

Recently, home service robots have attracted increasing attention due to the needs of an aging society with a declining birth rate. Home service robots can reduce the burden on people engaged in various types of repetitive and predictable work, such as waiters and housekeepers. The Kyushu Institute of Technology's home service robotics development team participated in the RoboCup@Home[1] League and the World Robot Challenge (WRC)[2] using the human support robot[3] developed by Toyota Motor Corporation. Object recognition technology is required for home service robots and recently using neural network-based deep learning[4] for image recognition has attracted attention. For example, You Only Look Once (YOLO)[5] provides real-time, high-precision object recognition.

To recognize objects in a personal living environment, a dataset that matches the environment must be trained in advance. However, deep learning generally requires a significant amount of data and manual generation of the dataset is time consuming.[6] In addition, manual annotation errors are an issue.

In this study, our objective is to realize a training dataset generation system that can easily recognize domestic objects with high accuracy and thereby reduce the time required to generate the dataset.

2. Previous research

2.1. *Semi-automatic Dataset Generation System*

Our robotics development team has been developing a semi-automatic dataset generation system[7] and using it in competitions.

Figure 1 shows an overview of object data capture using the previously developed semi-automatic dataset generation system. The system uses two RGB-D cameras to capture object images from different orientations. The object's background permeates using chroma key processing to make the background monochromatic. The dataset preparation flow is shown in Figure 2. After capturing the object

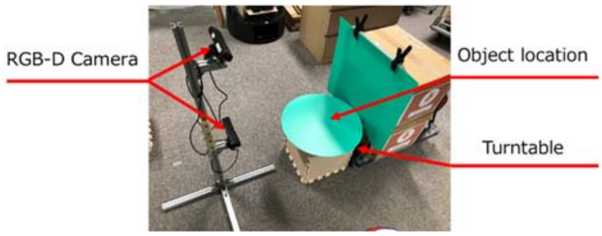


Fig. 1. Overview of object data capture system

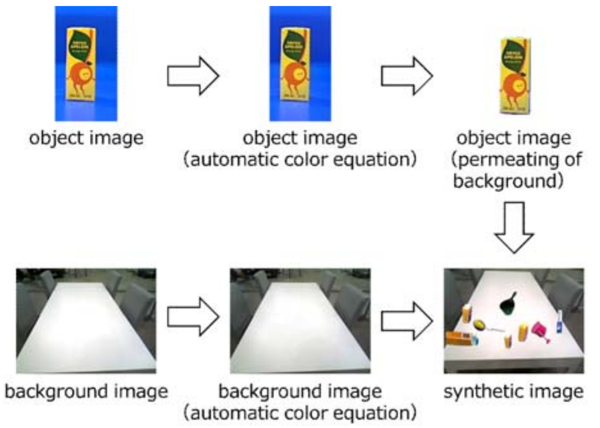


Fig. 2. Dataset generation process flow for the previous system

and background images separately, the automatic color equalization algorithm[8] is applied. Then, the background of the object images permeates. A synthesized image is generated automatically by randomly selecting an object image that has been permeated and synthesized it with the background image. Annotation information is automatically created simultaneously with the synthesized image.

2.2. *Problems*

The previous system uses chroma key and a fixed camera. Consequently, it has the following problems.

(i) The background may not permeate well depending on the direction of light or shadows.

(ii) When objects must be placed in multiple patterns, more time is required to capture the images.

The background color is filtered out by HSV conversion based on the pixel value. The threshold value for filtering is determined by empirical rules. Therefore, as described in problem (i), it may not be possible to completely filter out the background or

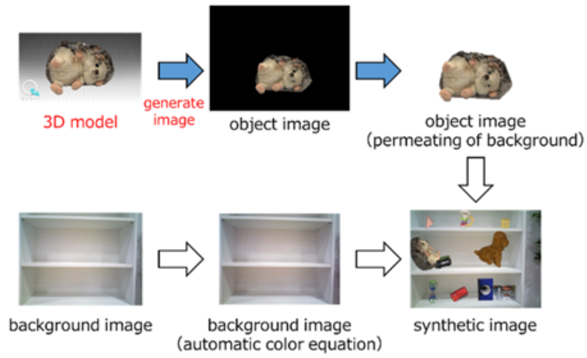


Fig. 3. Dataset generation process flow for the proposed system

the object image was broken. In such a case, the threshold must be adjusted, and the process needs to be repeated. However, threshold adjustment may also fail. In that case, the background must be removed one by one manually. Regarding the second problem (ii), since there are six possible placements on the paper pack, as shown in Figure 2, it is necessary to capture images of all six patterns. The time requires to capture all six patterns increases the time required to generate the dataset. In this paper, we propose a new method for the system used to capture object data.

3. Reducing Image Capture Time Using Three-Dimensional Scanning

In this paper, we propose a method to reduce the time required to capture images by using three-dimensional (3D) scanning. We used the *QLONE*[9] application, which can generate 3D models from scans captured by an iPhone camera. Scans are obtained by placing an object on a dedicated marker and rotating the object. Because the bottom of an object cannot be captured, two patterns of 3D models are created by inverting the top and bottom of each object. Thus, each object requires two or fewer shooting patterns. Multi-viewpoint images are automatically generated from the captured 3D models. At this time, the filtering process can easily be performed by making the background a different color from the color of the object. The dataset generation process flow of this system is shown in Figure 3. Note that only the image capture method differs from the existing system; thus the modified part is represented by the blue arrows in Figure 3.

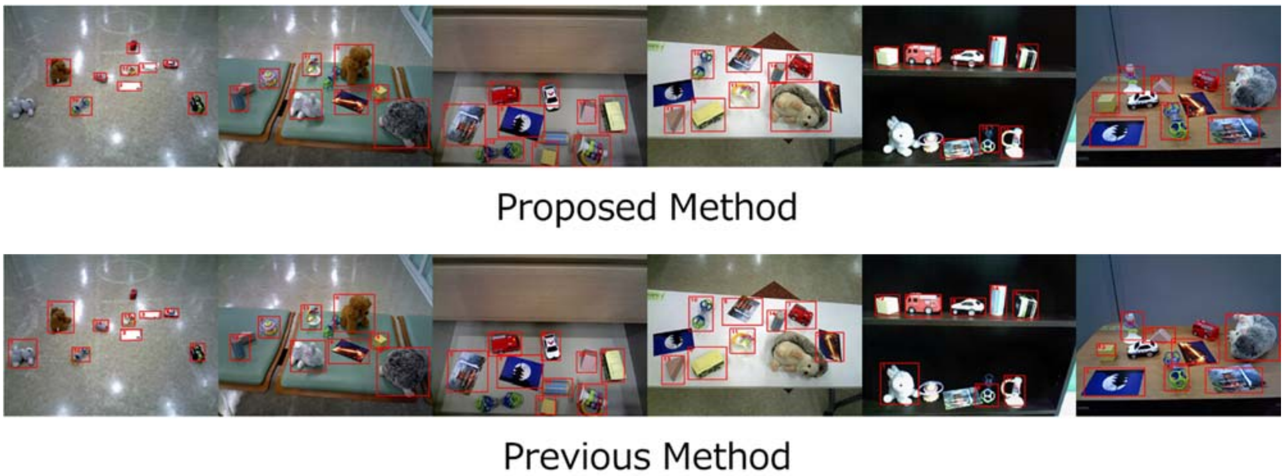


Fig. 5. Recognition result

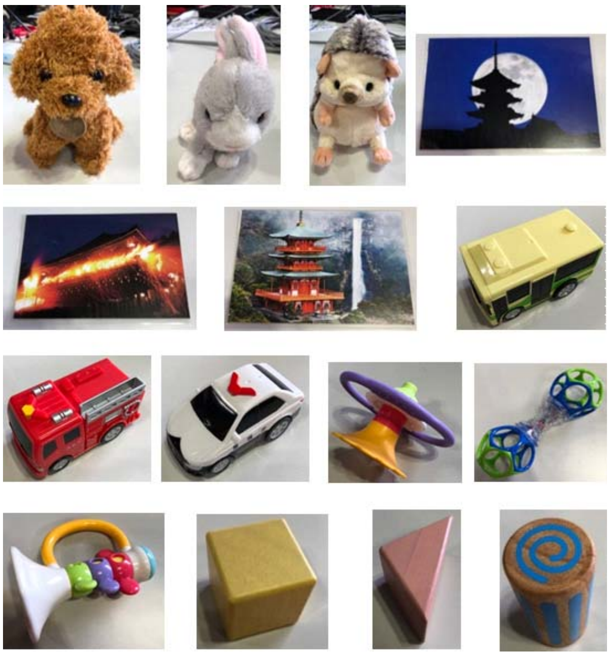


Fig. 4. Objects in WRC

4. Experiments

We compared the time required to generate the dataset using the proposed method and the existing method and the recognition accuracy of a trained YOLO model.

4.1. *Experiment environment*

The experimental environment was as follows.

- Total number of objects: 15 toys (used in WRC's Tidy Up task) shown in Figure 4
- Background image for synthesize: 18 types * 17 (shot in advance)
- Object recognition: YOLOv2
- PC: Intel Core i7-8700K, DDR4 32GB, NVIDIA GTX 1080

4.2. *Comparison of processing speed*

We compared the time duration from capturing the object images until all backgrounds have been filtered out. In the existing system, the total number of patterns to be placed was 39 for 15 objects. Shooting required approximately two minutes per pattern. Therefore, it took approximately 80 minutes to capture the image data, and it took approximately 70 minutes for all backgrounds to be completely removed. Thus, the total time required was approximately 150 minutes. On the other hand, the proposed system requires four minutes per pattern. One pattern is sufficient for objects that do not change even if the image is inverted vertically. Thus, it only takes approximately 100 minutes a total of 26 patterns. In addition, it takes approximately 10 minutes to generate images from 3D models and filter out the background; thus, the total time required is approximately 110 minutes. Note that the filtering process does not need to be done entirely manually. The above results demonstrate that the proposed system was

Table 1. Comparison of performance

Evaluation index	Proposed method	Previous method
Number of data	12600	15600
mAP (%)	60.72	64.77

able to generate a dataset at a faster speed than the existing system.

4.3. *Comparison of recognition accuracy*

The number of datasets generated was approximately 15,600 for the existing method and approximately 12,600 for the proposed method. We compared the recognition accuracy on YOLOv2 that was trained for up to 10,000 epochs using the dataset generated. The recognition results are shown in Figure 5. The results of comparing the proposed method with existing methods are shown in Table 1. We used mean Average Precision (mAP) as the comparison metric. The mAP value obtained by the proposed method was four points less than that of the previous method.

5. Conclusion

In this paper, we have proposed a dataset generation system using 3D scans. We generate a dataset of objects that were used in the WRC and to train YOLOv2. As a result, compared to the previous method, using 3D scans, dataset generation time can be reduced. In addition, the mAP obtained by the proposed method was only four points less than that of the previous method, which is not a significant drop in accuracy. The proposed method to generate images from 3D models can be used to generate datasets for home service robots.
We applied *QLONE* which is iPhone's 3D scanning application. *QLONE* requires a human operator; therefore, the process cannot be completely automated. A future challenge is to construct a system that enables the generation of 3D models and datasets without any human intervention.

Acknowledgements

This research is supported by the New Energy and Industrial Technology Development Organization (NEDO) and JSPS KAKENHI grant number 17H01798.

References

1. T. Wisspeintner, T. van der Zant, L. Locchi, and S. Schiffer, "RoboCup Home: Scientific Competition and Benchmarking for Domestic Service Robots," Interaction Studies, pp.392-426, 2009.
2. H. Okada, T. Inamura and K. Wada, "What competitions were conducted in the service categories of the World Robot Summit?," Advanced Robotics, 2019.
3. T. Yamamoto, K. Terada, A. Ochiai and F. Saito, "Development of Human Support Robot as the research platform of a domestic mobile manipulator," ROBOMECH Journal, 2019.
4. G. E. Hinton, S. Osindero and Y. W. Teh, "A fast learning algorithm for deep belief nets," Neural Computation, vol.18, no.7, pp.1527-1544, 2006.
5. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Real-Time Object Detection," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp.779-788, 2016.
6. J. Hestness, S. Narang, N. Ardalani, G. Diamos, H.Jun, H. Kianinejad, M. M. A. Patwary, Y. Yang, and Y. Zhou, "Deep Learning Scaling is Predictable, Empirically," arXiv:1712.00409, 2017.
7. Y. Ishida and H Tamukoh, "Semi-automatic Dataset Generation for Object Detection and Recognition and its Evaluation on Domestic Service Robots," Journal of Robotics and Mechatronics, vol.32, no.1, 2020.
8. A. Rizzi, C. Gatta and D. Marini, "A new algorithm for unsupervised global and local color correction," Colour Image Processing and Analysis. First European Conference on Colour in Graphics, Imaging, and Vision (CGIV 2002), pp.1663-1677, 2003.
9. "QLONE, the all in one tool for 3D Scanning," https://www.qlone.pro/

Anomaly Detection Using Autoencoder Trained with Reversed Color Models

Obada Al aama
Department of Life Science and Systems Engineering, Graduate School of Life Science and Systems Engineering,
Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan

Hakaru Tamukoh
Department of Human Intelligence System, Graduate School of Life Science and Systems Engineering,
Kyushu Institute of Technology,
2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan
E-mail:aama.obada-walidal425@mail.kyutech.jp, tamukoh@brain.kyutech.ac.jp
www.lsse.kyutech.ac.jp

Abstract

Autoencoders (AEs) have been applied in several applications such as anomaly detectors and object recognition systems. However, although the recent neural networks have relatively high accuracy but sometimes false detection may occur. This paper introduces AE as an anomaly detector. The proposed AE is trained using both normal and anomalous data based on convolutional neural network (CNN) with three different color models HSV, RGB and TUV (own model). As a result, the trained AE reconstruct the normal images without change, whereas the anomalous image would be reconstructed reversely. The training and testing of the autoencoder in case of RGB, HSV, and TUV color models were demonstrated and Cifar-10 dataset had been used for the evaluation process. It can be noticed that HSV color model has been more effective and achievable as an anomaly detector rather than other color models based on Z-test and F-test analyses.

Keywords: Convolutional neural network, Autoencoder, Anomaly detection, Color models.

1. Introduction

The Artificial intelligence (AI) is widely used and has been existed over many decades. It uses the information originating from sensors, images, languages and texts. Analyzing this information giving hypothesis leading to make decision[1]. AI can be viewed as a set that contains machine learning (ML) and deep learning (DL) [2]. DL is often categorized as supervised or unsupervised [3]. Autoencoders (AEs) are one of the DL methods which trained in an unsupervised fashion to automatically extract features of training data [4]. Moreover, anomaly detection is one of the most important applications of AEs [5]. One of the training methods that has been used for anomaly detection is a convolutional neural network (CNN). CNN has been applied in various modern applications, and it is often implemented in image analysis [6], speech and face recognition [7] and autoencoders [8] with great success.

The aim of this study is to use CNN-autoencoder trained with three different color models; Hue Saturation Value (HSV), Red Green Blue (RGB) and our own model (TUV) to improve the detection accuracy especially the anomalous one.

2. Research Concept

The main concept focuses over using the autoencoder trained with reversed color models in order to detect the anomaly data.

2.1. Autoencoder

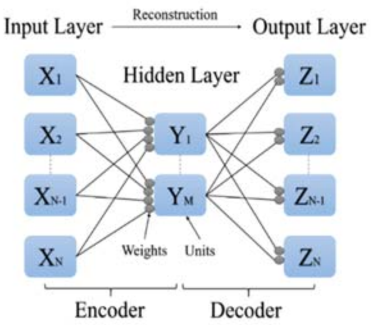


Fig. 1. Structure of autoencoder.

AEs are neural networks that aims to copy their inputs to their outputs. It is used to automatically extract features of training data. AEs are applied for object recognition systems that being used the anomaly detectors [5]. To improve a recognition accuracy, the anomaly detectors has the ability to remove anomalous objects before recognition process to reduce misrecognition. AEs are composed of three fully connected layers: input, hidden, and output layers. These layers are trained to reconstruct input data on the output layer as shown in figure 1.

2.2. Anomaly detection

The idea in anomaly detection based on machine learning, is to model the normal behavior of data in the training period, and then try to fit the test data using the trained model. In case a large inconsistency is found between the fitted model and the trained model, the test data is regarded as anomalous.

When using autoencoders, which applies dimensionality reduction to the input data, for anomaly detection, we assume that the data contains variables that can be represented in lower dimensions. These variables are also assumed to be correlated with each other and would show significant difference between normal and anomalous samples [9].

There are two types of training data for autoencoder to detect anomaly images; labeled and unlabeled data. Based on the type of data, the anomaly detection algorithm differs. In case of labeled data, conditional distributions can distinguish between correct and anomalous data. Accordingly, the probability of the conditional distribution determines whether the data is correct or anomalous. On the other hand, a generative model trained with correct data is employed as a detector for unlabeled data. The inability of the model to generate a correct output for anomalous data is utilized to detect anomaly.

3. Methodology

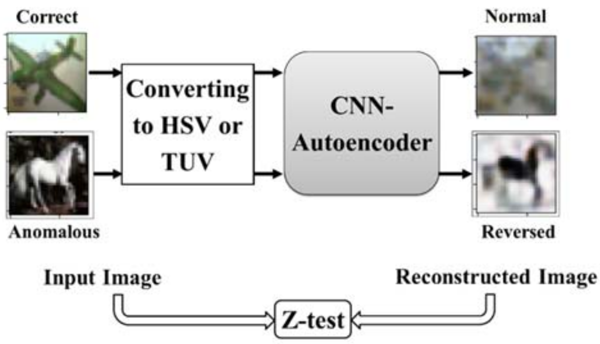


Fig. 2. Block diagram of training method**.**

The autoencoder reconstruct the input to the output even if the input was an anomalous data, and the mean square error (MSE) between input and output will be small in case of normal or an anomalous input, and the detection will be difficult especially for the anomalous one. Our goal is to maximize the difference of reconstruction error by reconstructing the anomaly classes reversely. Therefore, the MSE will be bigger and the anomaly detection will be easier as shown in figure 2.

3.1. Detection algorithm

The first step of the algorithm is to convert the training dataset images from RGB color model to HSV [10] or TUV (using equations illustrated below) color models.

$T = S \times \sin H,\ U = S \times \cos H,\ \ V = V$

where, H, S and V are the component of the HSV image. Secondly the anomalous data is reversed as a result of the next step. Consequently, the AE is trained using the new training dataset based on CNN.

The proposed training patterns for AE are as follow; 1- the first case the autoencoder is trained by class 0 as normal and other classes as anomalous, 2- the second case the classes 0 and 1 will be normal and other classes are reversed. Consequently, the number of normal classes will increase for each next case. Finally the autoencoder will be trained with all classes as normal for the last case.

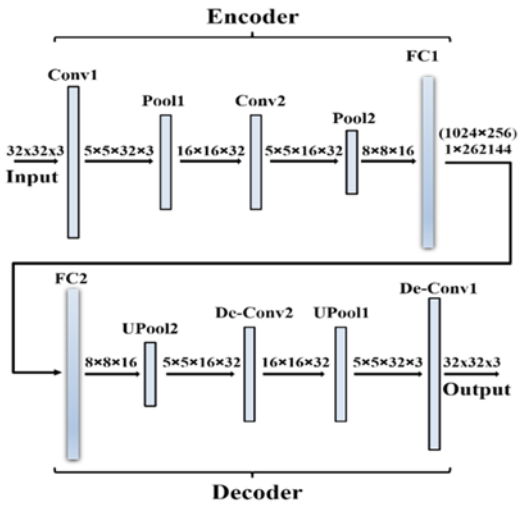


Fig. 3. Structure of convolutional neural network autoencoder.

Final step of the algorithm is to evaluate the performance of the AE using an inference dataset. The figure 3 illustrates the structure of CNN-autoencoder. As shown in the figure, the input image with size 32×32×3 is firstly convoluted in the first layer by a 5×5 filter. Consequently, the image dimensions are reduced through a pooling layer from size 5×5×32×3 to size 16×16×32. Next, another convolution layer is applied followed by a pooling layer to change the size of the image from 5×5×16×32 to 8×8×16. Finally, the encoding process is finalized with a fully connected layer with the output size of 1×262144 (1024×256). In order to decode the image, the reverse of the previous process is applied and finally a reconstructed image with the same dimensions as the input image is the same as the output.

3.2. Cifar-10 Dataset

The CIFAR-10 dataset is a set of images that can be used to teach a computer how to recognize objects, it contains RGB images with 32 by 32 pixels' size. It has ten classes and each class contains a different type of images. The dataset divides into a 50,000 images training set and 10,000 images testing set. Each set has an equal distribution of elements from each one of the ten classes[11].

3.3. Evaluation of performance for autoencoder

For significance validation, both F-test and Z-test were conducted. The Z-score value can be calculated based on the following formula [12]:

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Where $(\bar{x}_1, \bar{x}_2)$ are the average of the input and output pixels values respectively, (σ_1, σ_2) are the standard deviation for the input and output values respectively, (n_1, n_2) are the sample size of the input and output respectively.

A p-value is used in hypothesis testing to help accepting or reject the null hypothesis. It is evidence against a null hypothesis. The smaller the P-value, the stronger the evidence that you should reject the null hypothesis.

The used hypothesizes are; **if** P-value > 0.05 → There is no significant difference between the input-reconstructed image → Normal

if P-value < 0.05 → There is significant difference between the input-reconstructed image → Anomaly

4. Results and discussion

4.1. Training and testing loss

The difference between the test data input images and the reconstructed images were calculated in each epoch. The relationship between the testing loss and the epochs is shown in Figure 4. It is worth mentioning that the value of the test loss is almost similar or close to train loss value and this indicates to a good reconstruction process.

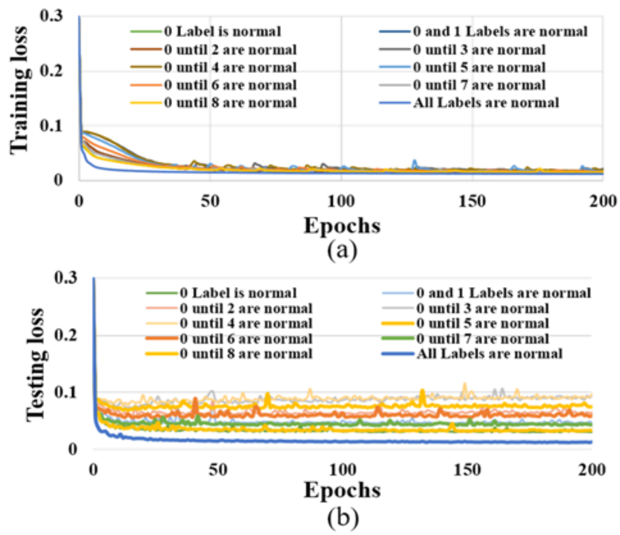


Fig. 4. Training and testing loss against epochs for HSV color model. (a) Training loss, (b) Testing loss.

4.2. Z-test

The P-value of each color model was calculated at zero mean value for each class and the results were shown in table 1 and table 2 for two cases; in the first one only class 0 is normal while the rest are anomaly, and in the second case all classes are normal.

Table 1. P-value for class 0 as normal and classes 1-9 as anomalous hypothesis for HSV, RGB and TUV.

class 0 normal, classes (1-9) anomaly	**Color model type**		
	HSV	**RGB**	**TUV**
class 0	0.5819	0.2013	0.2941
class 1	0.0057	0.0298	0.0012
class 2	0.000001	0.0100	0.0084
class 3	0.0076	0.0008	0.0142
class 4	3.09E-13	0.0222	0.0010
class 5	0.0005	0.0182	0.0109
class 6	0.0053	0.0322	0.0103
class 7	0.0213	0.0178	0.0478
class 8	0.0166	0.0424	1.4209E-12
class 9	0.0267	0.0323	0.0377

Table 2. P-value for class 0-9 as normal hypothesis for HSV, RGB and TUV color models.

All classes are normal	**Color model type**		
	HSV	**RGB**	**TUV**
class 0	0.4440	0.1899	0.1262
class 1	0.8405	0.6838	0.1086
class 2	0.6534	0.7221	0.0653
class 3	0.7256	0.8480	0.0853
class 4	0.3155	0.1685	0.0645
class 5	0.5290	0.1529	0.0792
class 6	0.8100	0.5489	0.0635
class 7	0.1732	0.4731	0.4457
class 8	0.5756	0.3557	0.2570
class 9	0.8039	0.8849	0.0792

From table 1 it is clearly that HSV in class 0 is better than RGB and TUV as the P-value is 0.5819 which is larger than 0.2013 and 0.2941. This because in the normal class the difference between input-output image should be small as proven by P-value result. In contrast, most of P-values of anomaly classes in RGB and TUV are bigger than in HSV which denotes the HSV detects the anomaly classes more effectively than other color models. Similarly, table 2 shows that, in general, HSV in all classes is better than RGB and TUV as the detection in HSV is more achievable than other models.

4.3. F-test

An anomaly detection performance is usually evaluating by employing the F-test using the recall and precision. Accomplishing high recall and high precision is not easy at the same time because the recall and precision goals are often conflicting.

Figure 5 shows the F-test against the threshold for HSV, RGB and TUV respectively. The first mean point in comparing with the previous work is the horizontal range for the F-test. This range in the proposed method was small, therefore, the correct threshold for the detection

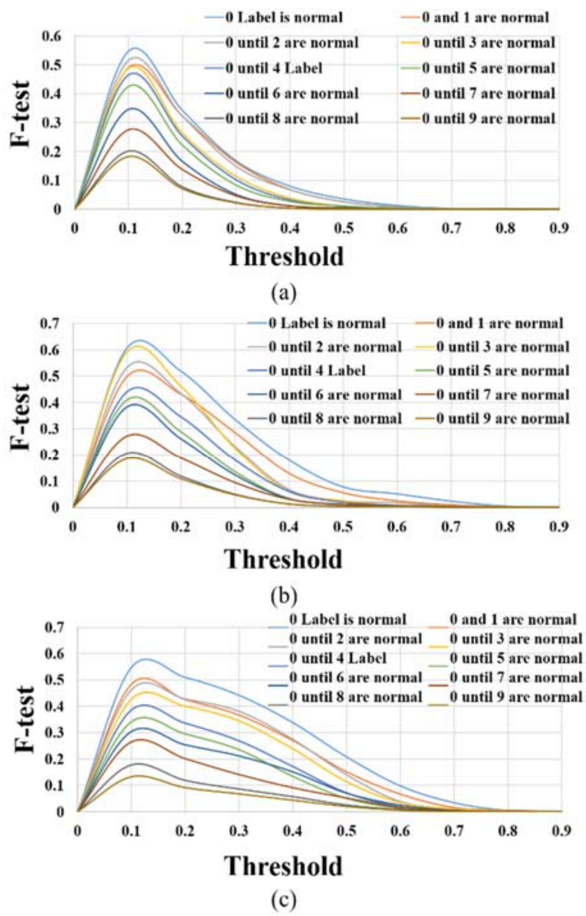


Fig. 5. (a) F-test for RGB, (b) F-test for HSV, F-test for TUV

could be defined easily. The second point is that F-value in the previous work was bigger in case of RGB than the value in our proposed method and that indicates increasing in the accuracy. In proposed mothed the F-test for HSV color model is better than other models,

Table 3 represents the comparison between our results and the previous results, the proposed CNN-autoencoder has been trained with three different color models HSV, RGB and TUV, whereas the stacked autoencoder trained with one color model which is RGB. Besides, the reconstruction quality results using the proposed CNN-autoencoder for the same color model (RGB) were better than the reconstruction results using the stacked autoencoder in spite of using the same dataset for the evaluation process.

Table. 3. General comparison between previous and proposed methods.

		Previous method [13]	Proposed method
Training Color Model		RGB	RGB, HSV and TUV
Structure		Stacked AE	CNN-AE
Dataset		Cifar-10	Cifar-10
Reconstruct-ion Quality	***RGB***	Moderate	Good
	HSV	NA	Good
	TUV	NA	Moderate
F-test	***RGB***	Good	Moderate
	HSV	NA	Good
	TUV	NA	Moderate
Z-test	***RGB***	NA	Moderate
	HSV	NA	Good
	TUV	NA	low

5. Conclusion

This research investigates the anomaly detection using CNN-autoencoder trained with three different color models. The trained AE has reconstructed the correct input normally, whereas the anomalous input has been reconstructed reversely. The results at 200 epochs training show that HSV color model has been more effective in anomaly detection rather than other models based on Z-test and F-test analyses.

Acknowledgements

This research was supported by JSPS KAKENHI Grant Numbers 17K20010.

References

1. A. Kaplan and M. Haenlein, "Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence," *Business Horizons*, 2019, vol. 62, pp. 15-25.
2. O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad, "State-of-the-art in artificial neural network applications: A survey," *Heliyon*, 2018, vol. 4, p. e00938.
3. W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, "A survey of deep neural network architectures and their applications," *Neurocomputing*,2017, vol. 234, pp. 11-26.
4. M. Y. Azar, V. Varadharajan, L. Hamey and U. Tupakula,"Autoencoder-based feature learning for cyber security applications," *IJCNN*, 2017.
5. 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*, 2017 pp. 665-674.
6. P. Fonseca, J. Mendoza, J. Wainer, J. Ferrer, J. Pinto, J. Guerrero, et al., "Automatic breast density classification using a convolutional neural network architecture search procedure," *in Medical Imaging*: *Computer-Aided Diagnosis*,2015, p. 941428.
7. Y. Sun, X. Wang, and X. Tang, "Deep learning face representation from predicting 10,000 classes," *in Proceedings of the IEEE conference on computer vision and pattern recognition, 2014,* pp. 1891-1898.
8. P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," *in Proceedings of the 25th international conference on Machine learning*, 2008, pp. 1096-1103.
9. V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM computing surveys (CSUR)*, 2009, vol. 41, p. 15.
10. L. Feng, L. Xiaoyu, and C. Yi, "An efficient detection method for rare colored capsule based on RGB and HSV color space," *in Granular Computing (GrC), 2014 IEEE International Conference*,2014, pp. 175-178.
11. A. Torralba, R. Fergus, and W. T. Freeman, "80 million tiny images: A large data set for nonparametric object and scene recognition," *IEEE transactions on pattern analysis and machine intelligence*, 2008, vol. 30, pp. 1958-1970.
12. Website:"https://statmagic.info/Content/Help Content/two-sample-prop.html", 29/07/2019.
13. A. Suzuki and H. Tamukoh, "Reverse Reconstruction of Anomaly Input Using Autoencoders, ", *Proceedings of Interntional Symposium on Intelligent Signal Processing and Communication Systems*,2018, pp. 431-435.

Emergence of Adaptive behavior in simulations by using Abstract Rewriting System on Multisets

Yasuhiro Suzuki
University Department, University Name, Address
City, State ZIP/Zone, Country[*]

E-mail: ysuzuki@nagoya-u.jp
www.ysuzuki.info

Abstract

We have developed an Artificial Intelligence system by using a model of chemical reaction, Abstract Rewriting System on Multisets, ARMS, where "intelligence" means that the reaction system can "select" specific molecules to sustain their reactions. We have implemented the reaction system by using an ARMS sand have obtained several molecules modified mutated DNA sequences that can sustain the reactions. We confirmed that reaction behaviors in the time series of concentration of non-mutated input molecule and mutated input molecule show oscillations; it would show that the system selects higher concentration one in between non-mutated and mutated one according to its concentration. Since the system exhibits adaptive autonomous behaviors, this DNA reaction networks system realize ARMS.

Keywords: List four to six keywords which characterize the article.

1. Introduction

DNA molecular computing has recently advanced to the study of molecular robots. On the other hand, artificial intelligence is shifting from research to social implementation. Shortly, molecular robots and artificial intelligence will fuse. The purpose of this research is to realize artificial intelligence using the underlying technology of DNA molecular computation.

For this reason, we use the most commonly used DNA molecular reaction called molecular displacement called strand displacement reaction. As for intelligence, we aim to create a function that can autonomously adapt to environmental changes. ARMS is a theoretical computation model based on the Gillespie algorithm, which is a rigorous physicochemical algorithm. In this research, the design for artificial molecular intelligence performed using physicochemical simulation.

1.1. *Self-sustained Chemical Reaction Networks, CRN*

We utilize the results of artificial life research that can adapt to environmental changes. In artificial life, various adaptive self-sustained autonomous reaction systems have been proposed, such as Eigen and Schuster's Hypercycle[1], Kauffuman's NK network[1], Fontana's lambda chemistry[1] and so on. In this research, we take the simplest Self-sustaining Chemical Reaction Network, CRN;

$$A + B \rightarrow C + D, \quad (1)$$
$$C + E \rightarrow A + F. \quad (2)$$

Here, the concentration of A keeps on sustaining, we assumed that the concentrations of B and E are sufficient and D and F flow out immediately and do not interfere with other reactions nor activities of other chemicals.

1.2. *Toehold mediated Strand Displacement Reaction, TMSD*

DNA strand displacement is a reaction between a double-strand DNA and single strand DNA, in which a strand in

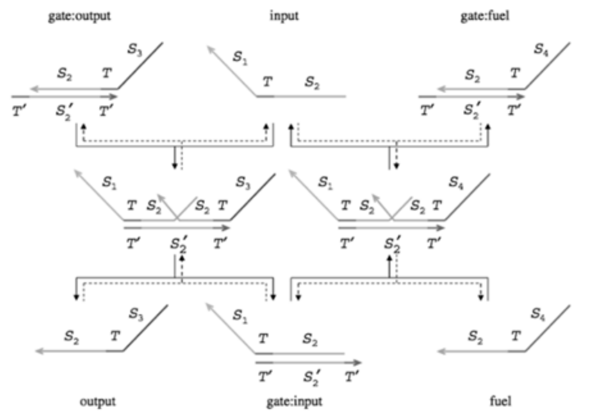


Fig. 1. Reprinted from the Journal of the Royal Society Interface, 2011 [4] (Open access)

a double-stranded DNA is replaced with the same or nearly identical single strand DNA. DNA strand displacement reaction is composed of the three single strands named the 'invader' 'incumbent' and the 'substrate'; this reaction is a swapping reaction between the invader and the incumbent strands on the substrate strand. The invader corresponds to an input signal, while the incumbent corresponds to an output signal. Toehold exchange reaction TER, which is Known as a class of strand displacements, since this reaction allows sequence- dependent control, it is peculiarly useful. In TMSD, the shorter incumbent forms a partial duplex with the longer, complementary substrate; then, the invader hybridizes with the toehold, which is the unbound region of the partially duplexed complement.

1.3. *Multiset Rewriting System on Multisets, ARMS*

We have been proposed a Chemical Reaction Networks, CRN, based on the Abstract Rewriting System, ARS; the ARS is a theoretical model of computation; we expand ARS on rewriting system on the multiset. A multiset is defined as a simple set and a map, which returns the duplication of element. We denote the duplication (multiplicity) of an element as M (a), for $a \in A$ and in case $c \notin A$, M (c) = 0; for example M (a) and M (b) of {a, a, b, b} are 2, and M (c) = 0; in the mathematical description, a multiset is described as; < sup, M () >, in which sup is a simple set of elements, in this paper we describe a multiset by denoting the same alphabet in its number of multiplicity such as {a, a, b, b} or a vector w =(M (a1) M (a2) . . . M (an)).

The union of two multisets M1, M2 is the same as the union of simple set and in vector description, the union of multisets is addition of vectors w1 and w2. And inclusion of sets is also the same as the simple set, when $M1(a) \leq M2(a)$ for all $a \in A$, the multiset M1 is included in M2 and we write $M1 \subseteq M2$.

A reaction rule is a pair of multisets, we denote A# as a set of all combinations of multisets over A and in the combinations, an empty multiset is included. A reaction rule $l \rightarrow r$, l, r $\in$ A# is described as a pair of multiset likewise chemical equations or a pair of its vector expression; and in some case, we can describe a reaction rule as a vector r, r = −l + r, it is simple and good for examining the dynamics of an ARMS, but this description cannot illustrate when there are the same species of element in the left-hand side and right-hand side such as a, b → a, c; in this case l = (1, 1, 0) and r = (1, 0, 1) and r = −(1, 1, 0) + (1, 0, 1) = (0, −1, 1).

A reaction is described as the rewriting of a multiset, if the left-hand side of a reaction rule is included in a multiset, these elements in the multiset are excluded and the right-hand side of the rule is merged to the multiset; the case when the multiset is a, a, b, b and the reaction rule is a, b → c, d, the left-hand side of the rule is included in the multiset, {a, b} $\subseteq$ {a, a, b, b} so the {a, b} is excluded from the multiset and it is transformed to {a, b} and the left-hand side of the rule {c, d} is merged to the set and we obtain {a, b, c, d} by this reaction.

Model

We model self-sustaining CRN with the seesaw gate reaction (Fig.1.) based on TMSD. In this reaction, first, TMSD occurs between the single strand DNA named' input 'and the double strand DNA named' gate:output

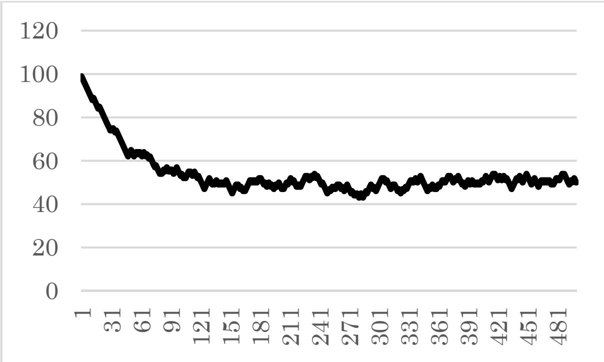


Fig. 2. Time change of the number of input molecules. The horizontal axis of the graph is the number of steps (one reaction is one step), and the vertical axis is the number of input molecules.

'and the single strand DNA named' output 'and the double strand DNA named " input:gate "are produced. Since the input:gate reveals a scaffold (T) for inter- action with the single stranded DNA named' fuel ', SDR occurs next on the fuel and input:gate, and the input and the gate:fuel are produced. Reactions of the first half and the second half occur like a seesaw, while the input consumed in the first half is supplemented in the second half; hence the reaction network sustains the input.

The self-sustaining CRN is implemented as;

input + gate:output ↔ input:gate + output (1),
input:gate + fuel ↔ input + gate:fuel (2).

In order to add perturbation to this reaction, an input 'molecule is introduced. input 'is an input molecule containing a miss-match pair. It is assumed that the binding force is weaker than the full match DNA strand as follows;

input’ + gate:output ↔ input’:gate + output (3),
input’:gate + fuel ↔ input’ + gate:fuel (4).

2. Method

We model and simulate the behavior of this reaction system using ARMS. This reaction system consists of eight reactions, including the reverse reaction. ARMS needs to determine the reaction rules to be applied, not random, but reasonably. Therefore, this model is regarded as a stochastic process model based on chemical kinetics (as for detail refer e.g. [3]).

We set the reaction rate constants as follows: the rate constants for the forward and reverse reactions are the same, the rate constant for the input' molecule is 100 times slower than the input molecule. This model assumes that the input' molecule has a mismatched pair. Therefore, we express this assumption by the difference in rate constants, which means that the affinity of strand matching will be lower, and the reaction will be slower.

3. Result

In the absence of perturbation, that is, without input', the reaction reaches an equilibrium state (Fig.2.). Add the

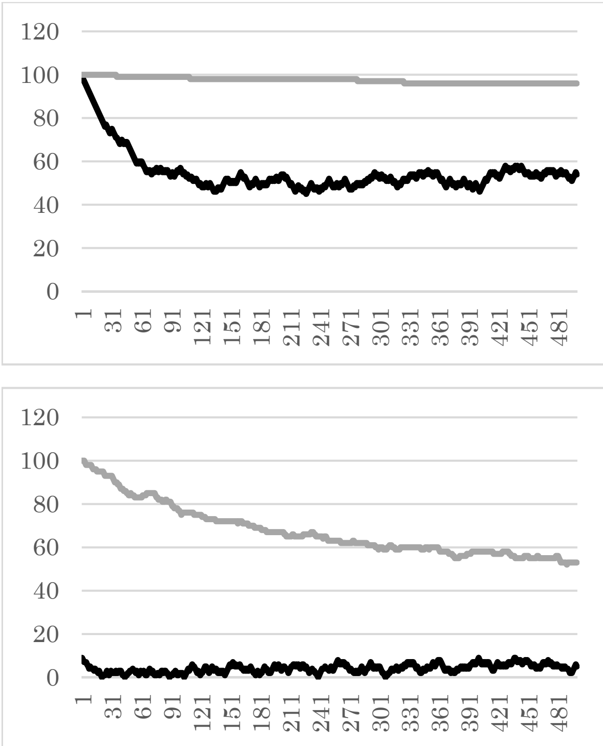


Fig.3. Time change of the number of input molecules and the number of input‘ molecules. When the same number of input and input’ molecule, input’ molecule does not react (top). (below) When the number of molecules containing mismatch is large, if the number of molecules of input' is 100 times that of input, the reaction system will use the molecules of input'. This result gives the DNA reaction system will be adaptive.

input 'molecule and add perturbation. If the input molecule and the input' molecule are equivalent, the input' molecule hardly reacts. This simulation result is obvious because the reaction speed of the input' molecule is 100 times slower than the input molecule. If

Fig. 4. 2-bit memory using DNA reaction. This memory realizes 2bit by on / off of input and input '. The equilibrium state of the input molecule and the input 'molecule is assumed to be (0, 0). And we regard the bit is set when the equilibrium state is changed by adding the input molecule or the input' molecule into reactions. In this simulation, at first i, which denotes the input molecules and i', input molecule, including mismatch, reached the equilibrium state. The equilibrium state is denoted as (0, 0) = (i, i'). Then 50 of input molecules added and (0, 0) transformed to (1, 0) and 50 of input molecules removed then (1, 0) back to (0, 0). Such adding and removing molecules realizes the states of bit change.

you set the number of input molecules to 100 times the amount of input molecules, the input molecules will be used. The simulation of this result predicts the environmental adaptation of a real DNA reaction system. If the input molecule containing the mismatch is high, the reaction system will use the molecule containing the mismatch (Fig.5.).

4. Discussion

The simulation results show that when the amount of input molecules is half the amount of input 'molecules, the concentrations of both become the same in the equilibrium state, and the equilibrium state shifts when the number of molecules is increased. Realize memory using this reaction system.

The case where the amount of the molecule is in the equilibrium state is 0, and the case where the equilibrium state is shifted by adding the molecule is 1. This method can realize n-bit, but for simplicity, this paper shows the example of 2-bit (Fig.4,5.). We confirmed through simulation that the DNA memory works. At first, when 30 molecules were added to the equilibrium state of the input and input 'molecules at the same amount, the shift of the equilibrium state was confirmed (Fig.5.). Then we confirm, do the bit on/off operations? Therefore, the input molecule and the input molecule were added and deleted. All states of input and input 'bit are (0,0), (0,1), (1,0), (1,1). It has confirmed that the transition between the four status bits can realize by adding and removing the numerator. We know that performing "removal" operations on real DNA reaction systems is not realistic.

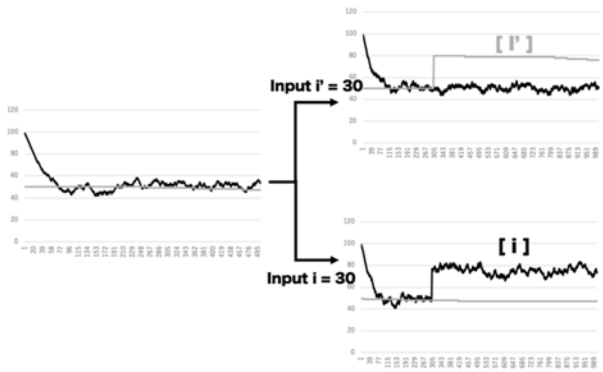


Fig. 5. This is the caption for the figure. If the caption is less than one line then it is centered. Long captions are justified manually.

However, it is more realistic to delete molecules as a result of using them instead of "deleting" them, for example, the DNA circuit by using this technique (Fig.4). In this paper, we show the realization of bit, but also show that this technology can realize "internal state" by input/output, which means that molecular artificial intelligence can be realized.

References

1. C. Adami, *Introduction to Artificial Life*. Springer Verlag 1997.
2. L. Qian, E. Winfree, A simple DNA gate motif for synthesizing large-scale circuits, Journal of the Royal Society Interface, Vol. 8, 2011.
3. M. Umeki and Y. Suzuki, A Simple Membrane Computing Method for Simulating Bio-Chemical Reactions, Computer and Informatics, 27, 529-550, 2008.

Implementing the Euler and Runge-Kutta method by using Abstract Rewriting System on Multisets

Yasuhiro Suzuki
*Department of Complex Systems Science, Graduate School of Informatics, Nagoya University, Furocho Chikusa Nagoya City, 464-8601/Japan**

E-mail: ysuzuki@nagoya-u.jp
www.ysuzuki.info

Abstract

In this paper, we show that by using a model of chemical reaction, Abstract Rewriting System on Multisets, ARMS, Euler method, and Runge Kutta method are implemented smoothly. ARMS is a flexible computational model, and it enables us to implement Multi-Agent Systems or P Systems quickly. Hence, by using the proposed method, we can apply the Eular or Runge Kutta method for them. In this paper, we take the Lotoka-Volterra model, for example, of the Multi-Agent system and show how to implement by using Euler and Runge-Kutta method then compare these results.

Keywords: List four to six keywords which characterize the article.

1. Introduction

Discrete mathematics does not have a theory equivalent to the dynamic system of physics-mathematics. However, discrete mathematical models of rule dynamics, such as Multi Agent Systems, are used in areas such as artificial intelligence, artificial life, complex systems, and mathematical economics. Since discrete mathematics does not have the theory of dynamical systems, we cannot treat rule dynamics theoretically. Therefore, all research on rule dynamics is simulation. The results of simulation differ depending on the conditions, and the results are probabilistic because many rule dynamics are stochastic. In artificial life and complex systems, all simulations are rule dynamics, but we cannot believe the results.

The reason we can't believe the results of the simulation is that the rules are applied randomly. In natural systems, rules are generally not applied randomly. For example, in ecosystems the application of rules depends on the size of the population. In a chemical reaction system, the application of the rule changes depending on the concentration. Congestion also results from the car moving anisotropically to avoid collisions, that is, the probability of movement is not random.

Therefore, almost all rule dynamics systems are inherently wrong, because they apply rules randomly. To solve this essential problem, I brought chemical kinetics into rule dynamics. Then, we formulated the rule dynamics as an abstract system of chemical reactions, which is the Abstract Rewriting System on Multisets, ARMS[1].

ARMS is a computational model of discrete mathematics, Abstract Rewriting System, an abstract chemical system based on ARS and physical chemistry. ARMS is a hybrid computing system based on both discrete and continuous mathematics. Therefore, a small number of molecules can be regarded as a discrete system, and a large number of molecules can be regarded as a continuous system.

The rule dynamics is constructed with ARMS, and then a large number approximation yields a differential equation. The differential equations obtained by this large number approximation can be analyzed by the theory of dynamical systems. Therefore, the large number approximation of ARMS avoids the inherent difficulties of rule dynamics.

Furthermore, solving the differential equation numerically gives an analytical solution of ARMS. Note, however, that ARMS is similar to Gillespie's method[2], and Gillespie already mentioned in his paper that Gillespie's method can be solved by Euler's method or Runge-Kutta method.

To my knowledge, Gillespie does not state that it can actually be solved by Euler's method or the runaway method. In this essay, we show how to solve ARMS by Euler's method or Runge-Kutta method.

2. Abstract Rewriting System on Multisets, ARMS

We have been proposed an artificial chemistry, based on the Abstract Rewriting System, ARS; the ARS is a theoretical model of computation; we expand ARS on rewriting system on the multiset. A multiset is defined as a simple set and a map, which returns the duplication of element. We denote the duplication (multiplicity) of an element as M (a), for a ∈ A and in case c /∈ A, M (c) = 0; for example M (a) and M (b) of {a, a, b, b} are 2, and M (c) = 0; in the mathematical description, a multiset is described as; < sup, M () >, in which sup is a simple set of elements, in this paper we describe a multiset by denoting the same alphabet in its number of multiplicity such as {a, a, b, b} or a vector w =(M (a1) M (a2) . . . M (an)).

The union of two multisets M1, M2 is the same as the union of simple set and in vector description, the union of multisets is addition of vectors w1 and w2. And inclusion of sets is also the same as the simple set, when M1(a) ≤ M2(a) for all a ∈ A, the multiset M1 is included in M2 and we write M1 ⊆ M2.

A reaction rule is a pair of multiset, we denote A# as a set of all combinations of multisets over A and in the combinations, an empty multiset is included. A reaction rule l → r, l, r ∈ A# is described as a pair of multiset likewise chemical equations or a pair of its vector expression; and in some case, we can describe a reaction rule as a vector r, r = −l + r, it is simple and good for examining the dynamics of an ARMS, but this description cannot illustrate when there are the same species of element in the left-hand side and right-hand side such as a, b → a, c; in this case l = (1, 1, 0) and r = (1, 0, 1) and r = −(1, 1, 0) + (1, 0, 1) = (0, −1, 1).

A reaction is described as the rewriting of a multiset, if the left-hand side of a reaction rule is included in a multiset, these elements in the multiset are excluded and the right-hand side of the rule is merged to the multiset; the case when the multiset is a, a, b, b and the reaction rule is a, b → c, d, the left-hand side of the rule is included in the multiset, {a, b} ⊆ {a, a, b, b} so the {a, b} is excluded from the multiset and it is transformed to {a, b} and the left-hand side of the rule {c, d} is merged to the set and we obtain {a, b, c, d} by this reaction. By using vector expression, a reaction is the addition of vectors as w −l + r, in which l and r are vector expression of a reaction rule; the case when M = {a, a, b, b}, w = (2, 2, 0, 0) and the reaction rule is a, b → c, d, ({1, 1, 0, 0}, {0, 0, 1, 1}); and the reaction is denoted as {2, 2, 0, 0} − {1, 1, 0, 0} ={1, 1, 0, 0} and {1, 1, 0, 0} + {0, 0, 1, 1} = {1, 1, 1, 1}.

3. Euler method and Runge-Kutta method

Euler's method is a numerical method for solving ODEs (see the class book for details). Euler's method uses initial values or differential equations $y' = f(t, y)$ and $y(t_0) = y_0$. time t is discretized as $t_n = t_0 + nh$, $y_{n+1} = y_n + hf(t_n, y_n)$. If the function is sufficiently smooth and it is close to h from $y(t_i)$, the exact solution is $y(t_i + h) = y(t_i) + f'(t_i)h + , ...$" (Taylor expansion), in which case the Euler method used in the right-hand side items after third term ignored.

The Runge-Kutta method (4th order) uses the Taylor expansion and the 4th order term, which is more accurate than the Euler method, and is defined as follows.

$$
\begin{aligned}
y_{i+1} &= y_i + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \\
k_1 &= hf(x_i, y_i) \\
k_2 &= hf(x_i + \frac{h}{2}, y_i + \frac{k_1}{2}) \\
k_3 &= hf(x_i + \frac{h}{2}, y_i + \frac{k_2}{2}) \\
k_4 &= hf(x_i + h, y_i + k_3).
\end{aligned}
$$

4. Solving Lotoka-Volterra equation by ARMS

Lotoka-Volterra equation, LV is the fundamental equation of mathematical ecology. Expressing LV eq. In rule dynamics is as follows;

herbivore -> herbivore, herbivore,
herbivores, carnivores-> carnivores, carnivores,
carnivores->.

These rule dynamics modeled by ARMS, that is, represented by a stochastic process of reaction kinetics. In the following, the number of individuals described with []. The denominator function, D([herbivore], [carnivore]) of the stochastic process, that is, all combinations of molecular collisions for a reaction to occur, is

D([herbivore], [carnivore]) = k_1 [herbivore] + k_2 [herbivore] [carnivore] + k_3 [carnivore].

a, b, c are reaction rate constants for each reaction. And the reaction probability function vector, Vrec is

Vrec={1/ D([herbivore], [carnivore])}(k_1[herbivore], k_2[herbivore][carnivore], k_3[carnivore]) .

The expression of the stochastic process of LV eq was obtained. We can simulate this stochastic process. Simultaneous approximation gives simultaneous differential equations.

4.1. *Deterministic ARMS with Euler method*

The unit of the reaction rate constant is mol / sec. Therefore, multiplying the reaction rate constant by sec. gives the amount of mol that changes. We denote [herbivore] is a, [carnivore], b and the vector of (a, b) at the t_j step, $M(t_j)$ and sec.,τ;

$$
\begin{aligned}
M(t_{j+1}) &:= M(t_j) + \left(ak_1\tau \begin{pmatrix} 1 \\ 0 \end{pmatrix} + abk_2\tau \begin{pmatrix} -1 \\ 1 \end{pmatrix} + bk_3\tau \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right) \\
&= M(t_j) + \left(a \begin{pmatrix} 0.1 \\ 0 \end{pmatrix} + ab \begin{pmatrix} -0.2 \\ 0.2 \end{pmatrix} + b \begin{pmatrix} 0 \\ -0.1 \end{pmatrix} \right) \\
&\qquad (a = M(t_j)_a, b = M(t_j)_b).
\end{aligned}
$$

, where the second term denotes that when $\tau = 0.1$, $k_1=k_3=1$ and $k_2=2$.

Deterministic ARMS results are uniquely determined, because Deterministic ARMS is not a stochastic simulation (Fig.1.).

The rule dynamics LV model is converted into a kinetic equation as;

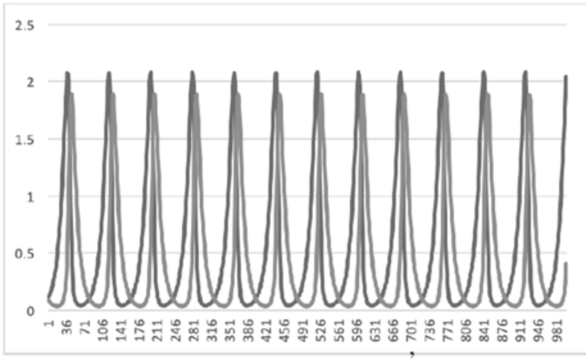


Fig. 1. Deterministic ARMS results. The vertical axis is the number of individuals a and b, and the horizontal axis is the number of steps.

$d[a]/dt = k_1a - k_2ab$,
$d[b]/dt = k_2ab - k_3b$.

Solving this simultaneous differential equation by Euler's method gives

$$
y_{n+1} = y_n + hf(t_n, y_n) = y_n + \tau \begin{pmatrix} k_1a - k_2ab \\ k_2ab - k_3b \end{pmatrix} = y_n + \begin{pmatrix} (k_1\tau)a - (k_2\tau)ab \\ (k_2\tau)ab - (k_3\tau)b \end{pmatrix}
$$

it is the same as Deterministic ARMS. Deterministic ARMS was shown to be consistent with Euler's method. As Gillespie has already pointed out, this is self-evident because a large number approximation to a stochastic ARMS yields a differential equation.

4.2. *Deterministic ARMS with Runge-Kutta method*

Euler's method is a simple numerical solution of differential equations. Therefore, we solve Deterministic ARMS using Runge-Kutta method, which is a numerical method with higher accuracy.

Scheme of solving LVeq.by Deterministic ARMS with Runge-Kutta method is as follows;

$$\Delta_1 = \tau\left(k_1 a\begin{pmatrix}1\\0\end{pmatrix} + k_2 ab\begin{pmatrix}-1\\1\end{pmatrix} + k_3 b\begin{pmatrix}0\\-1\end{pmatrix}\right)$$

$$\Delta_2 = \tau\left(k_1(a+\Delta_1^a\times\frac{\tau}{2})\begin{pmatrix}1\\0\end{pmatrix} + k_2(a+\Delta_1^a\times\frac{\tau}{2})(b+\Delta_1^b\times\frac{\tau}{2})\begin{pmatrix}-1\\1\end{pmatrix} + k_3(b+\Delta_1^b\times\frac{\tau}{2})\begin{pmatrix}0\\-1\end{pmatrix}\right)$$

$$\Delta_3 = \tau\left(k_1(a+\Delta_2^a\times\frac{\tau}{2})\begin{pmatrix}1\\0\end{pmatrix} + k_2(a+\Delta_2^a\times\frac{\tau}{2})(b+\Delta_2^b\times\frac{\tau}{2})\begin{pmatrix}-1\\1\end{pmatrix} + k_3(b+\Delta_2^b\times\frac{\tau}{2})\begin{pmatrix}0\\-1\end{pmatrix}\right)$$

$$\Delta_4 = \tau\left(k_1(a+\Delta_3^a\times\tau)\begin{pmatrix}1\\0\end{pmatrix} + k_2(a+\Delta_3^a\times\tau)(b+\Delta_3^b\times\tau)\begin{pmatrix}-1\\1\end{pmatrix} + k_3(b+\Delta_3^b\times\tau)\begin{pmatrix}0\\-1\end{pmatrix}\right)$$

$$\delta M(t_j) = \frac{\Delta_1 + 2\Delta_2 + 2\Delta_3 + \Delta_4}{6}$$

The results of the Deterministic ARMS with Runge-Kutta method are more accurate than the Euler method (Fig.2.).

5. Conclusion

ARMS is a stochastic process system, but it is a deterministic system focusing on the unit of the reaction rate constant. Deterministic ARMS is similar to a differential equation. Therefore, numerical solution of differential equations can be used to solve Deterministic ARMS, such as Euler method and Runge-Kutta method.

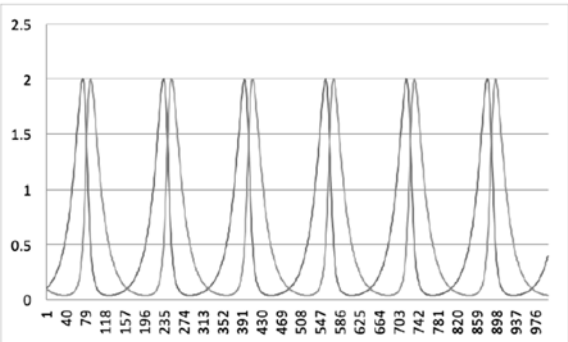


Fig. 2. Deterministic ARMS results with Ruge-Kutta method. The vertical axis is the number of individuals a and b, and the horizontal axis is the number of steps.

This paper shows the effectiveness of considering the rule dynamics system as a chemical reaction system. There are three reasons. First, a stochastic process of reaction kinetics can be obtained from rule dynamics. Second, a system of differential equations can be obtained from a stochastic process system. Of course, you can formulate a differential equation directly without going through a stochastic process. It is common practice in physics and chemistry to formulate a stochastic process (Chapman-Kolmogorov equation or Master equation) from a chemical system and obtain a continuous system from a large number approximation. This physicochemical ordinary way gives reasonably stochastic and continuous systems from the rule dynamics. This method is effective when large numbers cannot be approximated. If large number approximation cannot be performed, the simulation of the stochastic process system is performed. The stochastic system for this can be obtained from the physicochemical ordinary way.

However, when there are spatial fluctuations in the case of a few systems, specifically when molecules are unevenly distributed in the space, it is necessary to discuss whether it is OK to use ordinary differential equations, and it is necessary to use stochastic partial differential equations In some cases. In that case, for example, a cellular ARMS system in which ARMS is diffusion-coupled can be used. When you build a cellular ARMS, you can approximate it to a PDE or a cellular automaton composed of ordinary differential equations. In that case, the computational load generally increases, but the computational load can be dramatically reduced by using deterministic ARMS.

References

1. Yasuhiro Suzuki, et.al, (1996), Symbolic Chemical System based on abstract rewriting system and its behavior pattern, Journal of Artificial Life and Robotics, 1997, vol1, 211-219.
2. Gillespie, Daniel T. (1977). "Exact Stochastic Simulation of Coupled Chemical Reactions". The Journal of Physical Chemistry. 81 (25): 2340–2361.

Extracting tactile sensation from body movement and converting it into vibrotactile using the Tactile Score bit

Yasuhiro Suzuki
Department of Complex Systems Science, Graduate School of Informatics, Nagoya University,
*Furocho Chikusa Nagoya City, 464-8601/Japan**
E-mail: ysuzuki@nagoya-u.jp
www.ysuzuki.info

Abstract

Body motion speed and acceleration are extracted using motion capture and converted to tactile score bits and TS bits by discretizing the magnitude. The TS bit is Sensitivity information, which is information that can obtain through discretizes the extent and the duration of its size for a specific feature value over time. By using the TS bit, the sensitivity can extract from the temporal change in function. In this paper, we will introduce how to convert a dancer's body movement to the TS bit and TS bit to vibrotactile.

Keywords: Tactile Score, TS bit, vibrotactile, SHOKKAKU

1. Calculus of Senses

Our sense be a dynamical system in multi-dimensional sensory space, where each dimension corresponds to the sort of sense such as visual, audio, tactile, and so on. Our sensory systems are composed of differentiators and we perceive sensory stimulations through differentiating sensory stimulations.

In order to calculate differentiation of sense, we have to set measurable quantity of the sense, we can use not only quantity measure such as brightness of picture, loudness of sound but also can use quality measure such as value of Semantic Differential, SD method for measuring impressions or emotion such as beautiful, fear, happy.

The most important concept of the differentiation of sense is measuring the differences in time or / and space of sense; it shares the concept of differentiation in mathematics but this is not mathematical concept and a mathematical equation of sense is not required; in calculating differentiation of sense, we use differentiable measure. Algorithm of differentiation of sense is as follows;

I. Differentiation of Sense
1. select a sense to calculate differentiation,
2. set the width of differentiation,
3. measure the differences by the time width (2.);
3-1 (if it is required) measure of the differences of 3.
For example, we will differentiate a music piece;
1. a sense to calculate differentiation = the frequencies of sound,
2. the time width of differentiation = 0.5 second,
3. the differences by the time width (2.);
and by setting the duration of quarter note is 0.5, we obtain the music score of D, D, G, G; this differentiation of sound corresponds to drawn the music piece in musical notation. Then we differentiation the sound measure and

obtain the spectral pitch, which is the amount of vibration at each individual frequency and it corresponds to pitch of music note (Fig.1.).

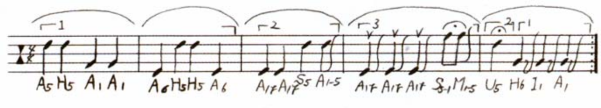


Fig. 2. A Tactile Score, where is generally described using staff notation. Special symbols are used as needed. This score is for massage, and each note has a logo. At the top of the staff notation, a key symbol and a number are attached, which means that the movement of the right hand and the left side is shifted, 1 means simultaneous movement, two means slight movement, and three means massive action.

II. Transform Sense

1. select a differentiated sense and the sense to transform (e.g. the volume of sound into the brightness of movie)
2. transform the sense and reconcile time width each other,

Transform sensory such as audio sensory to visual has been developed well, for example, the sound-on-film recording, which is the most prevalent method of recording analog sound on a film print; in short, transform the volume of sound (audio sense) into brightness (visual sense) of a film and synchronize the sound and film.

III. Integral of Sense

And also once a sense is differentiated we can calculate integral of sense as follows;

1. select the differentiated sense (time width were defined),
2. define the interval of calculating integral.
3. calculate an area of sense at each time width and calculate the sum of areas in the interval.

2. Tactile Score

Sensitivity can be differentiated, as shown above. The difference is discretization, and discretization is computerization. For example, a cheerful "Hello," there is no difference as a character language between the absence of energetic "Hello." However, it is a difference in spoken word, and the difference is sensitivity to sense the cheerful "Hello," performed by the difference of sensibility. Discretizing the change of voice volume cheerful "Hello," the voice volume increases diverging.

On the other hand, there is no healthy, "Hello," the voice volume becomes smaller with time. In other words, by making the time change of the voice volume into a difference, it is possible to computerize the feelings of energy and lack of energy.

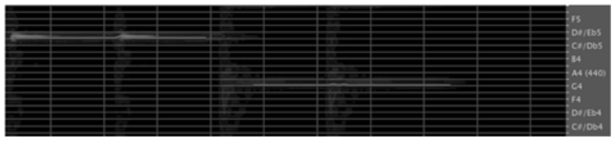


Fig. 1. Spectral pitch, the horizontal axis illustrates time and the vertical, frequencies that correspond to pitches of music notes. D, D, G, G and duration of each note is the same as 0.5 second.

The staff notation used to describe differentiated sensitivity (Fig.2.). The reason for using the staff notation is that it is readily available, and there is no essence in discretizing it into five or seven levels. The center of the staff, that is, the third line is the standard voice volume. The ascending lines are small, and the descending lines are loud. The length of each of the vowels of "Hello" is represented by the note. Although an example of how to say "hello" in this description, generally time variation of the intensity and size, all can be described by this method.

The musical score that describes this differentiated sensitivity is called Tactile Score[1] (Fig.2.). Also, the information specified by Tactile Score with the Sensitivity differentiated called Tactile bit and T-bit. T-bit is information that expresses sensitivity.

By performing the differentiation of the solid sensibility described in the previous section and expressing it with T-bits, arbitrary Sensitivity can turn into information.

3. Art Performance by using T-bit

The application range of the difference of Sensitivity and the conversion to the score is wide. For example, it is possible to differentiate body movements and convert them into tactile scores. There are various ways to do this. As an example, motion capture of the body movements of a dancer. Measure the X, Y, and Z coordinates of three points: wrist, ankle, and elbow. The center of gravity of X, Y, and Z obtained, and the speed of body movements obtained from a distance between the centers of gravity during Δt. Then the acceleration obtained from the change in speed. The velocity and acceleration are discretized into five levels and converted into tactile scores.

On the other hand, prepare multiple vibrotactile senses. Vibration and tactile sensation are assumed to have a substantially constant amplitude change over time. Then, the amplitude changed according to the score extracted from the dancer's movement. By measuring body movement, we can make it possible to obtain the score from the dancer's movement and change it into a vibrotactile sensation. This method can be used not only for dancers but also for moving objects in general.

SHOKKAKU 2019 was held at Kyoto University Graduate School of Human and Environmental Studies, where Butoh dancer Norihito Ishi (San-Kai-Jyuku, company) and composer Eric Maestri have created works using this method. We captured the motion of Ishii's body and converted it into a score then the tactile score turned to vibratory haptics. The generated vibrotactile signals are audio signals where E. Maestri applied a sound effect to the sound vibration generated as if to perform live and created a vibration tactile sensation to be presented to the venue and the audience.

The converted tactile sensation used to vibrate the audience's face through the Haptic Vibration Display (HVD, Fig.4.). Furthermore, a vibrator was attached to the percussion instrument Thunder Sheet, and the Thunder Sheet vibrated at a high volume.

The HVD originally developed to display vibration haptics, Tactile Score Vibration, and TSV generated from tactile scores. HVD contains multiple sound exciters. A sound exciter is an oscillator that responds to frequencies below the mid-range of the input audio signal. The configuration is almost the same as a general sound speaker, except that only a cone paper removed from the sound speaker. We use Dayton Audio's sound exciter. The input signals of our HVD are left and right channels for music and vibrotactile sense, respectively.

Thunder Sheet is a particular type of percussion instrument used to play the sound of thunder, as used in Richard Strauss's Alpine Symphony. Thunder Sheet has multiple shapes, we used 2M in height, 1M in width, and the material is a thin metal plate. We usually play by tapping with a mallet, but we attached a sound exciter to the Thunder Sheet and vibrated. By vibrating at a high volume, a 2-square-meter metal vibrating plate vibrates, so that tactile sound can be generated.

Fig. 3. From SHOKKAKU 2019, SHOKKAU 2019 was held at Kyoto University Graduate School of Human and Environmental Studies from June 29 to 30, 2019. In the center of the photo is Ishii performing his work of Butoh “empty garden”. The mask worn by the audience is a Haptic Vibration Display. The iron plate at the top center of the photo is the Thunder Sheet, and four sheets were used.

Fig. 3. From SHOKKAKU 2019, SHOKKAU 2019 was held at Kyoto University Graduate School of Human and Environmental Studies from June 29 to 30, 2019. In the center of the photo is Ishii performing his work of Butoh “empty garden”. The mask worn by the audience is a Haptic Vibration Display. The iron plate at the top center of the photo is the Thunder Sheet, and four sheets were used.

In the venue, Thunder Sheets were arranged one by one on four sides to surround the audience. And 40 HVDs were used. This allows the viewer to indirectly sense the vibratory haptics generated from Ishii's body movements from the ears and body through thunder Sheets, and directly vibrate the face with HVD. Since the HVD has open eyes, the audience can watch it while wearing it. The HVD can also wear on the head.

4. Brille Tactile Score

T-bit is also useful for barrier-free engineering. In particular, it has excellent potential as a communication technology for the visually impaired. The visually impaired mainly communicates in Braille. Braille is a character media, and there is a limit to the expression of emotions. Even though pure emotions can be expressed, it is challenging to express delicate feelings. Besides, it is not possible to represent the "Hello" without a cheerful "Hello" and Genki previously described only in Braille. Many people who are visually impaired can speak. Therefore, emotional expression is possible by voice. However, it is difficult for people with visual or hearing impairment or deafblindness to speak, and in that case, there is a significant restriction on the expression of emotions.

We worked with Norimasa Kawasora, a deafblind person, to create a Braille Tactile Score.

Braille Tactile Score is a vector of three numeric elements: touch, strength, and size. The three components are at the level of each element. Touch has four levels from good to harmful; the force has five degrees from stable to weak, and size has ten levels from small to large.

By using ordinary Braille and this Braille Tactile Score, deafblind people can express their emotions and sensibilities. At SHOKKAKU 2019, K. Norimasa showed “rain” and “snow” in Braille Tactile Score. His work, "Knowing the Rain," consists of a sentence expression in Braille and its Braille Tactile Score.

"Knowing the Rain" by Norimasa Kawasora

Light rain starts: (1, 1, 1)
A small and gentle touch,: (1, 1, 2)
The rain is getting stronger, and I don't know when to stop. : (1, 1,4)
From the light rain, the rain falls on the skin. : (4, 4, 4)

Gradually. : (4, 4, 4)
4 4 4

The rain was said to be fine the next day when it was in the rainy season compared to the snow, and the following day was fine: (3, 3, 3)
It can be a sudden heavy rain. Rain continues. : (2, 2, 1)
The rain gradually stops at last. : (1, 1,1)

Then, the following Braille Tactile Score is obtained.
1 1 1
1 1 2
1 1 4
4 4 4
4 4 4
3 3 3
2 2 1
1 1 1

We convert Braille Tactile Score into a vibrotactile sensation. We transform the elements of the Braille Tactile Score, "touch," "strength," and "magnitude," respectively, into the waveform, magnitude of the amplitude, and pitch of frequency.

"Touch" is the waveform of the vibrotactile sensation (from smooth to jagged), "Strength" is the magnitude of the amplitude, "Size" corresponds to high and low frequency (large is low frequency, small is high frequency).

Kawasora's "Kowing the Rain" converted to vibrotactile sensation. And this work was presented by using HVD and Thunder Sheets performed at SHOKKAKU2019.The text is to be typeset in 10 pt Times Roman, single spaced with interline spacing of 13 pt. Text area (excluding running title) is 6.75 inches across and 8.8 inches deep. Final pagination and insertion of running titles will be done by the publisher, so make sure that no page numbers are given in your paper and only the running titles provided in this template (authors’ names and paper title) are used.

Acknowledgement

This research supported by SPIRITZ, Kyoto University, and the author would like to thank, for the co-production at SHOKKAKU2019, to Butoh dancer, Norihito Ishii, Eric Maestri, and executive producer Professor Masaaki Sakagami (Kyoto University).

References

1. Y.Suzuki and R.Suzuki (2013) *Tactile Score*, Springer Verlag

A Tactile Sense Centered Virtual Reality game by using Biometric feedback

Yoshihito Ushida
Department of Imformatics, Nagoya University, Furocho, Chikusa
*Nagoya City, 464-8601/Japan**

Yasuhiro Suzuki
Department of Complex Systems Science, Graduate School of Informatics, Furocho Chikusa
Nagoya City, 464-860, Japan
E-mail: ysuzuki@nagoya-u.jp
www.ysuzuki.info

Abstract

In this research, we try to construct a Virtual Reality, VR system centered on the tactile sensation that has used as a secondary. Tactile sensation has a more significant effect on sensibility than audiovisual. Therefore, if tactile sensation and audiovisual sense are combined, presence can be given to a virtual object in the VR space. This system makes a user's biological information accessible in VR space. Then, a virtual object linked to the audiovisual data presented. A player has to use the virtual object to play this game with physical body movements. And the system feeds back biometric information and deliver the change of biometric data through the sensory presentation to the user.

Keywords: List four to six keywords which characterize the article.

1. Introduction

In conventional Virtual Reality (VR) technology and applications, visual information is mainly used, and auditory and tactile information have been used side by side. In this research, we will construct a VR system that focuses on tactile sense and secondary to audiovisual perception. The tactile sense has a strong influence on sensitivity as compared with the audiovisual sense. Therefore, using the tactile as the center, we aim to give the presence to the virtual object in the VR space. In order to do that we measure the user's biometric information, make it tangible, and create a virtual object in cooperation with audio-visual information. Through feedback sensory presentation via biometric information, virtual objects to generate a sense that users are cooperating with virtual objects.

2. Virtual Reality technology

Conventional Virtual Reality (VR) technology and applications have been developed by using auditory and tactile sensations as a secondary focus on visual information. In this research, we try to build a VR system that uses audio-visual as a secondary function, mainly on tactile sensation. Since the tactile sense has a stronger influence on the impression than the audio-visual sense, the purpose is to give a sense of presence to the virtual object in the VR space by using the audio-visual sense secondary to the tactile sense. Therefore, by making the measured biological information of the user tactile and presenting it as a virtual object associated with the audiovisual data, the biological information and the presentation of the sensation are fed back to generate a feeling that the virtual object is associated with the user.

2.1. *Virtual Reality in Entertainment and Media Arts*

2016 is called Virtual Reality, the first year of VR.[1] Multiple types of home VR Head Mount Display, HMD (Oculus Rift, HTC Vive, PlayStation®VR, etc.) are released, and many VR commercial facilities are also available. Open (eg[3]). With the rapid spread of such VR technology into society, VR technology is beginning to permeate everyday life.

Most of the current VR technology uses audiovisual stimuli, and tactile stimuli are in the process of development. For example, many VR gloves that provide real-time tactile feedback in conjunction with audiovisual stimuli have developed (e.g.,[4,5,6]), but in many cases, the direct tactile presentation has performed. Instead of these techniques, in several systems, the player touches the system itself directly, and the users instantly shook under the 3D image.

Further, a tactile sensation presented using a force sense generated by a change in acceleration (for example, [7,8]). On the other hand, game systems using only tactile stimulus presentations are beginning to put into practical use. Nintendo Switch ™, a game system from Nintendo released in 2017, has been accepted by society as a game that can be established only with a tactile presentation by vibrating vibration without providing audiovisual presentation (April to September 2017) Sold 4.89 million units[3]). In the 1-2-Switch, which is the launch title of the system, the game controller regarded as a small box with marbles, and by tilting it, the vibration display simulating the rolling and collision of the sculptures the number of statues in the small box is displayed. The tactile information used as the main sensation presentation and the audiovisual data used as a secondary one. The immersion is low because no monitor or Head Mount Display, HMD used.

The presentation of tactile information is considerably less than visualization, but attempts have made to make information accessible for visually impaired people[10] and to make vital information available. For example, Watanabe and his colleagues have created a workshop aimed at making the movement of the heart (heart sound) tangible and touching it to re-recognize the self as life[9]. In this workshop, "Heart Picnic," participants listen to their own heart sounds with a stethoscope, and at the same time, touch the heart sounds made tunable with a vibrotactile presentation device. Participants are allowed to perform (slightly tight) gymnastics and exercises to realize the change in heart sounds due to physical activity. Through the work, some of the participants gradually feel the vibrotactile presentation device as “the alter ego of their heart” (based on the results of a questionnaire survey[9]).

The last point of this workshop is "instructing to switch off the device." By this instruction, the participants are trying to “stop” the device. Since they have accustomed to their own hands so that this work aims to “recognize life and self.

3. Tactile Virtual Reality game, Hertoid

The purpose of this research is to give a sense of presence to a virtual object in the VR space by feeding back the

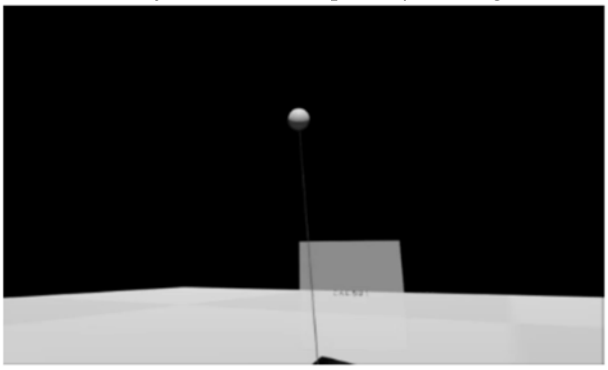


Fig. 1. From the game screen. The user and the heart (sphere) are connected by wires. The user feels the heartbeat from the tactile feedback of heart sounds.

baptized biological information based on the idea of the "heart picnic." Therefore, we created a virtual object in the VR space that made the user's biometric information tangible and constructed a tactile VR game system, Heartoid, using the virtual object.

This system is a VR game that touches a virtual beating object (sphere) that makes the player's heart sound tangible. This sphere corresponds to a vibratory tactile sensation presentation device in a hearty picnic, and by touching this sphere existing in the VR space via the controller, the player can touch the tunable heart sound. The tactile presentation device used was a tactile presentation device built into the controller of HINE's VINE VR. The sphere regularly vibrates finely based on the results of real-time measurement and analysis of the heartbeat sound and the movement (acceleration) of the player, and the amplitude of the vibration synchronized with the magnitude of the heart sound — the subtle vibration selected according to the player's biological condition and movement.

The player's body and the sphere are connected by a thin thread (Fig. 1). The thread visually complements the tunable heart sound, suggesting that the player's heart sound sent to the sphere, and suggesting a relationship between the player and the area with the thread. While the string is connected, the area floats in the air against gravity and is alive.

3.1. *Concept of the Game*

The concept of this system is the same as a hearty picnic, where you feel your own "life" by touching your artificial heart movements in a virtual space. Through interaction, we hope to deepen the awareness of other "life" by feeling the change of the vibration type generated by itself and feeling the existence of other beings but never other people's affairs. Therefore, it designed to be able to obtain various interpretations through various interactions with each player sphere.

3.2. *Structure of the Game*

1. Introduction (about 10 seconds)
The explanation about a sphere is given. Display "Android mimics the movement of your heart."
2. Contact 1 (about 1 minute)
The player takes no action and touches the floating sphere.
3. Introduction to gripping motion (about 30 seconds)
The operation of grasping and moving a sphere will be described. Here, the user recognizes the existence of minute vibration and its change.
4. Contact 2 (about 1 minute)
The player touches the sphere and feels the heartbeat of his heart through the vibration of the sphere.
5. Introduction to the in-game game (about 1 minute)
The explanation about the target game is given. During the game, the player interacts by grabbing and throwing a sphere. Also, if you change the volume (volume) of your voice while holding the sphere, you can inflate the sphere in proportion to the volume of voice.
6. Target game (about 3 minutes)

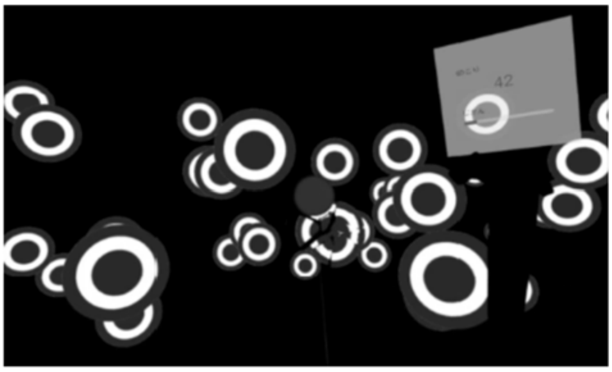


Fig. 2. Snapshot of the game. The black circle in the center is the sphere corresponding to the heart. The ring around it is the target. The user throws a sphere to break the target. Fifty targets generated.

The goal is to hit the sphere and break 50 targets in time. The heart rate fluctuates during the game. Also, the larger the sphere, the easier it is to hit multiple targets at the same time. Therefore, it is necessary to make physical movements and make utterances. Therefore, vocalization and respiratory heart rate fluctuate. The change of the mental and physical state is presented by tactile sensation and made tangible.

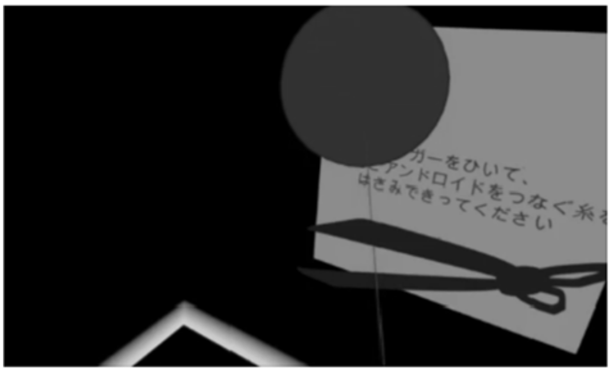


Fig. 2. Ending of the game. When the target application game ends, the controller becomes scissors. The user cuts the sphere connected to himself with the scissors using the scissors. When the thread reduced, the sphere falls by gravity. The sphere does not beat and does not return to me when thrown.

7. Ending

The player cuts the thread connecting the sphere to himself. At the end of the game, the controller becomes scissors. Use the scissors to cut the thread. When the thread is reduced, the sphere drops according to gravity and stops vibrating. The ball can be grabbed and thrown but does not respond in response to the user's heart rate. Also, throwing a sphere will not return to the player's hand.

4. Conclusion

In this research, we demonstrated the possibility of developing a haptic content game that affects sensibility by measuring and analyzing biological information and presenting haptics in real-time (making haptic). A user evaluation test will be performed in the future. Still, a user who tried for the preliminary survey said, “The sphere (corresponding to the heart) is alive because the changes due to physical movement are fed back by tactile feedback. Some comments have been foreseen that it affects sensibility, such as "I get a feeling." And "I hesitate to cut the wire at the end."

As future tasks, in addition to user evaluation, improvement of the execution speed of biological information analysis in this system will be significant for broader application. At present, it is not possible to support game designs that require review at every game frame (less than 0.0032 seconds for 30 FPS). Therefore, by speeding up the processing, it is desired to improve the method so that the technique can be applied to a game application requiring high-speed processing such as a shooting game.

References

1. First year of VR-Providing a place for hands-on experience Information service companies, creating new value, Nikkan Kogyo Shimbun August 15, 2016 (10 pages), 2016
2. Nintendo raises net income to 85 billion yen this quarter “Switch” is strong, Nikkei QUIC News, Nihon Keizai Shimbun October 30, 2017, 2017
3. Walk through the city "and get rid of zombies VR game facility in Abeno, Nihon Keizai Shimbun, August 4, 2017
4. H2L,Inc：UnlimitedHand, http://unlimitedhand.com/
5. Cerevo：Taclim, https://taclim.cerevo.com/ja/
6. VRgluv ： Force Feedback Gloves for VR, https://vrgluv.com/
7. Sunshine 60 Observatory: VR in the Sky (Virtual Reality), , http://www.sunshinecity.co.jp/campaign/cp/vr/
8. VR ZONE SHINJUKU: Ski downhill ski rodeo,VR ZONE SHINJUKU：https://vrzone-pic.com/activity/skirodeo.html
9. J. Watanabe, Y. Kawaguchi, A. Sakakura, H. Ando, “Heart Picnic”: Workshop on Touching the Beat, Transactions of the Virtual Reality Society of Japan, Vol.16. No.3, 2011 [10]
10. Y. Kono, M. Ishii, Y. Koike, M.Sato, Proposal of Tactileization System for the Visually Impaired, ITE Technical Report / 23.31 (1999)

Method of Extracting Sensibility from Time Series Data and Converting it to Vibrotactile

Yasuhiro Suzuki
Department Complex Systems Science, Graduate School of Informatics, Nagoya University
Furocho, Chikusa, Ngoya City, 464-8601, Japan

E-mail: ysuzuki@nagoya-u.jp,
www.ysuzuki.info

Abstract

We propose a method to convert sensory information into non-contact tactile sensation by vibrating of very low frequency sound. The vibrator does not directly touch the object in a non-contact tactile sensation. Then, the object is vibrated by giving an extremely low frequency. Previous research has shown that the resonance frequency of the chest and abdomen is about 10 Hz. Therefore, the chest and abdomen can be vibrated in a non-contact manner by using an extremely low frequency of about 10 Hz. Until now, very low frequencies have hardly used in engineering. We propose new uses for embedding very low frequencies in sound, we name it the Deep Micro Vibrotactile, DMV.

Keywords: Tactile Score, non-contact tactile, low frequency sounds, Deep Micro Vibrotactile, DMV

1. Contact Tactile Sense

Haptic Technology offers products such as haptics, tactile engineering, video game controllers, entertainment in amusement parks and cinemas, and direct skin and body contact.

Aesthetics and philosophy, on the other hand, discuss speculative haptics, such as visual haptics. This whitepaper does not focus on caring touches. Humankind has forgotten to consider the concept of contact time. The idea of time in the sense of touch here is a sense of time that changes with time. Music is a temporary change in auditory stimuli caused by sound. If a sound has no concept of time change, the sound will be a single sound. Music that has only one sound is not commonly called music. Some composers, such as Jacinto Shelsi, use long tones extensively but still use multiple long tones, and the sound changes over time.

In general, haptics does not change over time. Haptic engineering and tactile material research deal with soft, stiff haptic presentations that do not change over time. For example, no material changes the feel when touched. The only tactile that changes over time are a massage. Massage is not a single tactile stimulus, but multiple tactile stimuli that change over time. There are various theories about the origin of massage. One of the oldest records of massage is a wall painted on the tomb of an Egyptian doctor in BC.

Compared with tactile sensation, the hearing has evolved, and the concept of time change has become unmatched. Sheet music has been used since BC to describe music that is time-varying. The current sheet music was invented in the Middle Ages by an Italian monk and music teacher Guido D'Alezzo. The massage of the time cannot be reproduced from the mural of the doctor's grave in Egypt, but the medieval music can be replicated by playing the music score of the time.

Tactile sense directly makes an appeal to the deep feeling without language or words1. Tactile sense has a language, however, the way of using the language is obscure. If we could have full command of the language of tactile sense and express the feeling and tell it by using the word, cold

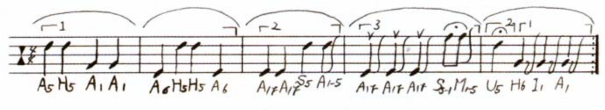


Fig. 1. A Tactile Score, where is generally described using staff notation. Special symbols are used as needed. This score is for massage, and each note has a logo. At the top of the staff notation, a key symbol and a number are attached, which means that the movement of the right hand and the left side is shifted, 1 means simultaneous movement, two means slight movement, and three means massive action.

media would be changing into intimate, organic ones.

1.1. *Tactile Score*

Humankind has forgotten to describe the time-varying haptics. Therefore, we have not developed a method to describe time-varying haptics. Because there is no description method, humans could not express and pass on the massage. Therefore, we proposed the Tactile Score[1], a method for describing time-varying haptics (Fig.1.).

There are various elements in the tactile sense, but it is difficult to describe them all. Music has multiple factors, but the score specifies only the pitch and duration. Tactile Score represents only the strength and length of pressure (reasonable force), ignoring other aspects. Tactile Score generally uses a staff notation. The center of the staff and the third line are normal strength; ascending is weak, descending is intense pressure. The length is described using musical notes. It is not necessary to discriminate the maximum durability and the minimum weight into five or seven steps. However, it is not essential to make the discretization in five or seven steps to use the staff notation. The reason for using staff notation is that it is the most popular.

1.2. *Tactile Score Vibration, TS vibration*

Tactile Score can describe the time-varying tactile sensation. Arbitrary messages can be represented, and it is possible to investigate how to construct massages that were previously unknown. From the research, the principle of the construction method of massage was obtained. Tactile Score has advanced the study of massage dramatically but faced great difficulties. Since the massage is performed using human hands, one person can massage only one person, so it is impossible to massage several people at the same time as music. Besides, to conduct experiments, humans must perform the massage; hence, it is impossible to generate the same tactile stimulus repeatedly, no matter how proper the technique is.

Tactile Score describes the change in pressure intensity and length over time. By changing the amplitude of vibrotactile and its time by So Tactile Score, the Tactile Score can convert to vibrotactile sensation. Therefore, we use a vibrotactile impression, which has little change in pressure and length over time so that the waveform of the vibration can design touch. We call such

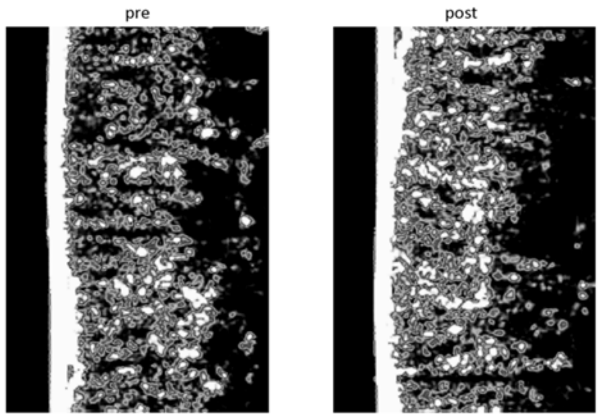


Fig. 2. Using the proprietary tactile vibration display HVD, collagen aggregation was observed when Tactile Score Vibration and TSV were subjected to facial vibration massage. There are several types of HVD. In the HVD used in this preliminary experiment, two sound exciters were used to shake the face up and down. In the figure, the left (pre) is before giving TSV and the right (post) is after giving TSV. On the right, the number of black spots is reduced because collagen was aggregated.

vibrotactile as Tactile Score Vibration, TSV. TSV can digitize the tactile sensation via Tactile Score. TSV has applied to many fields such as beauty and healthcare, and social implementation is progressing (Fig.2.).

2. Non-Contact Tactile Sense

Non-contact tactile sensation is to vibrate an object without touching it. Non-contact tactile sensation uses very low frequency. Human audible frequencies are above 20Hz and below are non-audible frequencies. The resonance frequency of the human chest and belly is about 10Hz. Therefore, the human breast and abdomen can be vibrated by a non-audible ultra-low frequency of about 10Hz without touching.

The very low frequency has not received much attention except for noise engineering. In noise engineering, very low-frequency sound, which is non-contact tactile sense, insulation has mainly performed. Also, few people complain of poor health, primarily concerning the very low frequencies generated by wind turbines.

The Japanese Ministry of the Environment surveyed the sound generated from wind power generation facilities. And they found that the very low-frequency sound produced from wind power generation facilities did not have a very high sound pressure level and was below the human perception threshold. Comparing other environmental noises, the noise generated by the wind farm was not necessarily prominent in the low-frequency range[2].

Besides, past studies on wind turbine noise and human health effects in Japan and abroad have broadly arranged. Review papers have reviewed by experts, published in medical journals, etc., and reports by governments of various countries have compiled. They could not confirm any clear link between the extremely low-frequency sound and low-frequency sound generated from the wind power generation facility and the health effects.

2.1. *40Hz vibrotactile suppresses Alzheimer disease*

Until now, low frequencies have received little attention except for noise engineering, as described above. But in April 2019, a neuroscientist at the Massachusetts Institute of Technology in the United States found that hearing 40 Hz to Alzheimer rats reduced beta-amyloid accumulation. In other words, they showed that listening to 40 Hz suppressed the progress of Alzheimer's disease. They surveyed several frequencies and found that 40Hz was the most effective. There is no effective treatment for Alzheimer's disease, and no new drugs have successfully developed. Although 40Hz is not an inaudible sound, it is a sound close to a vibration whose pitch is almost unrecognizable.

3. Tactile Sense Embedded Sound

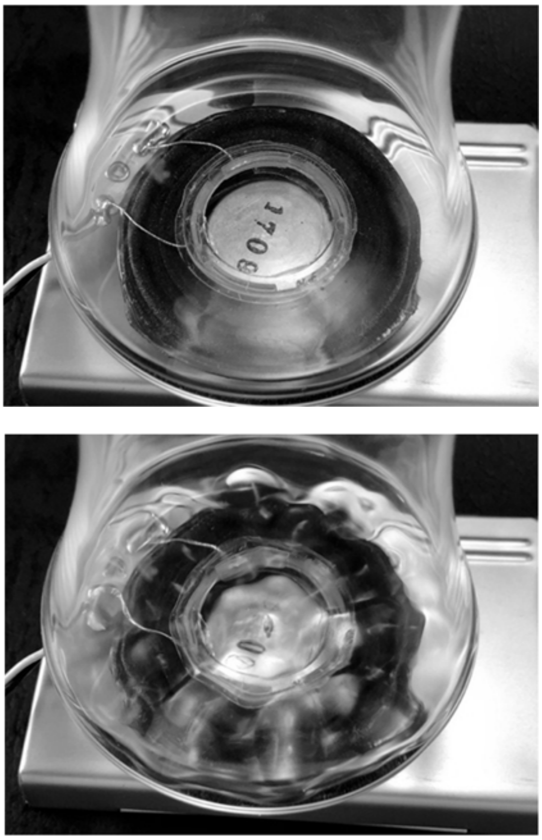


Fig. 2. Top) When only music from the sound exciter is applied to the container filled with water. Bottom) When sound mixed with music and vibration (non-audible low frequency) is used. The vibration component is inaudible low frequency and cannot be heard. However, when exposed to water, the water vibrates due to the vibration components contained in the sound, causing ripples.

In music, it distorts the sound and removes inaudible low-frequency or vibrational components. On the other hand, Goro Noguchi, a national singer, focuses on the vibration component, which is a very low frequency. Noguchi has discovered from years of experience that recognition of vibrations (very low frequencies) can deliver singing voices far.

However, as mentioned above, music and vibration cannot coexist. Noguchi has succeeded in coexisting music and vibration using acoustic technology that controls vibration without affecting music. Through joint research with Noguchi, we are developing technology to incorporate vibrotactile sense into music (Fig2.).

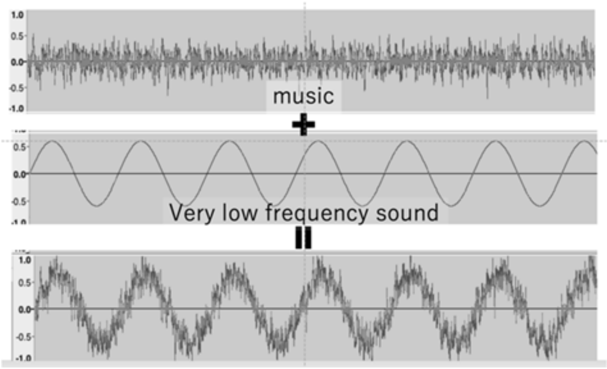


Fig. 3. When music and low frequencies are combined, the music is low frequency modulated.

3.1. *Low Frequency Modulation of Music*

By the combination of the very low-frequency sound, music is modulated by it (Fig.3.). If the frequency of the sound is high, the sound will be diffracted when it hits an obstacle.

On the other hand, if the frequency is low, it moves without diffraction even if there is an obstacle. Therefore, the range of low-frequency sounds is longer.

The minimum frequency generated by the minimum pitch of the pipe organ is 16Hz[3], which is non-audible sounds. For example, in Buch's "Toccata and fugues" this lowest note is used. Even if you can hear the music, the pitch will not be recognized as the frequency goes down. The reason for using non-audible sounds at high cost is to deliver audible sounds far away.

In the Edo period in Japan, the time of noon was signaled with a cannon every day at Edo Castle[4](now the Imperial Palace). The sound of this cannon all reached above Edo (now downtown Tokyo). This is because the sound of the cannon was transmitted farther by non-audible sounds.

When the low-frequency modulated music wave reaches the body, the audible component shakes the eardrum, and the low-frequency component shakes the body. We anticipate that the excitement of music is affected by low frequencies. We name the low frequency sound as Deep Micro Vibrotactile, DMV.

4. Conclusion

Haptic engineering has been dealing with tactile sensations delivered by contact. Acoustics has studied audible sounds. Therefore, non-contact tactile and non-audible sounds have not treated. This research proposes a method to evoke tactile sensation by vibrating by delivering low frequency to the body and suggests a new method that fuses non-contact tactile and non-audible sound.

It has confirmed that low frequency inhibits the progression of the pathology of Alzheimer rats and biological responses such as collagen aggregation. Such low frequencies can be embedded in music by Deep Micro Vibrotactile, DMV. In the future, DMVs will expand across fields, such as the beauty, medical, healthcare, and entertainment industries.

Acknowledgements

The authors would like to thank Dr. Mayumi Kotani and Akihito Fujita, president of Bloom Classic Inc. for their experiments on collagen aggregation for TS Vibration. We would also like to thank Mr. Goro Noguchi for his various cooperation in the collaboration of Deep Micro Vibrotactile, DMV.

References

1. Y.Suzuki and R.Suzuki (2013) *Tactile Score*, Springer Verlag
2. http://www.env.go.jp/air/teishuha/qa/#10
3. Low Frequency Sound, The Acoustical Society of Japan (2017), Colona Inc. (in Japanese)
4. M. Yamashita, Nazotoki Onkyo-Gaku in Japanese (Solve mystery in Acoustics) (2011), Maruzen Publishing

Proposal teaching materials the concepts and principles of machine learning for use in education

Shinichi Imai
Graduate school, Tokyo Gakugei University,
4-1-1, Nukuikita-machi, Koganei, Tokyo, 184-8501, Japan.

Yusuke Shiba
Faculty of Education, Tokyo Gakugei University,
4-1-1, Nukuikita-machi, Koganei, Tokyo, 184-8501, Japan.

E-mail: shimai@u-gakugei.ac.jp
http://www.u-gakugei.ac.jp/

Abstract

In this paper, Technologies such as AI, IoT, and big data are very important contents. However, IoT related content is still in the developing stage and is now evolving at a tremendous speed. It is very important to learn about AI, IoT, big data, etc. Therefore, we propose a teaching tool based on soccer penalty shoots using machine learning. This teaching tool can visually confirm that the goalkeeper robot does not shoot the ball by machine learning.

Keywords: Machine learning, Education, Teaching materials, Evaluation

.

1. Introduction

In recent years, as represented by IoT (Internet of Things) [1,2] and AI (Artificial Intelligence) [3,4], major evolution has often occurred rapidly in the ICT world. It is very important to learn about AI, IoT, big data, etc.
Under such circumstances, it is considered necessary for students of this university to learn the current AI technology level and what it can be used for, not only by classroom learning, but also by actually touching AI. The purpose of this study is to develop a robot that can learn while experiencing the concepts and principles of AI. There is no specific definition of AI at this time, and there are various interpretations by researchers. Therefore, in this study, we develop a robot teaching material that can learn the concept and principle of “machine learning” proposed in the early days of AI research while operating the robot.

As related research [5], network education, data analysis education, and problem solving that promote understanding from the viewpoint of data using an "IoT learning environment" constructed by cooperation of sensors, microcomputers, programming, and data storage servers in high school Methods of learning and programming education have been proposed. In this report, we were able to observe the improvement of programming ability through game production as an evaluation result, but there is a problem of a separate learning program in creative activities that understand and understand the data flow and network mechanism. In experiments using the device, the results calculated based on the theory can be visually observed as movements. Furthermore, you can feel that you are operating the device by listening to the operation sound of the device and touching switches and buttons. In other words, you can have an “experience” that you can see with your eyes, listen with your ears, and touch with your hands. References [6] and [7] insist on the need for experience and teaching material development based on cognitive science.

2. Machine learning teaching material

The learning material to be developed is a learning material that can visually understand machine learning based on the soccer penalty kick shown in Fig. 1. The penalty kick learning material is a mechanism that sensor and the Raspberry Pi. There are a switch for switching between a "learning mode" for performing learning and an "execution mode" for actually moving, and two switches for starting. In addition, three LEDs are installed for operation confirmation. Details of the hardware are shown in Table 1. Various functions will be described.

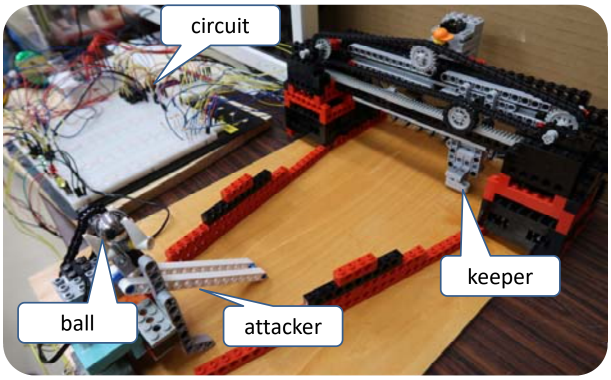


Fig. 1. Teaching materials

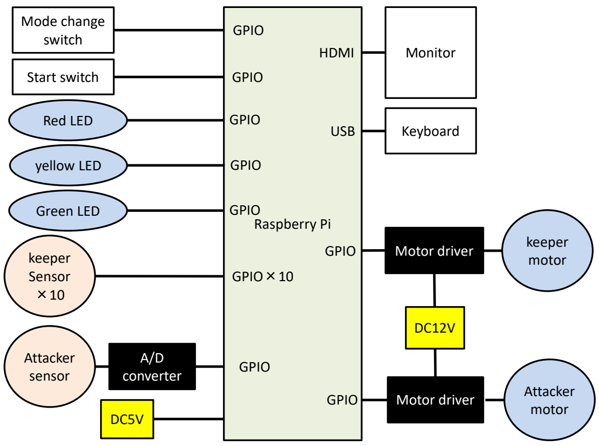


Fig. 2. System configuration

allows the goalkeeper to learn the position of the kicker's ball.
The development of teaching materials uses Lego Mindstorm[8], which is often used in robot learning

Table1 Teaching material specifications

Hardware	Raspberry Pi 3 Model B
Software	Python 3
Library	･RPi.GPIO ･Scikit-learn
A/D converter	mcp3008(10bit 8ch)
Sensor	Potentiometer(RV30YN40R) Microswitch(JF11210)×10
Actuator	LEGO Mind storm RCX motor×2
Other	Motor driver(TA7291P)

education. As a result, teaching materials can be created by combining blocks without specialized skills.

2.1. *Hardware*

As shown in Fig. 2, the hardware system configuration obtains the voltage of ten touch sensors that detect the position of the ball and the potentiometer that detects the position of the kicker, and commands the given control amount to the keeper motor. The microcomputer uses Raspberry Pi, and Python, which is often used for machine learning and AI, is used for programming. Since the Raspberry Pi does not have A/D conversion, analog sensors cannot be imported directly. For this reason, an A/D converter is installed between the analog sensor and the Raspberry Pi. There are a switch for switching between a "learning mode" for performing learning and an "execution mode" for actually moving, and two switches for starting. In addition, three LEDs are installed for operation confirmation. Details of the hardware are shown in Table 1. Various functions will be described.

●Keeper
The keeper moves to the left and right to catch the ball
● Goal
Collect balls that were not received by the keeper
● Attacker
Shoot the ball
Measure the angle of the attacker with a potentiometer

● Switch
Switch to switch between learning mode and execution mode
Switch to shoot the ball

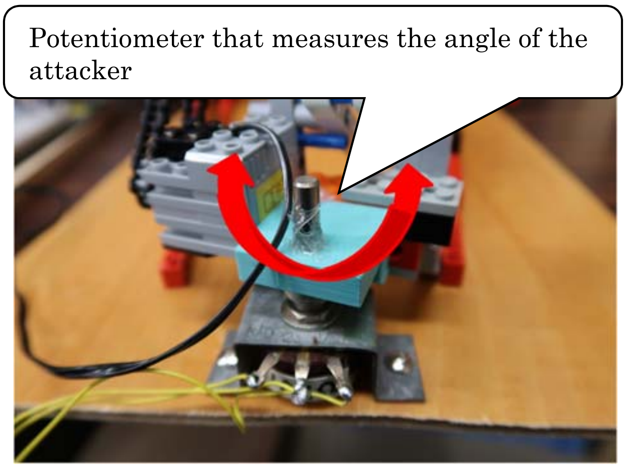


Fig. 3. Measuring the angle of the attacker

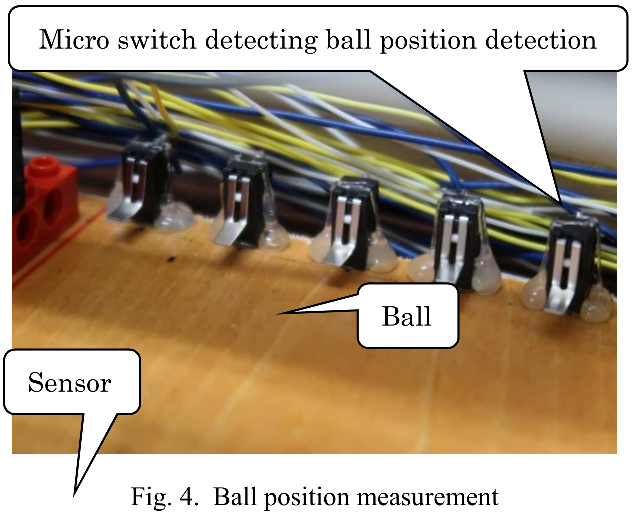


Fig. 4. Ball position measurement

● Check LED
Red: Loading
Yellow: Learning mode
Green: Run mode

2.2. *Linear regression model*

A linear regression model is used for machine learning. Here, learning is performed from the launch angle and the goal position. In this teaching material, the goal position is then estimated from the firing angle. The learned data is saved in csv format on the SD memory of Raspberry Pi.
The linear regression model models the relationship between the objective variable Y and the explanatory variablesX_i, i = 1,2 ..., p as follows.

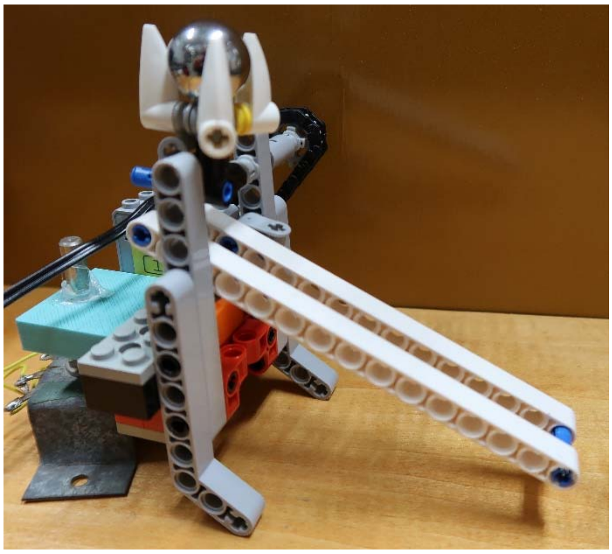


Fig. 5. Attacker

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p \qquad (1)$$

where, β_0 is the intercept, β_i is the coefficient of each explanatory variable, and p is the number of explanatory variables. In linear regression, a model with parameters of explanatory variable coefficients and intercept pairs $\{\beta_i\}_{i\in[0,p)}$is given.
Here, the objective variable represents the position of the keeper, and the explanatory variable is the position of the limit switch and the angle of the attacker.

3. Operation experiment

Explain the operation of the developed teaching materials. The operation has a learning mode and an execution mode.

3.1. *Learning mode*

1. Press the fire switch to collect the angle of the attacker shown in Fig. 3.
2. The ball contacts the micro switch and collects the goal position shown in Fig. 4.
3. Write each collected data to csv file.
4. Repeat steps 1 to 3 several tens of times.
5. Estimate the goal position from the angle.

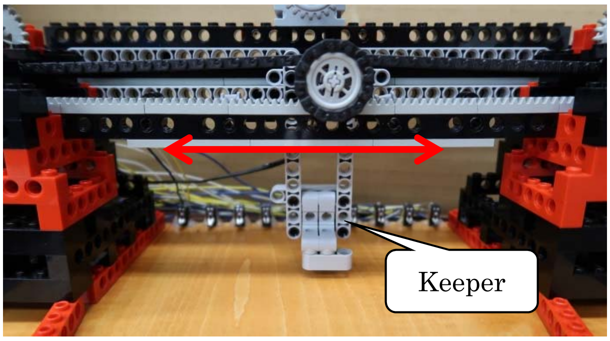


Fig. 6. Keeper

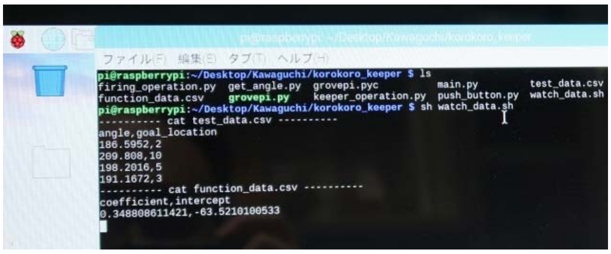


Fig. 7. Learning result

3.2. *Run mode*

1. Successful learning if the attacker (Fig.5) launches the ball and the keeper accepts shown in Fig. 6. (Successful learning if not accepted)
2. If not, return to learning mode and repeat learning.

3.3. *Operation check*

Since the developed robot is not learning anything at first, the goalkeeper robot cannot catch the ball, but by accumulating successful and unsuccessful experiences, inappropriate motion gradually decreases. Eventually, it became possible to catch the kicker's ball. Fig.7 shows the results of the learning materials. Here, the linear regression model of the learning result is displayed at the bottom.

4. Conclusion

In this research, we developed a teaching material that can visually learn machine learning, a kind of AI. Anyone can easily make the developed teaching materials by using a LEGO mindstorm. In addition, by using a Raspberry Pi in a computer, it has become possible to program in Python. As an operation check, the goal accuracy could be improved according to the number of learnings. In the future, I would like to support reinforcement learning and deep learning other than linear regression models. In addition, the class will be practiced using this teaching material.

References

1. K. H. Yeh, "A secure IoT-based healthcare system with body sensor networks," IEEE Access, Vol.4, (2016) pp. 10288-10299.
2. Kevin Ashton, "That `internet of things' thing", RFID Journal, Vol.22, No.7, (2009) pp.97-114.
3. B. Vallade, A. David, and T. Nakashima, "Three Layers Framework Concept for Adjustable Artificial Intelligence," J. Adv. Comput. Intell. Intell. Inform., Vol.19, No.6, (2015) pp. 867-879.
4. A. Nareyek, "AI in Computer Games," Queue, Vol.1, No.10, (2004) pp. 58-65.
5. H.Manabe, M.Omura, K.Hayashi, S.Kanemune, "How can we use 'IoT Learning System' for the information education?", 2016 Information Processing Society of Japan, (2016) pp.98-105. (in Japanese)
6. T.Asai, K.Osuka, M.Ishikawa and M.Inoue "A Motivational Lecture Using an Experimental Control System to Get Personal Experience and Its Effectiveness", Transactions of the Society of Instrument and Control Engineers Vol.48, No,10, (2012) pp622-631. (in Japanese)
7. R.Yasui, T.Sato, N.Araki and Y.Konishi "Control education by experiment using a floating device", The transactions of the Institute of Electrical Engineers of Japan. C, A publication of Electronics, Vol.138, No.5, (2018) pp.471-479. (in Japanese)
8. Official Lego Mindstorms, https://www.lego.com/en-us/themes/mindstorms?domainredir=mindstorms.lego.com (accessed 10 Dec, 2019)

Objective Evaluation of the Educational Effects on the Feedforward, Feedback and PID Control[*]

Yugo Tokura, Takao Sato, Ryota Yasui, Natsuki Kawaguchi, Nozomu Araki and Yasuo Konishi
Graduate School of Engineering, University of Hyogo
2167, Shosha, Himeji, Hyogo 671-2280, Japan
E-mail: {tsato, kawaguchi, araki, konishi}@eng.u-hyogo.ac.jp

Abstract

The aim of the present study is to improve the understanding of the control engineering for university students through the experience of the control experiment, and hence an experimental device is developed. The students are educated using the device, and the educational effect is evaluated. In the present study, the students take examinations for the control engineering before and after the experiment, and the effect is evaluated objectively.

Keywords: Control education, evaluation, feedforward, feedback, proportional, gain.

1. Introduction

In the engineering education, the control engineering is educated for the students who have taken mathematics, applied physics, mechanical engineering, electrical engineering and so on. The control engineering is not easy for students to understand intuitively only by the classroom lecture. Therefore, effective educational methods are demanded[1]. To realize the class of the control engineering through an experience of the control experiment, experimental devices have been developed[2-4]. In the conventional studies[3,4], the educational effect is evaluated subjectively. On the other hand, in the present study, questionnaire is answered before and after the experiment, and the difference between the questionnaires. As a result, the educational effect is evaluated objectively.

2. Educational objective

The educational objective for developing an experimental device is that engineering students understand the following control actions:

- Manual control
- Automatic control
- Feedback control
- Feedforward control
- Proportional control

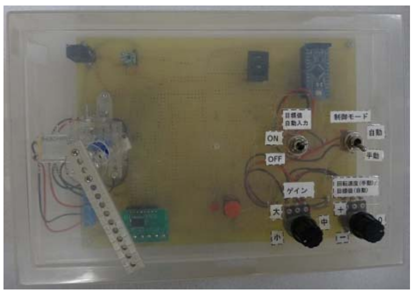


Fig. 1. Arm control device.

.

3. Experimental device

The developed experimental device is shown in Fig. 1. In the device, and its schematic diagram is illustrated as Fig. 2. In the device, the arm is rotated by a motor, and the control objective is to make the rotation angle of the arm be settled at the specified angle. In the manual control, the torque generated by the motor is changed by tuning a volume manually. In the automatic control, the rotation angle of the arm is measured, and the torque is decided by a computer using the measured angle.

Using the device, the students experience the control actions and can understand these meanings intuitively. The control actions implemented in the device are listed as follows:

- Manual control
  - Feedback control
  - Feedforward control
- Automatic feedback control
  - Small proportional gain
  - Middle proportional gain
  - Large proportional gain

4. Evaluation

To objectively evaluate whether the educational objective is achieved, the students answer the questionnaire shown in Table 1 before and after the experiment. The questionnaire consists of 19 questions, the correct answer rates are depicted in Fig. 3. In the figure, the blue bar shows the correct answer rate before the experiment, and the red bar shows after the

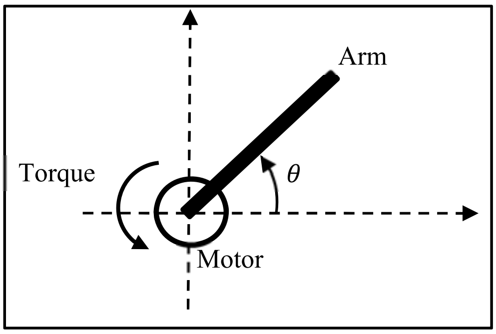


Fig. 2. Schematic diagram of the experimental device.

Table 1. Questionnaire before and after experiment

Term in control system	Question	Number in Figure 3
Manual control	Manipulated amount is decided by human	1
	Manipulated amount is decided by other than human	2
Automatic control	Manipulated amount is decided by human	3
	Manipulated amount is decided by other than human	4
Manual control vs automatic control	Manual control structure is simpler than automatic control	5
	automatic control structure is simpler than manual control	6
Feedback control	Stability can be improved	7
	Stability can be worsened	8
	Reference tracking performance can be improved	9
	Reference tracking performance can be worsened	10
Feedforward control	Stability can be improved	11
	Stability can be worsened	12
	Reference tracking performance can be improved	13
	Reference tracking performance can be worsened	14
Feedback control vs feedforward control	Feedforward control structure is simpler than feedback control	15
	Feedback control structure is simpler than feedforward control	16
Proportional control	Response speed can be improved	17
	Response speed can be worsened	18
	No steady-state error	19

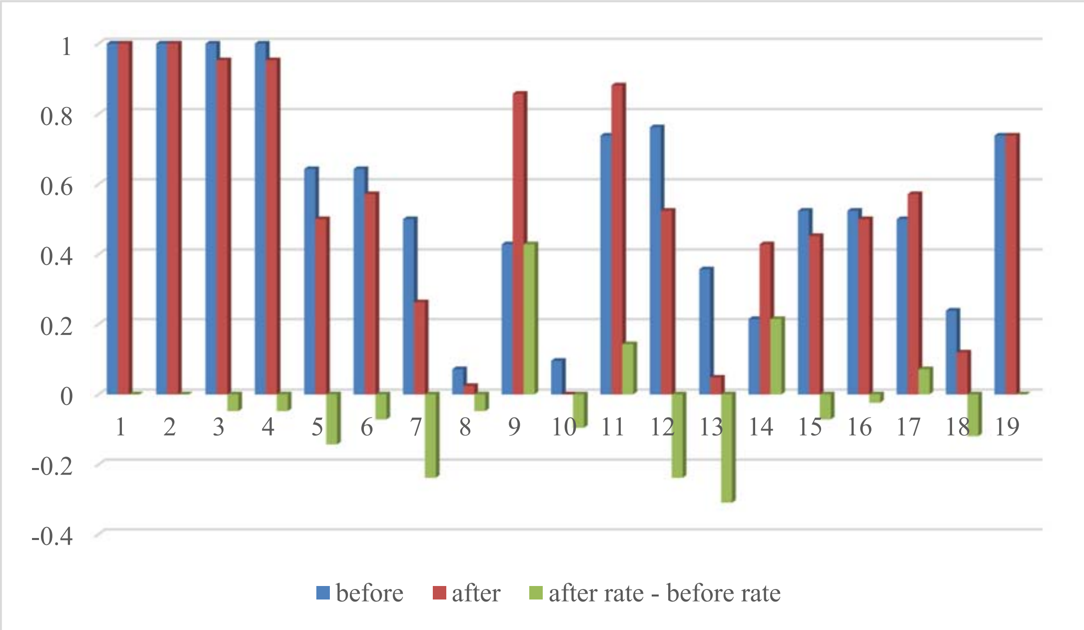


Figure 3. Correct answer rate before and after experiment

experiment. Further, the green bar shows the difference between the before and after experiments.
Fig. 3 shows that most of the students understand the mechanisms of the manual control and the automatic control and understand little about that the stability and the tracking performance can be worsened by the feedback control. Since the rates of Questions 9 and 14 of after experiment are increased compared with those before experiment, the number of the students who understand that the tracking performance can be improved and worsened by the feedback control and the feedforward control, respectively, is increased by experiencing the control experiment. On the other hand, from Questions 7, 12, and 13, the number of the students who misunderstand that the stability and tracking performance can be changed by the feedback control and the feedforward control.

5. Conclusion

In the present study, the educational effect using the experiment is evaluated objectively. A future work is to enhance the control engineering lesson to increase the correct answer rate after experiment.

Acknowledgment

This study was supported by JSPS KAKENHI Grant Number 16K06425.

References

1. H. Bonicke and C. Ament, Increased Usability through User Interface Modification of Development and Education Tools for Embedded Systems, *Proceedings of 2012 IEE Multi-Conference on Systems and Control*, Dubrovnik, 2012, pp. 675-680.
2. T. Sato, N. Kawaguchi, S. Nakatani, N. Araki and Y. Konishi, Control Education Through Flight Experiment and its Efficacy, *Proceedings of 11th IFAC Symposium on Advances in Control Education (ACE 2016)*, Bratislava, Slovakia, 2016, WeParallel A1.1.
3. R. Yasui, T. Sato, N. Araki and Y. Konishi, *Control Education by Experiment using a Floating Device*, IEEJ Transactions on Electronics, Information and Systems, **138**(5), 2018, pp. 471-479 (in Japanese).
4. K. Kawada and T. Tamai, *Development of Teaching Support Material for Nurturing Cooperation through Play*, Journal of Robotics, Networking and Artificial Life, **5**(3), 2018, pp. 169-172.

Programming Learning of Temperature Control for Science Class of Elementary School *

Yoshihiro Ohnishi*
Faculty of Education, Ehime University, 3, Bunkyocho
Matsuyama, Ehime 790-8577, Japan
E-mail: ohnishi@ehime-u.ac.jp
www. ehime-u.ac.jp

Takeshi Nakano**
Ehime University Elementary School, 1-5-22, Mochidacho
Matsuyama, Ehime 790-0855, Japan
E-mail: nakano.takeshi.hu@ehime-u.ac.jp

Teruyuki Tamai*
E-mail: tamai.teruyuki.xq@ehime-u.ac.jp

Shinnosuke Mori*
E-mail: mori.shinnosuke.my@ehime-u.ac.jp

Kawada Kazuo***
Faculty of Education, Hiroshima University, 1-1-1, Kagamiyama
Higashihiroshima, Hiroshima 739-8524, Japan
E-mail: kawada@hiroshima-u.ac.jp

Abstract

In the new study guidelines promulgated in 2017, it has been decided that programming education will be required in elementary schools from 2020. In the new study guidelines, new programming subjects will not be established, but programming will be handled within existing subjects such as mathematics and science. It is desirable to conduct independent research by analyzing information using programming using subjects related to existing subjects. Based on the above, we developed a teaching material on temperature control programming that uses heat generated by heating wire in the unit of "Use of electricity" in the 6th grade elementary school. In this paper, after considering various indicators of control, teaching materials using temperature control are introduced.

Keywords: programming, teaching material, temperature control

1. Introduction

In the new study guidelines promulgated in 2017, it has been decided that programming education will be required in elementary schools from 2020. In the new study guidelines, new programming subjects will not be established, but programming will be handled within existing subjects such as mathematics and science. It is desirable to conduct independent research by analyzing information using programming using subjects related to existing subjects. Based on the above, we developed a teaching material on temperature control programming that uses heat generated by heating wire in the unit of "Use of electricity" in the 6th grade elementary school. In the new study guideline "I noticed that there are tools that use energy efficiently by using temperature sensors, etc. around me. It is possible to learn the mechanism through experience through programming such as controlling the lighting of LEDs and light emitting diodes. " From these words, after learning the heat generated by the heating wire, students should actively investigate what kind of operation should be performed to keep the temperature constant and whether it can be automated by programming. In this paper, after considering various indicators of control, teaching materials using temperature control are introduced.

2. Performance Index

Determination of whether the controller is performing the desired performance has conventionally been made by the subjectivity of a veteran operator who is familiar with the plant. The controlled value of a process control system represented by a petrochemical plant etc. is given a reference value of a constant value such as temperature, pressure, flow rate. The control performance required at this time often emphasizes reducing the variance of the control error in the steady state from the reference value tracking property and the disturbance response property[1]-[3]. In these methods, there are the method of using the minimum variance control as a benchmark [4], the method of considering not only the output error but also the fluctuation of the manipulated variable [5]. As described above, control performance assessment having a control value of a fixed value target value is often performed based on output variance. Furthermore, minimum-variance control performance assessment index(MV-index) is shown by the following equation:

$$\eta = \frac{\sigma_{MV}^2}{\sigma_e^2} \qquad (1)$$

Where, σ_e^2 is the variance of control error signal. And, σ_{MV}^2 means the achievable minimum variance of control error signal. This can be calculated by the actual output without system model. The procedure of this calculation is considered on [6] The range is 0≤η≤1. The large value of MV-index η means that good control performance. However, long data is required for calculation of (1). Furthermore, it is difficult for elementary school students to understand the concept of the variance. Therefore, the following new index is introduced.

$$J = \sum_{i=0}^{t} \{a - |r(i) - y(i)|\} \qquad (2)$$

Where, a is a parameter that determines the target zone. This index can quantify how close the controlled value is to the reference value[7].

3. Experimental system

When constructing a system of temperature control, the control method is important. The selection of the heater is also important. For example, in the case of a temperature control system using a heating wire, the resistance of the heating wire is important. The temperature control learning system proposed in this study can use three heating wires of different thickness. Of course, a thick heating wire has a smaller electric resistance value and a large current flow, so the amount of heat generation due to Joule heat increases. The temperature control learning system is shown in Fig. 1. Switching of a heating wire uses FET 2SK4017. A computer is required to supply the gate voltage to the FET. In this study, we use Studuino compatible Arduino for this purpose. Studuino has Analog input which can do 10bit A/D conversion. And DC motor driver is also equipped. The schematic figure of temperature control system is shown as Fig.2.

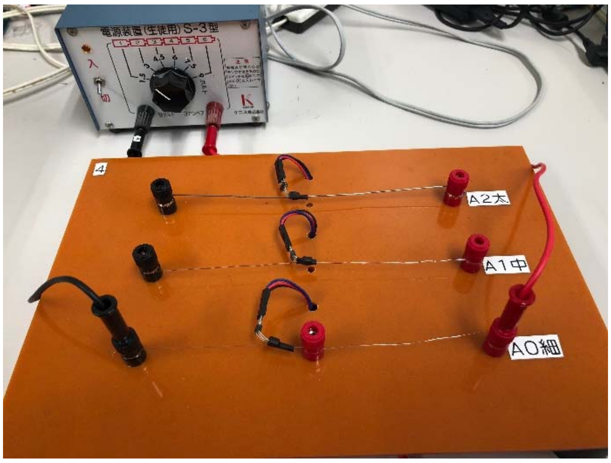


Fig.1 Temperature control system.

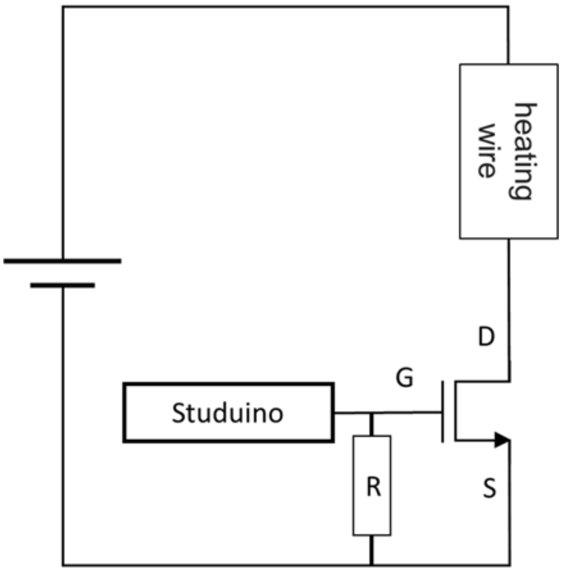


Fig.2 Schematic figure of temperature control system.

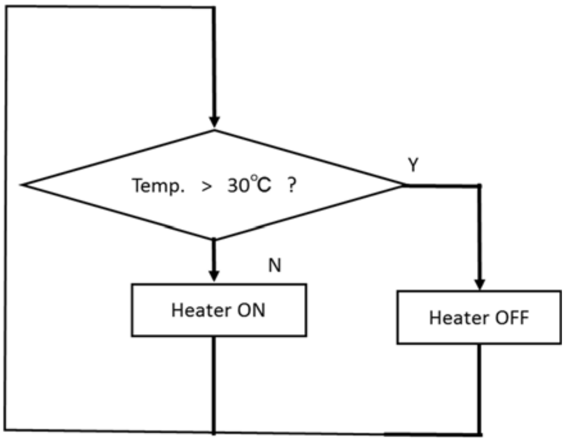


Fig.3 Flowchart of On-OFF control.

4. Experimental result

This session demonstrates the experimental result of the experimental system which is explained the previous session. The control algorithm used in this experiment uses ON-OFF control that a beginner makes first. Fig.3 is the flowchart of ON-OFF control. This section uses the same control law to confirm that the control performance differs depending on the difference in hardware. As hardware, a combination of thick heating wire and thin heating wire, high voltage and low voltage was prepared. First, when a low voltage is applied to a thin heating wire, the target temperature is not reached. The result is shown in Fig.4. In Fig4 "ten" is the calculation result of index(2). The "ten" in Fig.4 is lowest value in all results. Fig. 5 shows the result of applying a low voltage to a thick heating wire. The target temperature is reached, and the variance of the control error is also small. However, the rise time takes a lot. The "ten" in Fig.5 is better than Fig.4. Fig. 6 shows the result of applying a high voltage to a thin heating wire. Although the rise time is shorter than Fig. 5, the variance of the control error is slightly larger. Fig. 7 shows the result of applying a high voltage to a thick heating wire. Although the rise is very good, it cannot be stabilized around the reference value by the ON-OFF control. Finally, Fig.8 is shown as a reference result. Fig. 8 shows the result of PI control instead of the ON-OFF control shown in Fig.3. The "ten" of Fig.8 is the best of all results.

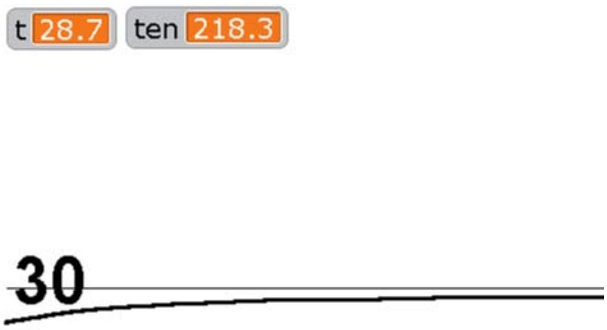


Fig. 4 Control result of applying low voltage to thin heating wire.

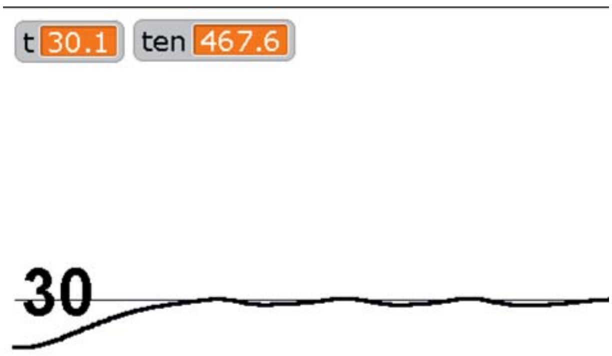
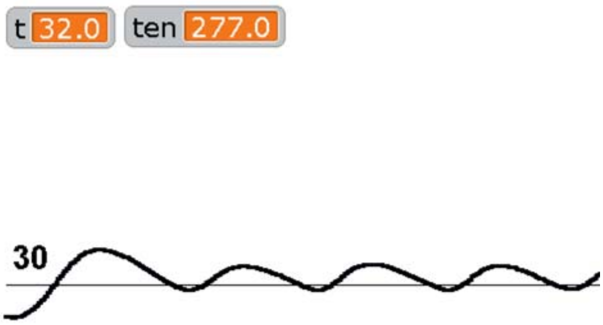
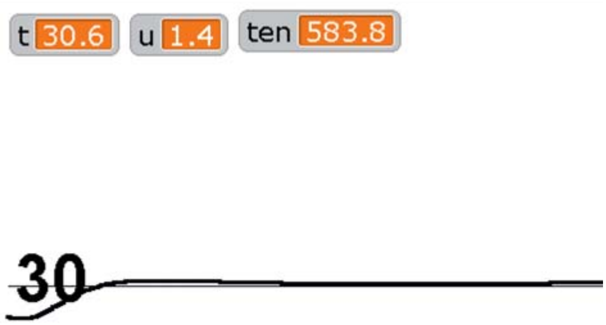


Fig. 5 Control result of applying low voltage to thick heating wire.

Fig. 6 Control result of applying high voltage to thin heating wire.



Fig, 7 Control result by applying high voltage to thick heating wire.



Fig, 8 Control result by applying high voltage to thick heating wire(PI control).

5. Conclusions

In this paper, temperature control experiments were explained as the content of sixth grade elementary school science. In order to check how close the temperature is to the reference value, an index related to dispersion was adopted. The effectiveness of the proposed method was evaluated by experimental examples.

Acknowledgements

The development of the experimental system was carried out by Mr. Motoki Sato a former student of Ehime university. This work was supported by JSPS KAKENHI Grant Number JP18K02980.

References

1. B. Huang and S. L. Shah: Performance Assessment of Control Loops: Theory and Applications, Springer, London (1999)
2. M. Jelali: An overview of control performance assessment technology and industrial applications, Control Engineering Practice, Vol.14, pp. 441-466 (2006)
3. M. Jelali: Control Performance Management in Industrial Automation, Springer, London (2013)
4. T. J. Harris: Assessment of Closed Loop Performance, Canadian Journal of Chemical Engineering, Vol.67, pp. 856-861 (1989)
5. M. J. Grimble: Controller performance benchmarking and tuning using generalized minimum variance control。Automatica, Vol. 38, pp.2111-2119 (2002)
6. L. Desborough and T. J. Harris: Performance assessment measures for univariate feedback control, The Canadian Journal of Chemical Engineering, Vol 70, pp.1186-1197 (1992)
7. Y.Ohnishi: Educational Effect Measurement of Control Technology based on Performance Index, Proceedings of the SICE Annual Conference 2019, pp. 1323-1327(2019)

Actuator Fault-Tolerant Control Using a Spiking Neuron Model

Masanori Takahashi
Department of Electrical Engineering & Computer Science, Tokai University,
9-1-1 Toroku Kumamoto, 862-8652, Japan
E-mail: masataka@ktmail.tokai-u.jp

Abstract

This paper presents a new design method for a self-repairing control system (SRCS) as an actuator fault-tolerant control system (FTCS). The proposed SRCS uses the well-known Izhikevich spiking neuron model as a fault detector. When the actuator fails, the neuron model is excited and then spikes occurs. Thus, counting up spikes makes it possible to find failures. Compared with the existing active FTCSs, the proposed method has the following advantages: (1) it is possible to set a maximum detection time in advance, and (2) the structure of the control system does not depend on the mathematical model of the plant. It is quite simple even if the plants have high orders. In this paper, several numerical simulation results are shown to confirm the effectiveness of the proposed SRCS.

Keywords: Self-repairing control, actuator failure, fault detection, dynamic redundancy, spiking neuron model.

1. Introduction

In the previous works[1,2], several types of the self-repairing control systems (SRCS) have been developed as one of active fault-tolerant control systems (FTCS). The SRCS can automatically detect the failure, and replace the failed sensor with the healthy backup so as to recover the system stability. Compared with existing active FTCSs, the SRCS has the following advantages: (1) the maximum time of detection can be specified arbitrarily in advance, and (2) the structure of the control system does not depend on the mathematical model of the plant. It becomes quite simple even if the plants have high orders. Unfortunately, the conventional SRCSs have utilized an unstable detection filter[1] which is not suitable for the concept of the strong stability[3]. Recently, as a remedy, the well-known Izhikevich spiking neuron model is used as fault detector. A faulty signal in the control loop excites the neuron model. Hence, just counting up the number of spiking waves makes it possible to find failures. Of course, the boundedness of all the signals in the neuron model is always guaranteed, and so the requirement of the strong stability can be satisfied. Moreover, it is not necessary to set a threshold for fault detection. However, only sensor failure has been considered, but actuator failure has not been solved.

In this paper, the SRCS using the spiking neuron model is modified to tolerate the actuator failure. Furthermore, the theoretical analysis on stability and the procedure for fault detection are shown. In addition, the effectiveness of the proposed SRCS is confirmed through numerical simulation.

Throughout this paper, with $x \in \mathbb{R}$, define the "sgn" function by

$$\mathrm{sgn}[x] = \begin{cases} 1 & (x \geq 0) \\ -1 & (x < 0) \end{cases}$$

2. Problem Statement

Consider a linear time invariant system of the form:

$$\Sigma_P : \begin{aligned} \dot{y} &= ay + bu + \boldsymbol{h}^T \boldsymbol{z} \\ \dot{\boldsymbol{z}} &= \boldsymbol{F}\boldsymbol{z} + \boldsymbol{g}y \end{aligned} \tag{1}$$

where, $y \in \mathbb{R}$ is the output, $u : \mathbb{R}^+ \to \mathbb{R}$ is the actual control input, and $\boldsymbol{z} \in \mathbb{R}^{n-1}$ is the state. Here, assume that the high frequency gain $b \in \mathbb{R}$ is positive. Moreover, $\boldsymbol{F} \in \mathbb{R}^{(n-1)\times(n-1)}$ is supposed to be a stable matrix (*i.e.*, all eigenvalues lie in $\mathbb{C}^-$).

For occasion of failure, the two actuators are prepared. One is the primary sensor #1, and the other is the backup #2. Then, the actual control input can be expressed as follows.

$$u(t) = \begin{cases} u_1(t) & (t \le t_D) \\ u_2(t) & (t > t_D) \end{cases} \tag{2}$$

where, $t_D \in \mathbb{R}^+$ is the detection time, and its detail will be discussed later. Each $u_i \in \mathbb{R}, i \in \{1, 2\}$ is the output of the actuator $\#i$. Obviously, in healthy case, we have $u_i = u_c$, where $u_c: \mathbb{R}^+ \to \mathbb{R}$ is the designed control input. Based on dynamic redundancy (2), the primary actuator #1 is usually utilized, but switched to the backup when the failure of the primary one is detected.

The failure scenario to be consider here, is expressed as follows.

$$u_1(t) = \varphi, t \ge t_F \tag{3}$$

where, $t_F \in \mathbb{R}^+$ is the unknown failure time, and $\varphi \in \mathbb{R}$ is the unknown stuck value. Such a failure occurs when the actuator gets stuck.

The problem is to design the SRCS, which can replace the failed sensor with the backup so as to maintain the stability and guarantee the convergence property of y:

$$\lim_{t\to\infty} \sup|y(t)| \le \lambda \tag{4}$$

for arbitrarily given $\lambda \in \mathbb{R}^+$.

3. Control System Design

First of all, the detection filter is introduced based on the spiking neuron model[4,5]

$$\Sigma_D : \dot{v} = \mathrm{sgn}[y](\theta_2 v^2 + \theta_0 - \eta w) + \theta_1 v + p(y + y^3)$$
$$\dot{w} = \varepsilon(\gamma_1 \mathrm{sgn}[y]v + \gamma_0 - w) \tag{5}$$
$$\text{if } \mathrm{sgn}[y]v \ge v_T \text{ then } \begin{cases} v \leftarrow \mathrm{sgn}[y]v_R \\ w \leftarrow w + w_R \end{cases} \tag{6}$$

where, $\theta_0 \in \mathbb{R}, \theta_1 \in \mathbb{R}, \theta_2 \in \mathbb{R}$, $\gamma_0 \in \mathbb{R}$, $\gamma_1 \in \mathbb{R}$ and $\eta \in \mathbb{R}^+$ are some constants. Also, $v_T \in \mathbb{R}$ is the threshold for the "auxiliary resetting" (for spiking), and the resetting rule (6) should be invalid before the steady state.

Next, the high-gain feedback controller is designed by

$$\Sigma_C : u_c = -\frac{p^2}{b}(y + v + y^3 + v^3) \tag{7}$$

where, $p \in \mathbb{R}^+$ is the feedback gain to stabilize both the plant and the detection filter. Then, the following lemma stands.

Lemma 1. *Consider the control system constructed by (1)-(7). If there is no failure, then all of the signals in the control system are bounded. Furthermore, regarding convergence of the plant output, the inequality (4) hold.*

Proof. Suppose that there is no failure, *i.e.*, $u = u_c$. Now, define the new variable: $s : \mathbb{R}^+ \to \mathbb{R}$ by

$$s := y + v \tag{8}$$

From (1), (5) and (8), it is shown that

$$\begin{aligned}\dot{s} = &-(p^2 - p - a)s - p^2 s^3 - (p - a)v - pv^3 \\ &+3(p^2 - p)s^2 v - 3(p^2 - p)sv^2 + \boldsymbol{h}^T\boldsymbol{z} \\ &+\mathrm{sgn}[y]\theta_2 v^2 + \theta_1 + \mathrm{sgn}[y]\theta_0 - \mathrm{sgn}[y]\eta w\end{aligned} \tag{9}$$

$$\dot{\boldsymbol{z}} = \boldsymbol{F}\boldsymbol{z} + \boldsymbol{g}s - \boldsymbol{g}v \tag{10}$$

$$\begin{aligned}\dot{v} = &-pv - pv^3 + \mathrm{sgn}[y]\theta_2 v^2 + \theta_1 v \\ &+\mathrm{sgn}[y]\theta_0 - \mathrm{sgn}[y]\eta w \\ &+ps + ps^3 - 3ps^2 v + 3psv^2\end{aligned} \tag{11}$$

Consider the positive definite function $V: \mathbb{R}^+ \to \mathbb{R}^+$ as,

$$V := \frac{1}{2}\{s^2 + \delta_1 \boldsymbol{z}^T\boldsymbol{P}\boldsymbol{z} + v^2 + \delta_2 w^2\} \tag{12}$$

where $\boldsymbol{P} \in \mathbb{R}^{(n-1)\times(n-1)}$ is the positive definite matrix which satisfies $\boldsymbol{F}^T\boldsymbol{P} + \boldsymbol{P}^T\boldsymbol{F} = -2\boldsymbol{Q}$ for any positive definite $\boldsymbol{Q} \in \mathbb{R}^{(n-1)\times(n-1)}$. Taking the time derivative of V give

$$\begin{aligned}\dot{V} = &-(p^2 - p - a)s^2 - p^2 s^4 - (p - a)sv - psv^3 \\ &+3(p^2 - p)s^3 v - 3(p^2 - p)s^2 v^2 + \boldsymbol{h}^T\boldsymbol{z}s \\ &+\mathrm{sgn}[y]\theta_2 sv^2 + \theta_1 s + \mathrm{sgn}[y]\theta_0 s - \mathrm{sgn}[y]\eta ws \\ &-\delta_1\boldsymbol{z}^T\boldsymbol{Q}\boldsymbol{z} + \delta_1\boldsymbol{z}^T\boldsymbol{P}\boldsymbol{g}s + \delta_1\boldsymbol{z}^T\boldsymbol{P}\boldsymbol{g}v \\ &-pv^2 - pv^4 + \mathrm{sgn}[y]\theta_2 v^3 + \theta_1 v^2 \\ &+\mathrm{sgn}[y]\theta_0 v - \mathrm{sgn}[y]\eta wv \\ &+psv + ps^3 v - 3ps^2 v^2 + 3psv^3 \\ &+\delta_2\varepsilon\gamma_1\mathrm{sgn}[y]vw + \delta_2\varepsilon\gamma_0 w - \delta_2\varepsilon w^2\end{aligned} \tag{13}$$

Assume that p is chosen so that

$$p^2 - p - a > 0 \tag{14}$$

Then, the time derivative of V can be evaluated as

$$\begin{aligned}\dot{V} \le &-\alpha_1 s^2 - \alpha_2 s^4 - \delta_1\alpha_3\|\boldsymbol{z}\|^2 - \alpha_4 v^2 - \alpha_5 v^4 \\ &-\frac{1}{2}\delta_2\varepsilon w^2 + 2\delta_2\varepsilon\gamma_0^2 + \delta_3\end{aligned} \tag{15}$$

where

$$\alpha_1 = p^2 - p - a - \frac{|a|}{2} - \frac{3}{4\delta_1}\|\boldsymbol{h}\|^2$$

$$-\frac{3\theta_1^2}{4\delta_3}-\frac{3\theta_0^2}{4\delta_3}-\frac{2\eta^2}{\delta_2\varepsilon}-\frac{3}{4\delta_1}\|\boldsymbol{Pg}\|^2$$

$$\alpha_2=\frac{1}{4}p^2-p-\frac{|\theta_2|}{4}$$

$$\alpha_3=\lambda_{\min}[\boldsymbol{Q}]-1$$

$$\alpha_4=p-\frac{|a|}{2}+\frac{3}{4\delta_1}\|\boldsymbol{Pg}\|^2$$

$$-|\theta_2|-|\theta_1|-\frac{3\theta_0^2}{4\delta_3}-\frac{2\eta^2}{\delta_2\varepsilon}-2\delta_2\varepsilon\gamma_1^2$$

$$\alpha_5=\frac{1}{4}p-|\theta_2|$$

Choose sufficiently large p. Then, $\alpha_i>0\ \forall i$. Hence, from (15), it follows that

$$\dot{V}(t)\le-\alpha V(t)+\beta,\ \ t\in[0,t_F) \qquad (16)$$

where

$$\alpha=\min\left\{2\alpha_1,\frac{2\alpha_2}{\lambda_{\max}[\boldsymbol{P}]},2\alpha_3,\varepsilon\right\}$$

$$\beta=2\delta_2\varepsilon\gamma_0^2+\delta_3 \qquad (17)$$

Solving the differential inequality (16), the following inequality can be obtained.

$$V(t)\le V(0)e^{-\alpha t}+\frac{\beta}{\alpha},\ \ t\in[0,t_F) \qquad (18)$$

Therefore, if no failure occurs, that is, $t_F=\infty$, then all the signals in the control system are bounded. Moreover, taking $|y|\le|s|+|v|$ into consideration, it follows that

$$\lim_{t\to\infty}\sup|y(t)|\le\lim_{t\to\infty}\sup 2\sqrt{V(t)}\le 2\sqrt{\frac{2\beta}{\alpha}} \qquad (19)$$

It is clear that $2\sqrt{2\beta/\alpha}<\lambda$ for sufficiently small δ_2 and δ_3. This mean that the inequality (4) holds if no failure occurs. The proof is completed. □

Remark. Regarding the filtered signal v, it follows that

$$\lim_{t\to\infty}\sup|v(t)|\le\lambda$$

By setting the threshold v_T as $v_T<\lambda$, the auxiliary resetting (6) is not performed after the steady state, as long as the actuator is healthy.

4. Fault Detection

The sign of the output y does not change within a time period after the failure. That is, there is a time $t_E>t_F$ such that

$$\mathrm{sgn}[y(t)]=\mathrm{sgn}[y(t_F)],\ \ t\in[t_F,t_E) \qquad (20)$$

Now, consider the case where $\mathrm{sgn}[y(t_F)]=1$. The signals in the detection filter obey

$$\dot{v}=\theta_2v^2+\theta_1v+\theta_0-\eta w+p(y+y^3)$$
$$\dot{w}=\varepsilon(\gamma_1v+\gamma_0-w) \qquad (21)$$

Refer to the following Izhikevich neuron model[4] of the same scale as (21).

$$\dot{v}=\theta_2v^2+\theta_1v+\theta_0-\eta w$$
$$\dot{w}=\varepsilon(\gamma_1v+\gamma_0-w) \qquad (22)$$

Here, suppose that the parameters in (22) are chosen so that the bursting pattern[4,5] appears as shown in Figure 1.

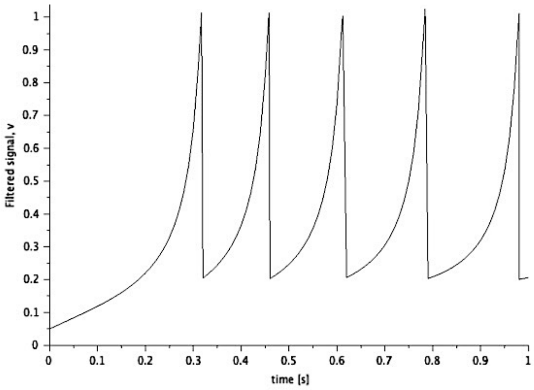


Fig. 1. Simulation results: the measured output and the actual output (top) and the filtered signal (bottom).

Note that it is possible to make the bursting time much shorter than t_E-t_F by changing time scale.

By comparing (21) with (22), it is shown that a stimulus injected into (21), is larger than the stimulus in (22) by the term $p(y+y^3)$. Generally, in a spiking neuron, a larger stimulus induces more spikes. Therefore, in the detection filter (21), a more oscillatory bursting pattern appears on the time period $[t_F,t_E)$. In the case of $\mathrm{sgn}[y(t_F)]=-1$, the same discussion as above holds.

If there is no failure, then, the above bursting does not occur because the filtered signal, v is suppressed smaller than v_T (see Remark).

Thus, by counting the number of the spikes, the failure can be found before the time t_E. Specifically, the detection time t_D is defined by

$$t_D:=\min\{t\mid c_R(t)\ge n_R\} \qquad (23)$$

where $c_R\in\mathbb{N}$ is the counted number of the spikes in the filtered signal v, and $n_R\in\mathbb{N}$ is the specified minimum number of spikes in the bursting pattern of (22).

After replacing the failed actuator, the boundedness of all the signals in the control system are guaranteed again, and the plant output can converge to the small region, that is the inequality (4) holds.

The overall control system is illustrated in Figure 2.

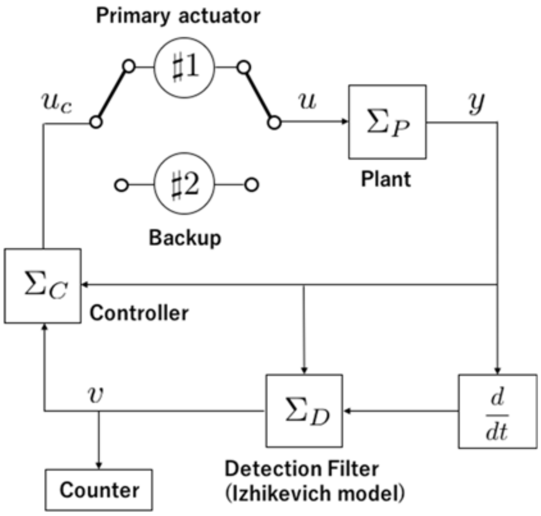


Fig. 2. Block diagram of the SRCS against actuator failures.

5. Numerical Examples

To confirm the effectiveness of the proposed method, the numerical simulation is explored.

Consider the following unstable plant.

$$\begin{aligned} \dot{y} &= -y + u + z,\ y(0) = 1 \\ \dot{z} &= -2z + y,\ z(0) = -1 \end{aligned} \tag{24}$$

The failure scenario is supposed that

$$t_F = 25\ [s],\ \varphi = u(t_F) \tag{25}$$

For the above plant, the parameters for the detection filter are selected as $\theta_0 = -0.06, \theta_1 = -0.6, \theta_2 = 4, \gamma = \delta = 1, \varepsilon = 0.02, \gamma_0 = -6, \gamma_1 = 20$. Also, the parameters for resetting are $v_T = 1, v_R = 0.2, w_R = 2$. The prespecified number of the spikes in the bursting pattern is supposed to be five per second, that is, $n_R = 5$ within almost 1 [s]. At last, the controller gain is chosen as $p = 6$.

The simulation results are shown in Figure 3 where the plant output y (top), the filtered signal v (middle), and the actual input u (bottom) are shown. From the simulation result, the failed actuator can be replaced at $t_D \cong 26$ [s]. Also, the plant output converges to a very small ball before and after the failure.

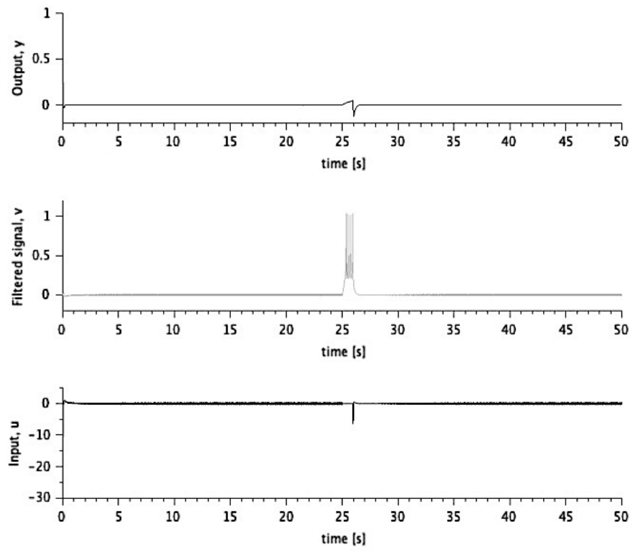


Fig. 3. Simulation results: the plant output (top), the filtered signal (middle) and the actual input (bottom).

6. Conclusion

This paper has presented the new SRCS that can find actuator failure by using the Izhikevich neuron model. From theoretical and numerical analysis, it is shown that fault-tolerant control can be accomplished.

Acknowledgements

This work is supported by JSPS KAKENHI Grant Number JP16K06429.

References

1. M. Takahashi: Self-repairing control against unknown sensor failures and biases, Asian J. Control, Vol. 15, Issue 4, pp. 1215-1223 (2013).
2. M. Takahashi: Self-repairing PI/PID control against sensor failures, Int. J. Innovative Computing, Information and Control, Vol. 12, No. 1, Feb. pp. 193-202 (2016)
3. A. Yanou, M. Minami, and T. Matsuno: "Strong Stability Rate for Control Systems Using Coprime Factorization", Trans. of the Society of Instrument and Control Engineers, Vol.50, No.5, pp.441–443 (2014) (in Japanese)
4. E. M. Izhikevich: Simple model of spiking neurons, IEEE Trans. on Neural Networks, Vol. 14, No. 6, pp. 1596-1572 (2003)
5. M. Takahashi: Izhikevich model-based self-repairing control for plants with sensor failures and disturbances, J. Robotics, Networking and Artificial Life, Vol. 6(2), pp. 105-108 (2019)

Deep object 6-DoF pose estimation using instance segmentation

Victor Pujolle
Computer science and system engineering, Kyushu institute of technology, 680-4 Kawazu
Iizuka, 820-0053, Japan
E-mail: victor.pujolle5@gmail.com,
https://sites.google.com/view/hayashilab-kyutech/home

Eiji Hayashi
Computer science and system engineering, Kyushu institute of technology, 680-4 Kawazu
Iizuka, 820-0053, Japan
E-mail: haya@mse.kyutech.ac.jp,
https://sites.google.com/view/hayashilab-kyutech/home

Abstract

Pose estimation algorithms' goal is to find the position and the orientation of an object in pace, given only an image. This task may be complex, especially in an uncontrolled environment with several parameters that can vary, like the object texture, background or the lightning conditions. Most algorithms performing pose estimation use deep learning methods. However, it may be difficult to create dataset to train such kind of models. In this paper we developed a new algorithm robust to a high variability of conditions using instance segmentation of the image and trainable on a virtual dataset. This system performs semantic keypoints based pose estimation without considering background, lighting or texture changes on the object.

Keywords: Pose estimation, deep-learning, keypoints localization, instance segmentation, virtual training, Factory automation

1.Introduction

The goal of the paper is to present a new framework to handle estimation of the 6 degree of freedom (6-DoF) pose (translation and rotation in 3D) of an object from a single image. It has many applications, especially in robotic manipulation. However, despite the need of a general and accurate framework, this task tends to be treated on a case by case basis. For example, approaches tend to differ depending of the object's texture.

Because 3D pose dataset are difficult and time consuming to make, training a network for general pose estimation in an uncontrolled environment with multiple object's instances is challenging. To tackle this problem, our approach combined previously existing pose estimation approach with a classic instance segmentation network. The algorithm can be divided in three stages. The first stage is the instance segmentation where we use a mask-R-CNN network[1] to provide semantic masks of objects' instances on an image. Because instance segmentation is already a well-studied subject this paper does not focus on this part. For this step, any state-of-the-art instance segmentation network would work, the better the accuracy, the better the final result. The second stage is the keypoints localization network. For this task we use a heavy convolutional network to predict of set of semantic keypoints[2,3,4]. This network takes as input the semantic masks extracted by the mask-R-CNN. Because the input is only a mask of the object, it is possible to train this network using virtual images generated randomly with a simple mesh of the object. This way, the creation of the virtual dataset does not require skills in 3D modelling to beat the reality gap[5].

The last stage is to solve the pose to perspective problem[5]. Because it is already a well-documented subject, this paper does not focus on this step. While this work focuses only of RGB input, it could be adapted fairly easily to RGBD inputs with some light modification in the network's architecture and for the virtual dataset generation.

2.Related work

6-DoF pose estimation is extremely useful for many robotic tasks and several approaches of the subject have already been studied. But these methods typically address the problem for highly textured objects. This leads to failure for textureless objects and reduces generalization capability. Other methods like dense pixel voting or gradient mapping[7,8] works on textureless objects. However, they need instance specific 3D model of every object, limiting their real applications. Methods using 2D landmarks localization on the image can works for textureless objects. Finding the pose of a rigid object given the position of *n* landmarks on the image is commonly referred as Pose to Perspective problem[6] (PnP).

3.Technical Approach

3.1Keypoints localization

The keypoints localization is made using the "stacked hourglass" network architecture[4] that has been proven effective for human pose estimation using keypoints It seems natural to think that the same architecture could be used for object keypoints localization. This hourglass network is plugged directly at the end of an already trained mask-R-CNN network. The mask-R-CNN reduces objects to masks, removing background, textures and specific details. Only the form of the object projected on the camera captor is left. This enables the stacked hourglass to consider only the form of the object and helps the generalization.

3.1.1 Network architecture

The network architecture (Fig 1) is mostly inspired of the architecture presented in Refs 2, 3 & 4 with few minor modifications. The network input is a mask extracted by the mask-R-CNN from an RGB image and the output is a set of heatmaps, one for each keypoint. The heatmap intensity at a given pixel indicates the confidence of the respective keypoint to be located at this pixel. The network is made of stacked hourglasses. Each hourglass consists of two processing stage. In the first stage, a series of convolution and max-pooling layers are applied the input image. Each max-pooling layer divides by two the resolution of the feature map. This process continues until the resolution reach a minimal resolution set by the user (4 × 4 in our model). After this down-sampling process, series of deconvolution layers[10] are applied to the feature map. A residual layer is also applied[11] to increase the accuracy of the network. This process continues until the feature map reaches the input resolution. Another hourglass can be stacked to the end of the first one to refine the output heatmaps. The ground truth labels used for the training are heatmaps made by applying 2D gaussian centered at every keypoint. The loss function used is the l2 norm. It is possible the use intermediate supervision at the end of each hourglass module the increase the accuracy and fasten the training by providing more signal for the gradient back propagation[12]. The output of the last module is used as the output of the whole network. The maximum of each heatmap indicates the localization of the respective keypoint.

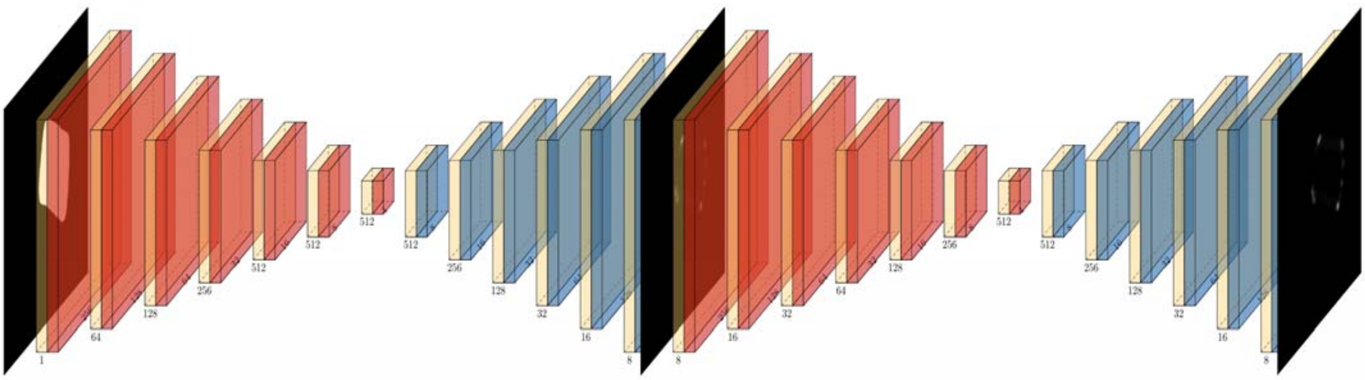


Fig. 1. Architecture of the keypoints localization network with intermediate supervision. Max-pooling layers are drawn in red, deconvolution layers in blue. The symmetric nature of this architecture makes possible the use of residual layers, which are not represented. The size of the feature maps and number of channels are indicated for each layer.

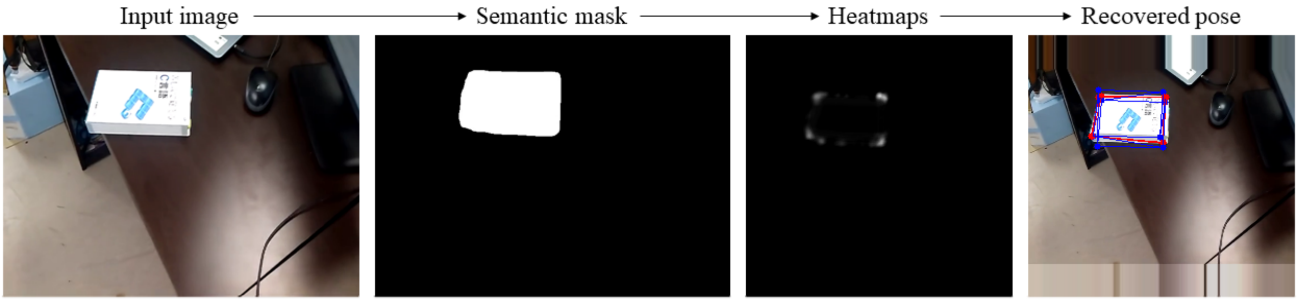


Fig. 2. Full workflow of the algorithm. The semantic mask is inferred by a mask-R-CNN. Then the heatmaps are created by the stacked hourglass network. Finally, the pose is recovered from the keypoints with the solvePnPRansac from the OpenCV library. On the final image, 2D estimated keypoints are drawn in red and the projection of the bounding box with the recovered 6-DoF pose is drawn in blue.

3.1.2 Design explanation

The main goal of this architecture is to provide a heatmaps with the same resolution of the input. The combination of top-down and bottom-up processing enable the network to use local and global features of the input. This is especially useful given the large variability in scale of the objects. The addition of intermediate supervision has been proven useful to increase gradient signal and therefore fasten the training process while reducing issues such as vanishing gradient[12]. Finally, residual layers, which have achieved state-of-the-art results[11] are used.

3.2Training

3.2.1 Virtual training

Training may become very challenging when no dataset is available for the task. In the case of pose estimation, most datasets focus on very few objects with limited number of instances in a very controlled environment. Because labelling data for pose estimation is extremely time consuming, virtual training has been proven efficient to train state-of-the-art pose estimation algorithms[5]. However, creating realistic scenes and 3D models of objects can also be very challenging. In this paper we use a virtual dataset created using only a simple mesh of every objects. This approach makes possible to create important dataset in a limited amount of time. To do so we compute mask projection of the object with random poses as illustrated in (Fig 3). This mask projection is made using OpenCV functions for projection and drawing. We also apply different kinds of blur to simulate the imperfection of the mask-R-CNN's outputs. These images are similar to the output of an instance segmentation network; therefore, they can be used to train the keypoints localization network. Because the mask-R-CNN is easy to train with a good accuracy, we consider only the training of the hourglass network. However, the virtual dataset is not a perfect representation of the mask-R-CNN's output and some issues with the reality gap may appear.

Fig. 3. Examples of virtual masks for 3 objects: a book, a bottle and a cup.

4.Experiments

In this section we focus on pose estimation for specific object instance. This case in useful for many robotic applications where the objects and the environment are controlled. However, because we did not create dataset made of real image, the evaluation of the results for real images is only visual. We took videos of a book, a bottle and a cup and tried to recover the pose of the object for every frames of these videos. The pose recovered by the algorithm. Then the bounding box of the object is projected again on the image to give a visual representation of the result. Our goal is to be able test the algorithm on standard datasets like PASCAL3D+[13].

Fig. 4. Illustration of the pose recovery. The output of the network is drawn in red, the projected bounding box in blue.

Our method achieves failry stable and accurate results for single object pose estimation but still need to be tested for multiple objects and occlusion cases. Because the first step of the algorithm is to perform instance segmentation of the image, a vast variability of objects and backgrounds are supported. However the drawback come from the accuracy of the mask-R-CNN that can be insufficient for a proper detection of the keypoints. We think that improvements of the instance segmentation's state-of-the-art will help improving the accuracy of our approach.

5.Summary

In this paper we proposed an original method to infer the 6-DoF of an object from a single RBG image. This keypoints localization network can be trained very easily using a dataset made of virtual images and can be plugged over any instance segmentation network with only few adjustments. With its easy training process and the good generalization capability, our framework can be useful for many robotic applications where the creation of a specific dataset of real annotated images cannot be done.

6.References

1, Kaiming He, Georgia Gkioxari, Piotr Dollár & Ross Girshick, Mask R-CNN, IEEE ICCV, 2017.
2.Georgios Pavlakos, Xiaowei Zhou, Aaron Chan, Konstantinos G. Derpanis & Kostas Daniilidis, 6-DoF Object Pose from Semantic Keypoints, IEEE ICRA, 2017.
3, Junsu Kang, Dayeon Kim, Victor Pujolle, Juyong Lee, Donghyeon Lee, Young Jin Heo & Wan Kyun Chung, 3D object pose recognition framework for robot task based on RBG image, Postech domestic conference, 2019.
4, Alejandro Newell, Kaiyu Yang & Jia Deng, Stacked hourglass networks for human pose estimation, European Conference on Computer Vision ECCV, 2016
5. Jonathan Tremblay, Thang To, Bala Sundaralingam, Yu Xiang, Dieter Fox & Stan Birchfield, Deep Object Pose Estimation for Semantic Robotic Grasping of Household Objects, Conference on Robot Learning (CoRL), 2018.
6. Lepetit, V. & Moreno-Noguer, EPnP: An Accurate O(n) Solution to the PnP Problem, International Journal of Computer Vision, 2009.
7. Peng, Sida & Liu, Yuan & Huang, Qixing & Bao, Hujun & Zhou, PVNet: Pixel-wise Voting Network for 6DoF Pose Estimation, Xiaowei, 2018.
8. S. Hinterstoisser, C. Cagniart, S. Ilic, P. F. Sturm, N. Navab, P. Fua, & V. Lepetit, Gradient response maps for real-time detection of textureless objects, PAMI, 2012.
9. A. Toshev & C. Szegedy, DeepPose: Human pose estimation via deep neural networks, in CVPR, 2014.
10. Zeiler & Matthew D., Deconvolutional networks, Computer Vision and Pattern Recognition (CVPR), IEEE (2010).
11. K. He, X. Zhang, S. Ren, & J. Sun, Deep residual learning for image recognition, CVPR, 2016.
12. C.-Y. Lee, S. Xie, P. Gallagher, Z. Zhang, & Z. Tu, Deeply supervised nets, AISTATS, 2015
13. Y. Xiang, R. Mottaghi, & S. Savarese, Beyond PASCAL: A benchmark for 3D object detection in the wild, WACV, 2014.

Autoencoder with Spiking in Frequency Domain for Anomaly Detection of Uncertainty Event

Umaporn Yokkampon[1], Sakmongkon Chumkamon[1], Abbe Mowshowitz[2] and Eiji Hayashi[1]
[1]Department 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, m-san@mmcs.mse.kyutech.ac.jp, amowshowitz@ccny.cuny.edu, haya@mse.kyutech.ac.jp
www.kyutech.ac.jp

Abstract

This paper proposes the autoencoder method with spiking raw data to the frequency domain to analyze and predict the anomaly case among the standard data set and compare it with original data. The dataset is the real-world data from factory automation. The combination of frequency domain and original data can improve the validity and accuracy in detecting an anomaly data. Therefore, analyzing time-series data using combination of autoencoder and the frequency domain can be efficient in detecting anomalies.

Keywords: Anomaly detection, Autoencoder, Data mining, Factory automation.

1. Introduction

Anomaly detection in time-series is an important real-world problem. Nowadays, sensors and internet of things devices are ubiquitous and produce data continuously. While the data gathered by these devices is valuable and can provide meaningful insights, there is a growing need for developing algorithms that can process these data efficiently. Anomaly detection is applied in network intrusion detection, credit card fraud detection, sensor network fault detection, medical diagnosis and numerous other fields.[1]

Among many anomaly detection methods, there are a few ways to reduce the dimensions of large data sets to ensure computational efficiency such as backwards selection, removing variables exhibiting high correlation, high number of missing values but by far the most popular is Principal Components Analysis (PCA). A relatively new method of dimensionality reduction is the autoencoder. Autoencoders although is quite similar to PCA but its autoencoders are much more flexible than PCA. Autoencoders can represent both liners and non-linear transformation in encoding but PCA can only perform linear transformation. Autoencoders are a branch of neural network which attempt to compress the information of the input variables into a reduced dimensional space and then recreate the input data set. Moreover, the autoencoders can apply dimension reduction in a hierarchical manner, obtaining more abstract features in higher hidden layers leading to a better reconstruction of the data[2].

In this paper, the autoencoder method with underlined spiking raw data has been proposed to detect any anomaly by frequency domain to analyze and predict. Moreover, we also compare and validate our proposed method with original data, which is from factory automation data and real-world data based on AUC, Precision, Recall and F1-Score criteria.

2. Proposed Method

We proposed the novel method using the autoencoder method with spiking raw data to the frequency domain to analyze and predict the anomaly case. Step by step procedure is as follows:
1. Divide all datasets into 2 groups.
2. For the second group, calculate the FFT of each dataset.
3. Spike plot by using FFT values.
4. Combine raw data and FFT values of each dataset. This is the input of the second group. For first group has only original data.
5. Input both groups of all datasets into the autoencoder process.

Fig. 1 describe the whole procedure algorithmically.

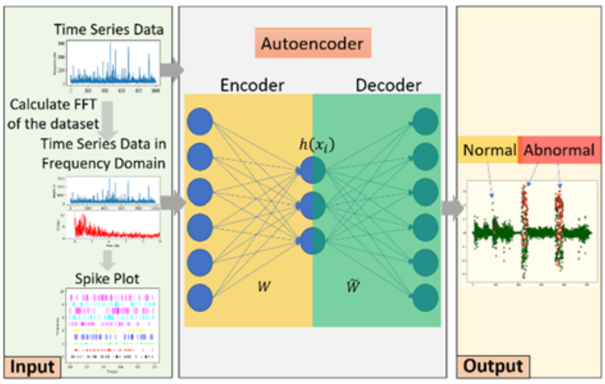


Fig. 1. Architecture of the proposed method.

2.1. *Autoencoder*

Autoencoder neural network is an Unsupervised Machine learning algorithm that applies backpropagation, setting the target values to be equal to the inputs. Autoencoder consists of 3 components: encoder, code, and decoder. The encoder compresses the input and produces the code, the decoder then reconstructs the input only using this code. Given a data point $x \in \mathbb{R}^d$ (d is the dimension of x), the loss function can be viewed as minimizing the reconstruction error between the training data and the outputs of the autoencoder[3] :

$$L(x, \hat{x}) = \| x - \hat{x} \|^2 \quad (1)$$

After training, autoencoder uses the reconstruction error as the anomaly score. The data with high anomaly score is considered to be 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 with anomalous data the autoencoder has not encountered.

2.2. *Frequency Domain (FFT)*

Frequency-domain analysis is a tool of utmost importance in signal processing applications. Frequency-domain analysis is widely used in such areas as communications, geology, remote sensing, and image processing. While time-domain analysis shows how a signal changes over time, frequency-domain analysis shows how the signal's energy is distributed over a range of frequencies.

Fast Fourier Transform (FFT) is a mathematical technique for converting a signal from the time domain into the frequency domain. Signals on a flaw detector, on an oscilloscope or on an acoustic emission display are typically time domain signals showing how the amplitude varies with time. When transformed into the frequency domain, the display shows how the amplitude varies with the frequency. This display is often referred to as the frequency spectra of the signal. A sine wave of a single frequency in the time domain will give a single line at that frequency in the frequency domain. A spike in the time domain will give a spread of frequencies in the frequency domain.

The FFT computes the discrete Fourier transform (DFT) in an efficient manner. The DFT is defined given by

$$H_k = \sum_{i=0}^{n-1} x_i e^{2j\pi ik/n} \quad (2)$$

where j is the imaginary number $\sqrt{-1}$, and n is the number of points in T and F.

2.3. *Spike Plot*

Spikes represent any number of horizontal or vertical line segments with fixed or variable heights. It is usual to use in time series plots. They may also be useful in more domain-specific cases, such as visualizing spike trains for neurophysiology or spectrograms in physics and chemistry applications.

3. Experiments

This section introduces the data sets we used and the evaluation metric we compared and evaluate to demonstrate the effectiveness of our proposed method in anomaly detection.

3.1. *Data sets*

To illustrate the effectiveness of our method, time series data is got from UCR public data set[4]. More details about datasets are shown in Table 1. All datasets are given in time series form and every data point is manually labelled. All datasets, we choose the minority class as anomaly class. We split 20% of the data as test data.

Table 1. Summary of the datasets.

Datasets	Length	Number of instances	Anomaly Ratio
SonyAIBORobotSurface2	65	980	0.38
ItalyPowerDemand	24	1096	0.49
Wafer	152	7164	0.11

3.2. *Performance Evaluation*

We evaluate the accuracy of anomaly detection method using Area under the curve of the receiver operating characteristic (AUC), Precision (Pre), Recall (Rec), and F1-Score, which are defined as follows:

$$\text{Pre} = \frac{TP}{TP+FP} \tag{3}$$

$$\text{Rec} = \frac{TP}{TP+FN} \tag{4}$$

$$\text{F1} = 2 \times \frac{\text{Pre識 ec}}{\text{Pre}+\text{Rec}} \tag{5}$$

where TP is the correctly detected anomaly, FP is the falsely detected anomaly, TN is the correctly assigned normal, and FN is the falsely assigned normal.

4. Results and Discussion

We evaluate the anomaly detection performance for both spiking raw data and combine frequency domain using the autoencoder method. For the second group, we calculated FFT of all datasets and spike plot. Example of spike plot is shown in Fig. 2. We showed the results of the wafer dataset.

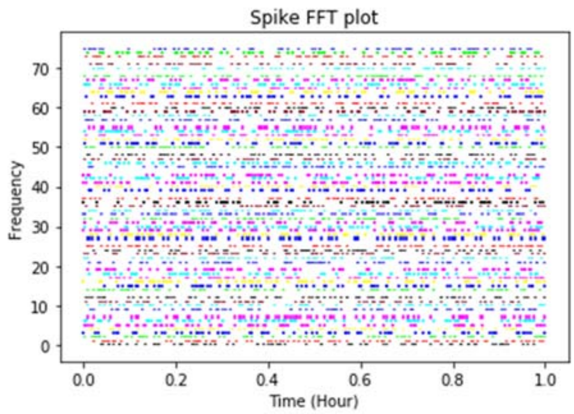
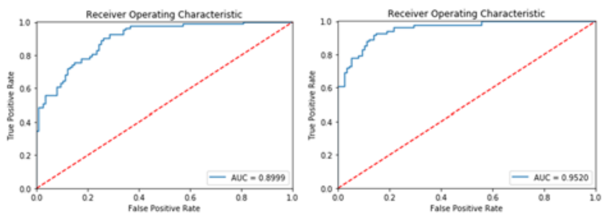
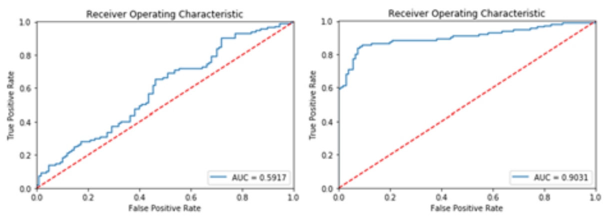


Fig. 2. The example of spike plot.

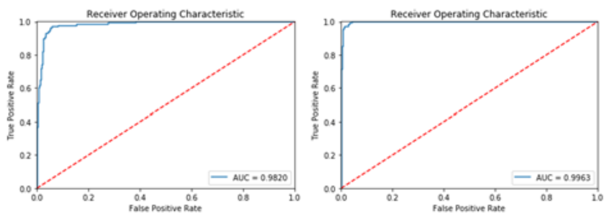
We performed experiments on accuracy on 3 data sets and used AUC as the criterion. From the Fig. 3. The left column is original data and the right column is raw data combined with frequency domain. We can see that our method could outcome AUC value more than original data from all datasets. It shows that our method has good performance in anomaly detection for time series data.



(a) SonyAIBORobotSurface2



(b) ItalyPowerDemand



(c) Wafer

Fig. 3. AUC comparisons between original data and combine frequency domain.

Table 2. Comparison between original data and combine frequency domain.

Datasets	Original Data				Combine Frequency Domain			
	AUC	Precision	Recall	F1-Score	AUC	Precision	Recall	F1-Score
SonyAIBORobotSurface2	0.8999	0.9043	0.7647	0.8287	0.9520	0.9565	0.8333	**0.8907**
ItalyPowerDemand	0.5917	0.7091	0.5166	0.5977	0.9031	**0.9727**	0.7279	0.8327
Wafer	0.9820	0.7349	0.9979	0.8464	**0.9963**	0.8008	**1.0000**	0.8894

The anomaly detection results and comparisons are summarized in Table 2. The results show that the method that we proposed outperformed the original data in all aspects and datasets. In particular, the wafer dataset gave recall is perfect.

5. Conclusion

In this paper, we propose the autoencoder method with original data compared with spiking raw data to the frequency domain to analyze and predict the anomaly case based on AUC, Precision, Recall and F1-Score criteria. The evaluation results show that our proposed method in anomaly detection improves accuracy in all aspects. Therefore, analyzing time-series data in the frequency domain can be efficient in detecting anomalies for time series data.

In the future we will use the variational autoencoder with time series data to the frequency domain to analyze and predict the anomaly case and compare with the autoencoder method.

6. References

1. Varun Chandola, Arindam Banerjee, and Vipin Kumar, Anomaly detection: A survey, *ACM computing surveys (CSUR)*. **41**(3), 2009
2. J. An and S. Cho. *Variational Autoencoder based Anomaly Detection using Reconstruction Probability*. Technical Report. SNU Data Mining Center, 2015.Xuhong Wang,
3. Ying Du, Shijie Lin, Ping Cui, and Yupu Yang. Self-adversarial variational autoencoder with gaussian anomaly prior distribution for anomaly detection. *CoRR*, abs/1903.00904, 2019.
4. H. A. Dau, E. Keogh, K. Kamgar, C.-C. M. Yeh, Y. Zhu, S. Gharghabi, C. A. Ratanamahatana, Yanping, B. Hu, N. Begum, A. Bagnall, A. Mueen, and G. Batista, "The ucr time series classification archive," October 2019, https://www.cs.ucr.edu/~eamonn/time series data 2018/.

Evaluation of the Relationships Between Saliency Maps and Keypoints

Ryuugo Mochizuki
Center for Socio-Robotic Synthesis, Kyushu Institute of Technology,
*2-4, Hibikino Wakamatsuku, Kitakyushu, 808-0196, Japan**

Kazuo Ishii
Center for Socio-Robotic Synthesis, Kyushu Institute of Technology,
2-4, Hibikino Wakamatsuku, Kitakyushu, 808-0196, Japan
E-mail: rmochizuki@lsse.kyutech.ac.jp, ishii@brain.kyutech.ac.jp
www.lsse.kyutech.ac.jp/~sociorobo/ja/

Abstract

Saliency is a property of images that triggers bottom-up attention. For example, if a location in an image is sufficiently different from its surrounding and worthy of paying attention, such characteristic of image is saliency. From the point of view, the location of larger saliency is outstanding visually. On the other hand, As Image Feature extraction method, such as SIFT or SURF, robust feature matching has been realized under the existence of changing size or rotation of observed target. For the consequence, its advantage has been introduced into image stitching and Visual SRAM. However, the amount of image features is susceptive to changing photographing condition, such as luminance variety, defocus-ing etc. We assumed that feature extraction stability is large in salient region because of steep bright-ness gradient. We evaluated the relationship between saliency and feature extraction stability.

Keywords: Saliency Map, Spatial-frequency, Invariant Image Feature, Filter Tuning

1. Introduction

In recent years, many attempts have been done such as the selection of desired information in input information[1][2]. If attention models can be constructed to select information, the intelligence and awareness of humans can be implemented in computers.

According to Itti et.al, saliency is defined as the property of images, which triggers bottom-up attentions. Saliency occurs by the local conspicuity over the entire visual scene[3]. In this model, input image is decomposed into luminance, color, and orientation components, then, each component is processed individually with Gaussian filter.

Considering that the saliency map is applied to environment recognition by mobile robots, various changes in photographing condition are expected to affect the input image. The change affects spatial frequency components of the image. If the spatial frequency changes, the response of Gaussian filter also changes, then, the effect reflects saliency map. Considering that the saliency map is used to select the keypoints of the image, Changes in the saliency map affect the results of feature selection, then, input data of detectors vary. Thus, recognition results are influenced according to the change in photographing conditions.

For keypoint extraction, small influence is desirable in spite of the variety of object size, angle and luminance. In case of the keypoint application for object detection, repetitively extracted keypoints are ideal to select.

In our research, we propose a method for generating saliency maps, which can absorb the effect of spatial frequency changes. If the parameters of the filters can be determined automatically, the effect of the spatial frequency change can be diminished in saliency maps (Fig. 1) We evaluated the relationship between saliency and keypoints.

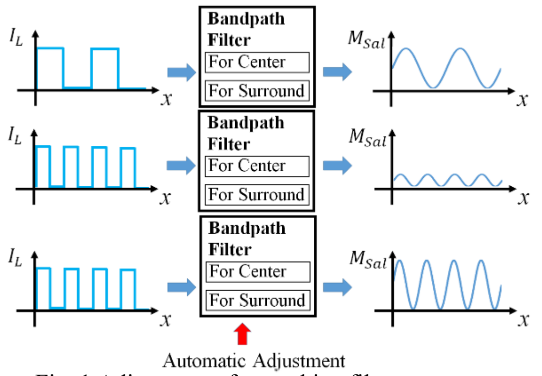


Fig. 1 Adjustment of smoothing filters against changing spatial frequency

2. Related Work

2.1. *Saliency Map*

Itti et. al. simulated human eye movement, and expressed the result as saliency maps [3]. In the process of saliency map creation input image is reduced by $1/2^n$ and nine resolutions of the images are obtained. The Center and the Surround can be obtained through the smoothing operation by a common Gaussian filter. This signal process is similar to the different responses from fovea and its neighbor in retina for the common stimuli. All the reduced images are enlarged to the same size, and the across scale difference image of the two conpornents is normalized and added to obtain a map for each component (i. e. Luminance, Color, Orientation). Saliency map is obtained through the addition of all the maps of the three components.

According to [4], saliency map changes if the parameter of the Gaussian filters are changed. The ratio of filter parameter σ_c/σ_s is crucial for the determination of saliency. Arbitral selection of σ_c/σ_s enabled high granularity in saliency map. However, in [3][4], the parameters cannot be adjusted depending on the variety of spatial frequency. As the result, saliency map can be affected in the event of spatial frequency change.

2.2. *Keypoint Extraction*

Keypoint extraction is often applied for object recognition tasks [5], image stitching tasks [6], etc. by robot vision. A keypoint has a co-ordinate, a descriptor which explains brightness gradient in the neighborhood. In the object recognition task, the database image and the newly observed image are searched. Recently, scale-knvariant keypoint extraction methods have been proposed, such as SIFT[7], and BRISK[8]. As the result, the stability of object detection has been improved. However, if photographing conditions (brightness of the environment, size of the observed object, focusing conditions, camera internal parameters, etc.) change, the number of extracted keypoints changes significantly. Stably extracted keypoints are desirable for the use of object detection tasks by robot vision..

3. Proposal of Saliency Map

3.1. *Outline*

In this research, we developed the theory of [4] to mitigate the effect of spatial frequency variation. The strategy is automatic adjustments of σ_c and σ_s. In the saliency map generation process (Fig. 2), the input image is decomposed into luminance, color, and orientation components in advance. For each component, the Center and the Surround are generated by the combination of integral image and box filters. The parameter of the filters are automatically adjusted so that the pixel values of the across scale difference are maximized. The across scale

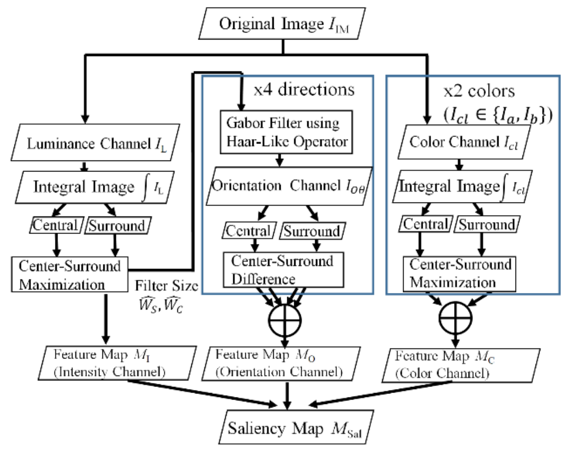


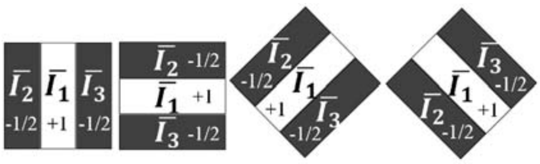
Fig. 2 Overview of proposed saliency map method

differences of all three components are merged to form saliency map.

3.2. *Decomposition of Input*

We utilize CIE-Lab color system to simplify the difference of complimentary color channel. I_L, I_a and I_b indicates luminance, color (Red-Green), color (Blue-Yellow) component, each other. For the obtainment of orientation component, Haar-Like Filters(Fig. 4[9][10]) are convoluted on I_L. The operations are expressed as Eq. (1)

$$I_\theta(\boldsymbol{p_p}) = \left|\overline{I_1(\boldsymbol{p_p})} - \frac{1}{2}\overline{I_2(\boldsymbol{p_p})} - \frac{1}{2}\overline{I_3(\boldsymbol{p_p})}\right| \quad (1)$$



(a)0[deg] (b)90[deg] (c)90[deg] (d)135[deg]
Fig. 3 Haar-like Filters

3.3. *The Center and Surround*

We align two box filters F_{Bs}, F_{Bc} centered with point $\boldsymbol{p_p}$ as Fig. 4 shows. The filters are used for convolution to generate the Center and Surround. The filter widths W_{Bs}, W_{Bc} can be variable up to W_{pmax} and fulfills W_{Bs}>W_{Bc}. This arrangement is same as [11]

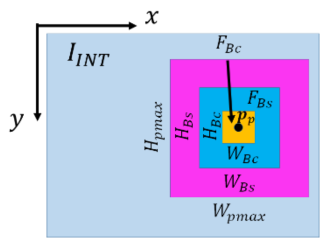


Fig. 4 Alignments of box filters

3.4. *Filter Adjustment*

To obtain across scale difference of luminance, color components, we maximize the pixel value of the difference $I_{cs}(\boldsymbol{p_p})$ as in [11] by changing $W_{Bs}(\boldsymbol{p_p})$, $W_{Bc}(\boldsymbol{p_p})$ according to Eq. (2) and Fig. 4.

$$\widehat{I_{cs}(\boldsymbol{p_p})} = \max_{W_{Bc}(\boldsymbol{p_p}),W_{Bs}(\boldsymbol{p_p})} I_{cs}(\boldsymbol{p_p})$$
$$= \max_{W_{Bc}(\boldsymbol{p_p}),W_{Bs}(\boldsymbol{p_p})} \left|I_c(\boldsymbol{p_p}) - I_s(\boldsymbol{p_p})\right| \quad (2)$$

Here, $W_{Bs}(\boldsymbol{p_p})$, $W_{Bc}(\boldsymbol{p_p})$ satisfies $\widehat{W_{Bs}(\boldsymbol{p_p})}$, $\widehat{W_{Bc}(\boldsymbol{p_p})}$.

On the other hand, for orientation component, to obtain across scale differences the sizes of the Haar-like Filters are set to $\widehat{W_{Bs}(\boldsymbol{p_p})}$, $\widehat{W_{Bc}(\boldsymbol{p_p})}$, then, the filters are convoluted with I_L. The responses of the Center and the Surround are denoted as $I_{\theta\_c}(\boldsymbol{p_p})$, $I_{\theta\_s}(\boldsymbol{p_p})$. The across scale differences of all directions are obtained and merged to map $M_{O\_\theta}(\boldsymbol{p_p})$.

3.5. *Saliency Map Generation*

Map M_C (for Color), and M_O (for Orientation) are obtained by Eq. (4)(5). Saliency map M_{Sal} is formed through the merge of M_I, M_C, M_O with Eq. (6). The functions f_{mix}, g_{mix}, h_{mix} for merging maps can be selected arbitrarily.

$$M_C = f(M_{Ca}, M_{Cb}) \quad (4)$$
$$M_O = g(M_{O\_0}, M_{O\_45}, M_{O\_90}, M_{O\_135}) \quad (5)$$
$$M_{Sal} = h_{mix}(M_I, M_C, M_O) \quad (6)$$

4. Evaluation of the relationship between saliency and keypoint extraction

4.1. *Outline of the Experiment*

In this experiment, we assume that the selected image keypoints are used for object detection. Thus, we evaluate the relationship between saliency $\boldsymbol{M_{Sal}}$ and feature stability $\boldsymbol{F_{Stb}}$. Suppose the number of small regions is N_q, $\boldsymbol{F_{Stb}}$ and $\boldsymbol{M_{Sal}}$ are expressed in line vector of N_q dimensions. However, we treat $\boldsymbol{M_{Sal}}$ and $\boldsymbol{F_{Stb}}$ as two dimensions (Fig. 5). Then, we calculate the relationship ϕ_i by obtaining inner product $\boldsymbol{F_{Stb}} \cdot \boldsymbol{M_{Sal}}$. The saliency maps were generated by conventional methods (i. e. Itti method, VOCUS2) and our proposal to compare ϕ_i. The source codes for the experiment are Simpsal[12] by Caltech for Itti method, and [13] for VOCUS2. We chose BRISK[8] as keypoint extraction method because descriptor is expressed in binary system. Such system is reported to require shorter time for matching than SIFT[7]. Furthermore, the descriptor has properties of rotation and scale invariance.

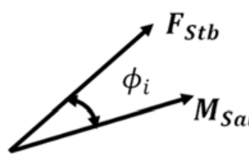


Fig. 5 Relationship between keypoint stability $\boldsymbol{F_{Stb}}$ and saliency $\boldsymbol{M_{Sal,i}}$

4.2. *Evaluation Function*

We consider two conditions of keypoints which have high stability under photographing condition variety. Firstly, the keypoints must extracted at the same location. We define the property as repeatability. Secondly, the descriptors must remain the same, that is, the similarity.

To evaluate keypoint stability, keypoint displacement have to be considered because of image flicker, resize of observed object size. For example, the combination of the same keypoints is considered as (I) or (II) in Fig. 6 in different photographing condition. We define a small region of W_qxH_q [Pixels] to search identical keypoints

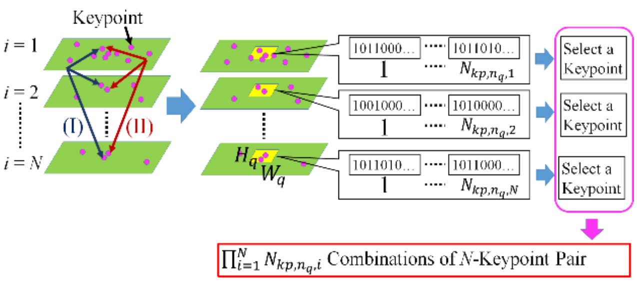


Fig. 6 Ambiguity in keypoint identification under changing photographing condition

Suppose $N_{kp,n_q,i}$ keypoints are extracted at n_q-th small region under i-th photographing condition, the variance σ_{kp,n_q} of keypoint number is obtained by Eq. (7). The average $\overline{N_{kp,n_q}}$ (for N variations of a parameter) of extracted keypoints is obtained by Eq. (8).

$$\sigma_{kp,n_q} = \frac{1}{N}\sum_{i=1}^{N}\left(N_{kp,n_q,i} - \overline{N_{kp,n_q}}\right)^2 \tag{7}$$

$$\overline{N_{kp,n_q}} = \frac{1}{N}\sum_{i=1}^{N} N_{kp,n_q,i} \tag{8}$$

r_{ftr,n_q} is obtained by the normalization of σ_{kp,n_q} r_{ftr,n_q} should be larger if σ_{kp,n_q} is smaller as Eq. (9) shows.

$$r_{ftr,n_q} = 1 - \frac{\sigma_{kp,nq}}{\max_{nq} \sigma_{kp,nq}} \tag{9}$$

To obtain similarity, we select two keypoints from the same small region (as seen in Fig. 6) and different photographing conditions, then calculate Hamming distance between the two descriptors. To obtain average Hamming distance of all combinations of the keypoint pairs, we use Eq. (10). The similarity s_{Dsc,n_q} is calculated with normalization by Eq. (11) so that the range satisfies [0, 1], and r_{Derr,n_q} is smaller as the distance is larger.

$$r_{Derr,n_q} = \frac{\min_{K} \sum_{l=1}^{N-1}\sum_{m=l+1}^{N} d_H(d_{n_q,l,k_l}, d_{n_q,m,k_m})}{c\binom{N}{2}} * \frac{1}{L_D} \tag{10}$$

$$\text{s.t. } \boldsymbol{K} = [k_1, k_2, k_3, \dots, k_{N_i}], l, m = 1, 2, \dots, N_i, l \neq m$$

$$s_{Dsc,n_q} = 1 - \frac{r_{Derr,nq}}{\max_{nq} r_{Derr,nq}} \tag{11}$$

Keypoint stability of F_{Stb,n_q} is calculated by the weighting of r_{Derr,n_q} and s_{Dsc,n_q} as Eq. (12) shows.

$$F_{Stb,n_q} = w r_{ftr,n_q} + (1-w) s_{dsc,n_q} \tag{12}$$

For the saliency M_{Sal,i,n_q}, The maximum response of M_{Sal} is searched within each small region. Maximum saliency and feature stability are expressed as N_q dimensions of line vectors (denoted as $\boldsymbol{M_{Sal,i}}$, $\boldsymbol{F_{Stb}}$, respectively). ϕ_i is calculated as the angle between the two vectors (Eq. (13)). To be noted that r_{ftr}, s_{dsc} are calculated only for the regions where keypoints are extracted more than twice during N variations of photographing conditions.

$$\phi_i = \cos^{-1}\left(\frac{\boldsymbol{M_{Sal,i}} \cdot \boldsymbol{F_{Stb}}}{\|\boldsymbol{M_{Sal,i}}\|\|\boldsymbol{F_{Stb}}\|}\right) \tag{13}$$

$$\boldsymbol{F_{Stb}} = \left[F_{Stb,1}, F_{Stb,2}, \dots, F_{Stb,n_q}, \dots, F_{Stb,N_q}\right]$$

$$\boldsymbol{M_{Sal,i}} = \left[M_{Sal,i,1}, M_{Sal,i,2}, \dots, M_{Sal,i,n_q}, \dots, M_{Sal,i,N_q}\right]$$

The average $\bar{\phi}$ is obtained according to Eq. (14).

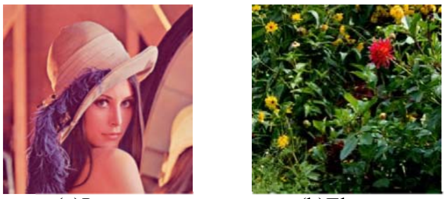
$$\bar{\phi} = \frac{1}{N}\sum_{i=1}^{N} \phi_i \tag{14}$$

4.3. *Method*

Fig. 7 shows the experimental images (Lenna, Flower, Tree, Things). These images were selected in the database of Caltech[14] and SIDBA[15]. The spatial frequency spectrums of the images are shown in Fig.8. Lenna is a well known for test image to be used image analysis. Flower has wider spectrum than Lenna with higher frequency component. As well as the comparison of Things and Tree, Things has higher frequency component than Tree.

The photographing condition to adjust to vary extracted keypoint number is $I_{Max,i}/I_{Max,1}$ for luminance, $W_{Obj,i}/W_{Obj,1}$ for object size, each other, whose range is from 0.5 to 1.0 with the step 0.1 of increase.

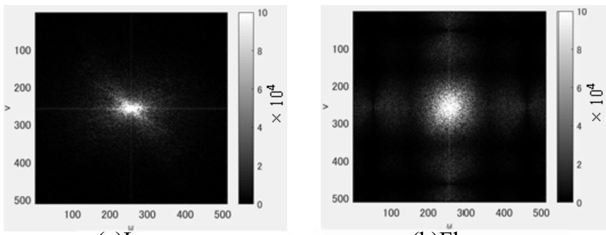
For changing $W_{Obj,i}$, we selected images of no white backbround, (i. e. Tree and Flower). We selected T_{FAST}=20 (T_{FAST}: Threshold of FAST Score[8]) and $I_{Max,1}$ = 255 during the adjustment of $I_{Max,i}$ and $W_{Obj,i}$. As the setting of the proposal, for Setting 1, $W_{pmax} = W_{IM}/4$ and for Setting 2, $W_{pmax} = W_{IM}/2$. W_{IM} indicates the image width. The resolution of the image is $W_{IM} \times H_{IM}$ = 512x512 [Pixel].



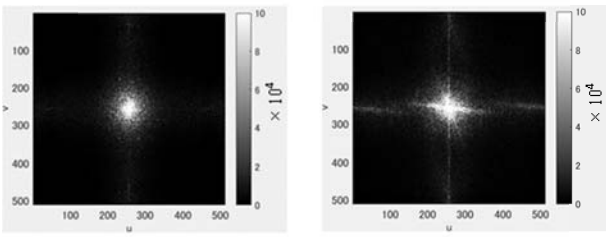
(a)Lenna (b)Flower

(c)Tree (d)Things

Fig. 7 Experimental Images (※Cited from [14][15])



(a)Lenna (b)Flower



(c)Tree (d)Things

Fig. 9 Spectrum of spatial frequency

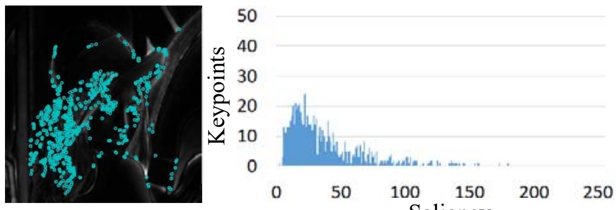
4.4. *Result and Discussion*

Table 1, 2 shows the relationship of $\boldsymbol{M_{Sal}}, \boldsymbol{r_{ftr}}$ and $\boldsymbol{M_{Sal}}, \boldsymbol{s_{dsc}}$ under variable $I_{Max,i}$, and $W_{Obj,i}$, each other.

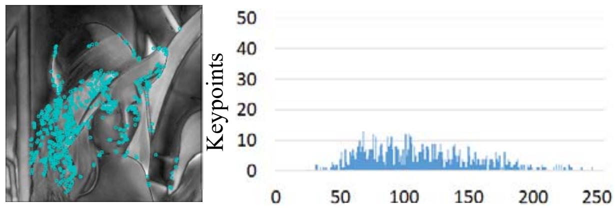
Flower has higher frequency component than Lenna, and Things has higher frequency component than Tree.

We discuss the comparison of $\bar{\phi}$ under variable $I_{Max,i}$. Referring to Table 1, 2 for VOCUS2 and Itti, $\bar{\phi}$ was large for high spatial frequency. While, for proposal, $\bar{\phi}$ is less influenced by spatial frequency change compared to conventional method.

Fig.10 (for Lenna), 11 (for Flower) shows the location of keypoints on saliency map (Left), the histogram which indicates the response of saliency at the locations respectively. There is difference in frequency component, however, for the case of the proposal, the location of the peak in the histogram is higher saliency than other saliency map. Thus the inner product in Eq. (13) becomes larger. The change of $W_{Obj,i}$ means the

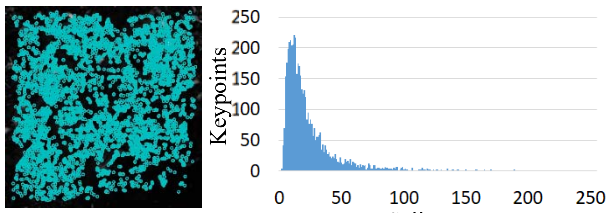


(a)Itti

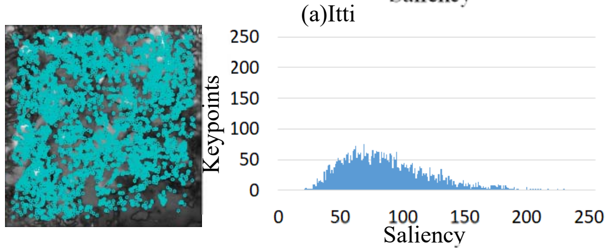


(b)Proposed method (setting2)

Fig. 10 Keypoint location (Lenna)



(a)Itti



(b)Proposed method (setting2)

Fig. 11 Keypoint location (Flower)

Table 1 Comparison of $\bar{\phi}$ (variable $I_{Max,i}$)

(a) $w = 1$ $(\boldsymbol{F_{Stb}} = \boldsymbol{r_{ftr}})$

Image	VOCUS2		Itti	Proposal	
	1/10	5/10		Set.1	Set2
Lenna	18.13	20.63	31.61	19.94	**18.04**
Flower	29.00	29.64	41.39	19.72	**18.98**
Tree	21.85	24.34	33.97	20.76	**20.10**
Things	31.24	29.77	28.36	19.60	**18.37**

(b) $w = 0$ $(\boldsymbol{F_{Stb}} = \boldsymbol{s_{dsc}})$

Image	VOCUS2		Itti	Proposal	
	1/10	5/10		Set.1	Set2
Lenna	**17.20**	18.45	33.15	20.05	17.65
Flower	28.85	29.32	42.18	19.66	**18.82**
Tree	18.97	21.53	33.90	19.24	**17.27**
Things	28.58	27.40	26.67	15.21	**13.11**

Table 2 Comparison of $\bar{\phi}$ (variable $W_{Obj,i}$)

(a) $w = 1$ $(\boldsymbol{F_{Stb}} = \boldsymbol{r_{ftr}})$

Image	VOCUS2		Itti	Proposal	
	1/10	5/10		Set1	Set2
Tree	22.62	23.22	34.06	**21.87**	22.16
Things	28.74	25.83	29.56	18.96	**17.68**

(b) $w = 0$ $(\boldsymbol{F_{Stb}} = \boldsymbol{s_{dsc}})$

Image	VOCUS2		Itti	Proposal	
	1/10	5/10		Set1	Set2
Tree	24.57	25.37	35.30	**23.24**	23.88
Things	28.30	27.06	29.74	21.52	**20.09**

image reduction, then high frequency component increases. The proposal recorded smaller $\bar{\phi}$ than others. As the results, the M_{Sal} of our proposal turned out to have larger correlation in feature stability and saliency, which means our proposal is more suitable for keypoint selection.

5. Conclusion

In this research, we proposed method of saliency map generation which consists of adaptive filter adjustment to spatial frequency to aim at preventing fluctuation of saliency caused by different input image and photographing condition change. Our saliency map method turned to be suitable for selecting keypoints less affectable by photographing condition change compared to other conventional methods.

References

1. A. A. Dalve, S. Shiravale, Real Time Traffic Signboard Detection and Recognition from Street Level Imagery for Smart Vehicle, International Journal of Computer Applications, 135(1), 2016, pp. 18-22.
2. H. Wang, X. Dong, J. Shen, X. Wu and Z. Chen, Saliency-Based Adaptive Object Extraction for Color Underwater Images Proceeding of the 2nd International Conference on Computer Science and Electronics Engineering, 2013, pp. 2651-2655.
3. L. Itti, C. Koch, and E. Niebur, A Model of Saliency-based Visual Attention for Rapid Scene Analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(11), 1998, pp. 1254-1259.
4. S. Frintrop, T. Werner, and G. M. García, Traditional Saliency Reloaded: A Good Old Model in New Shape, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 82-90.
5. M. Aly, M. Munich and P. Perona Bag of Words for Large Scale Object Recognition Properties and Benchmark Proceedings of the Sixth International Conference on Computer Vision Theory and Applications, 2011, pp. 1-8.
6. M. Brown, D. Lowe, Automatic Panoramic Image Stitching Using Invariant Features, International Journal of Computer Vision 74(1), 2007, pp. 59-73.
7. D. Lowe, Distinctive Image Features from Scale-Invariant Keypoints 60(2), 2004. pp. 91-110.
8. S. Leutenegger, M. Chili and R. Siegwart, BRISK: Binary Robust Invariant Scalable Keypoints, International Conference on Computer Vision, 2011, pp. 2548-2555.
9. P. Viola and M. Jones Rapid Object Detection Using a Boosted Cascade of Simple Features Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 511-518, 2001
10. R. Lienhart, A. Kuranov and V. Pisarevsky Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Detection, Joint Pattern Recognition Symposium DAGM 2003 Pattern Recognition, 2003, pp. 297-304.
11. R. Mochizuki, S. Yasukawa and K.Ishii Proposition of Saliency map Based on the Maximization of Center-Surround Difference, Proceedings of the International Conference on Artificial Life and Robotics, 24, 2019, pp. 487-492.
12. MATLAB Saliency –Caltech Vision- HomePage http://www.vision.caltech.edu/~harel/share/gbvs.php
13. Saliency System VOCUS2 Universität Bonn Home Page http://pages.iai.uni-bonn.de/frintrop_simone/vocus2.html
14. Caltech Database Homepage http://www.vision.caltech.edu/Image_Datasets
15. Kanagawa Institute of Technology http://www.ess.ic.kanagawa-it.ac.jp/app_images_j.html

Robot Motion and Grasping for Blindfold Handover

S. Chumkamon, K. Kawamoto, J. Inthiam, U. Yokkampon, E. Hayashi
Hayashi Laboratory, Kyushu Institute of Technology, Japan
E-mail: m-san@mmcs.mse.kyuetech.ac.jp, kawamoto@mmcs.mse.kyuetech.ac.jp, jiraphan@mmcs.mse.kyutech.ac.jp, may@mmcs.mse.kyuetech.ac.jp, haya@mse.kyutech.ac.jp

Abstract

Autonomous robots in human-robot interaction (HRI) recently are becoming part of human life as the number of service or personal robots increasingly used in our home. In order to fulfill the gap of HRI merit, we would like to propose a system of the autonomous robot motion and grasping creation for assisting the disabled person such as the blind people for handover tasks to help blind people in pick and place tasks. In this paper, we develop the robot motion to receive the object by handing over from the blindfold human to represent the blindness. To determine the target of the object and human hand, we implement the 6DOF pose detection using a point cloud and hand detection using the Single Shot Detection model in Deep learning for planning motion using 9DOF arm robot with hand. We finally experiment and evaluate the tasks from blindfold-robot handover tasks.

Keywords: Factory Automation, Motion Planning, Human-Robot Interaction.

1. Introduction

In order to push country economic, the smart automation in industrial sector is keys to improve nowadays. Since the merit of artificial intelligence research could facilitate various tasks in industrial. Therefore, we are able to develop the factory that could produce products faster with higher precision. In smart factory, the robot is important to assembly because it is more reliable and accuracy than human.

The manipulator robot actually is controlled by joint position however we usually control the robot in Cartesian space therefore we have to calculate the inverse kinematic (IK). One of popular method of IK recently is Trac-IK and KDL which can calculate the solution more than related method[1]. Since the robot also need to plan the trajectory according to the environments such as avoiding the collision or moving to the target. The recent motion planning method is Open Motion Planning Library which is popular, reliable and fast for solving the trajectory motion in various movement such as robot arm motion, vehicle, drone[2].

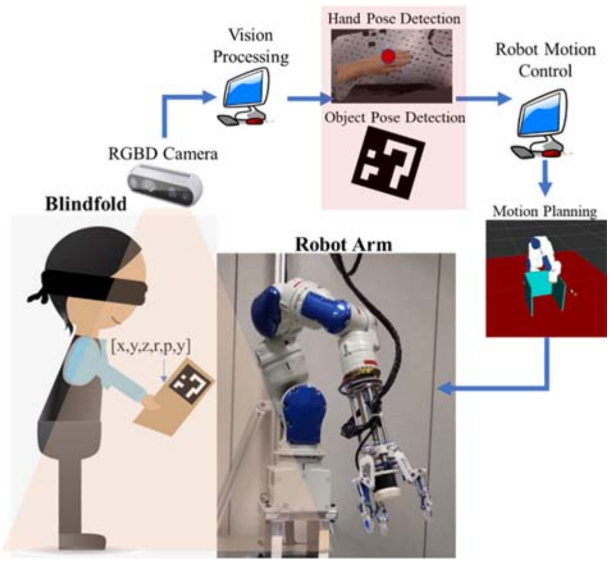


Fig. 1. The overview system of challenge experiment of handover tasks from blindfold

Since the robot could not move by itself, the user have to teach the robot to move according to the tasks. However, sometimes it is hard to teach the robot because of the difficult environment. Therefore, we proposed the teaching less system that the robot can generate motion according to the situation. We implement the perception system to let the robot can autonomously planning its motion depending on the tasks.

Collaborative robot (Cobot) becomes import recently since it helps the worker achieve the tasks easier and more reliable. There is the research on assembly tasks or Pick and Place recently[3]. However, the recent research is lack for implement the robot for disabled people therefore our application also can apply for service tasks for blind people. Since the blindness recently is lack support from the society, we therefore attempt to initiate the experiment that can apply for blindness using subject with blindfold.

In this paper, we propose the robot system that can generate motion to grasping the object from human in any poses by let the human blind their eye that the human cannot help the robot to grab the object. In addition, this system might be useful for blind people in service tasks such helping the blind people keep the thing to its place.

In the rested paper, we organize the section in 5 section including

2. Concept of Motion Generation for Handover from Blindfold

In this section, we explain the overview system including robot hardware configuration, robot perception, motion planning and grasping pose for handover.

2.1. *Hardware Configuration*

From the beginning, we have proposed the concept idea and the objective of this system. Consequently, we design the hardware configuration. Firstly, we utilize the robot arm which is industrial robot from Motoman Yaskawa as an operator for getting the object from blindfold. The robot was construct by 7 degree of freedom including one joints in base, two joints in shoulder, two joints in elbow, and two joints in wrist. For grippe, we use D-Hand robot which 2 DoF that can control finger angle and gripper state. The robot is shown in the Fig. 2.

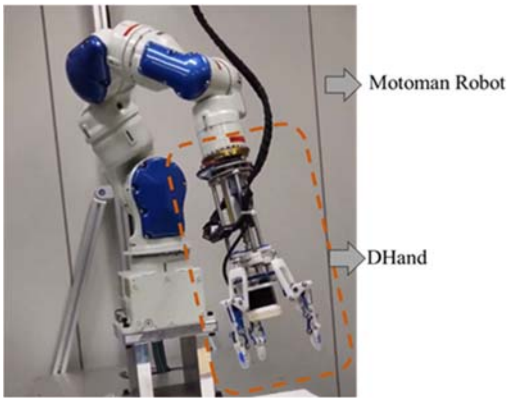


Fig. 2. The robot system and hand for grasping.

Fig. 3. Example color image and depth image from hand detection.

2.2. *Perception*

We develop perception system in order to estimate the target position for motion which is based on the hand detection and object position estimation. Firstly, we employ RGBD camera to getting image. In perception, hand detection is implemented using Deep Learning based Single Shot Detection[4] then we extract the location of hand from color image and depth image. We estimate the hand position by project the depth image in Cartesian space. Afterward we extract area of the points cloud of hand and object in the palm. We then remove the hand point by skin filter and get only object point and estimation the position. Fig 3 shows the result when the robot can detect the hand extracting in color and depth image. Fig. 4 shows the results in point cloud of position estimation of hand and object which the upper point represents object and lower point represents hand position.

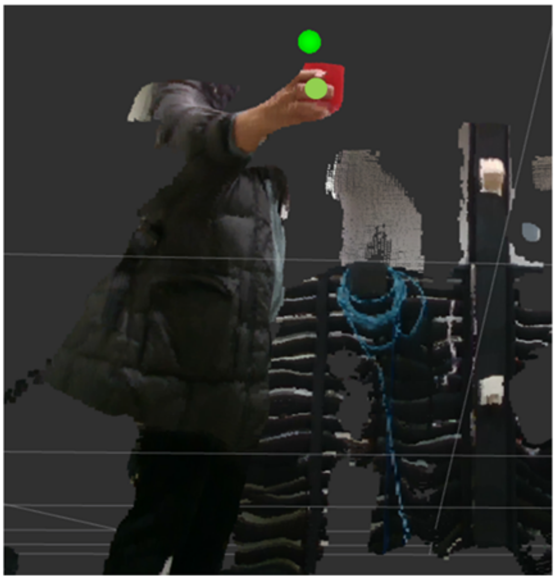


Fig. 4. 3D point cloud of hand and object position estimation for robot grasping.

2.3. *Motion planning*

Since the robot knew the position of the object, we then can plan the motion for the robot. Firstly, we set the home pose to initialize the robot joints to prepare for grasping object from blindfold. When the robot can detect hand and object then it would plan motion using OMPL and KDL inverse kinematic. In this implementation, we utilize MoveIt API for organize the motion plan and collision protection among the environment. After complete detection, then robot would approach the object prepare for grasping then move to grasp from blindfold who does know the current object pose and robot hand position that human cannot help the robot grip object easier. Therefore, our system is pretty unique among the current research that usually let human know robot hand and can help robot to grip object.

3. Experiment

From the above section, we explained our system and the configuration then in this section we explain the demonstration and evaluation of our system. Since we proposed this system for handover from blindfold in factory automation or human service sector. We then invite subject to do experiment and blind his eye to preventing subject helping robot to grasp which is unique experiment comparing to recent research and slightly challenge.

The experiment starts from the robot initializing position and wait for subject hand detection with object. In fig. 6 shows experiment snapshot during the handover from blindfold starting from T0 to T8. In T0 the robot prepares for hand detection in home pose. After the robot can detect hand then the robot moves to pre-grep position in T2. T3, the robot moves move to grasp from subject handover the object robot without seeing the robot because subject is blindfold. T4 after handover completely, the robot moves to post-grip position. Then in T5 and T6 robot move to place the bottle into the box at the pre-place position. T7 and T8, the robot put the bottle into the box softly.

In fig. 5 shows the 3D image captured during the handover experiment. The first, second and third show the side view, top view and front side view respectively.

4. Conclusion and Future Works

This paper proposed the autonomous motion planning of manipulator robot to utilize in factory automation for handover tasks. The handover tasks are important for assembly production since there are many worker and robot operating together. However, our paper also is designed to utilize for blind people as a service robot. Because this system can assist blind people for keeping things to their place. This might help blind people live easier.

In this paper, we already present the system for motion plan where the system can be employed for factory and human life such as blind people. For the perception in this system, we implement hand position estimation using deep learning to deploy for motion planning. For the motion system, we utilize The Open Motion Planning Library and KDL inverse kinematic which are capable to plan the motion to grasp the object from handover tasks. By the experiment and result, we conclude that the robot could generate the motion and grasping for handover task For the future work, we would like to the improve robust pose estimation for grasping using 6DoF pose detection. Since robot can know the object pose then the robot can grasp with the reliable position. Moreover, we would like to improve the reliable of system for real service task to assist blind people to improve human life.

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. Zhu, Z. and Hu, H., 2018. Robot learning from demonstration in robotic assembly: A survey. Robotics, 7(2), p.17.
4. Liu, Wei, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C. Berg. "SSD: Single shot multibox detector." In European conference on computer vision, pp. 21-37. Springer, Cham, 2016.

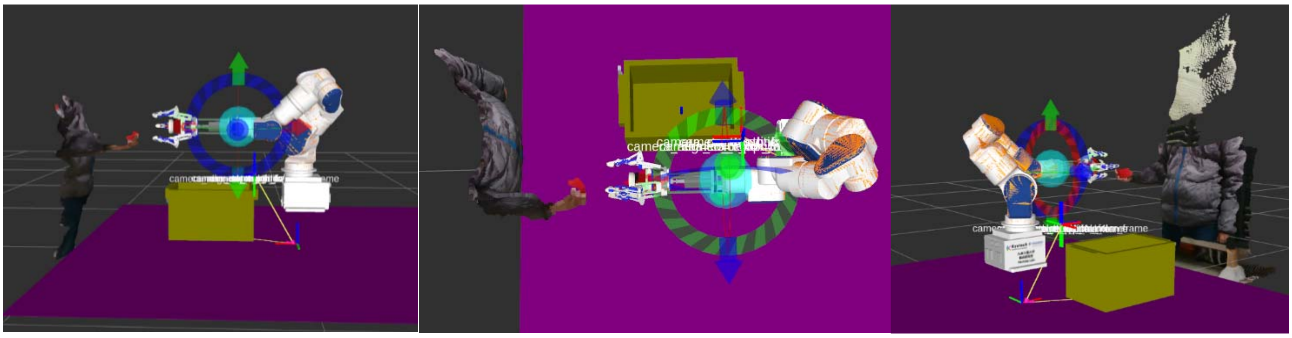


Fig. 5. 3D point cloud of environment and subject with robot model during handover experiment in different views

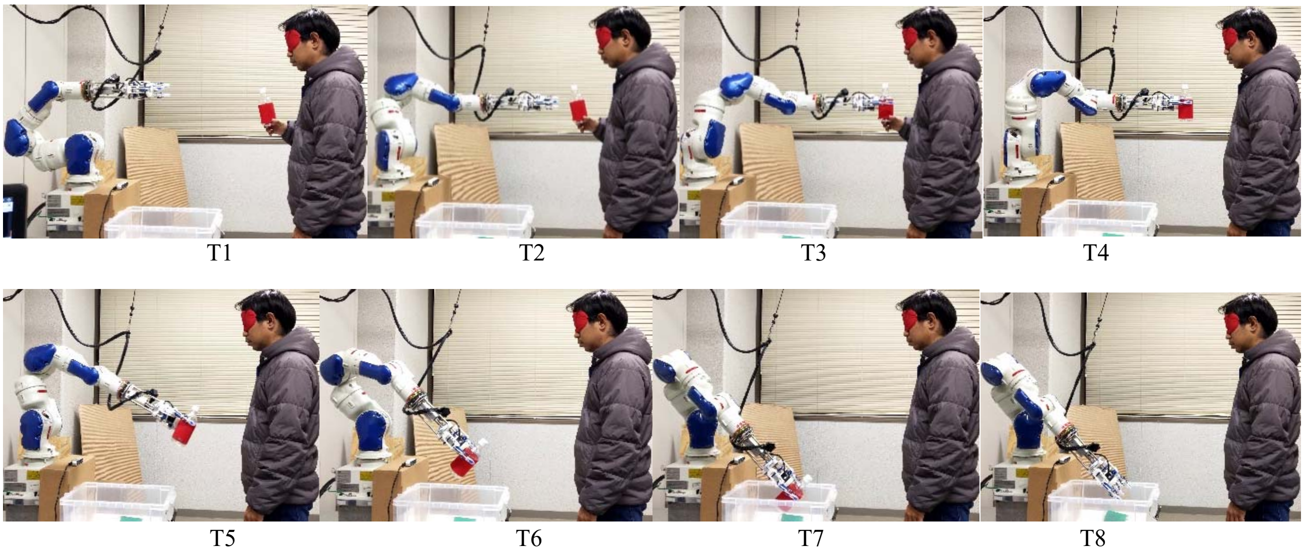


Fig. 6. The snapshots during object handover experiment from blindfold to robot

Gait Learning Method for Quadrupedal Robot Using Chaos Time-series Analysis

Yuehang Ma, Kaori Watanabe, Hidekazu Suzuki
Graduate School of Engineering, Tokyo Polytechnic University, 1583 Iiyama
Atsugi, Kanagawa 243-0297, Japan
E-mail: xmaancn@gmail.com, watanabe.kougei.karakuri@gmail.com, hsuzuki@em.t-kougei.ac.jp
www.t-kougei.ac.jp

Abstract

In the field of pet robots and robot-assisted therapy (RAT), characterization of animal motion is important for the development of robots resembling various animals. This paper presents a method for the generation of animal gait in quadrupedal robots. In this study, we employed AIBO as an experimental quadrupedal robot and generated the gait of the robot on the basis of an animal's gait and zoology. Moreover, we realized the stable gait on the ground by adjusting the minor deviation of parameters for each joint.

Keywords: Quadrupedal Robot, Gait Learning, Robot-Assisted Therapy, Genetic Algorithm

1. Introduction

Animals have for long been recognized as being a positive force in healing processes.[1] In recent years, animal-assisted therapy (AAT), which makes use of the healing effects of animals has attracted attention.[2] Examples of the expected results of this type of therapy are buffering actions for stress, improvement of sociability and shortening of the medical treatment period through mental healing. Thus, the introduction of AAT is being considered in hospitals and health facilities. However, it is difficult to employ AAT in such facilities because of the risks of the spread of infection from animals to patients and the necessity of proper animal training.

Robot-assisted therapy (RAT), in which robots resembling animals are used instead of real animals, is important for patient safety.[3] Pet robots resembling various animals, such as the dog robot as "AIBO", seal robot "Paro[4]", *etc.*, are used in this type of therapy. Banks *et al.* reported no difference between the effectiveness of a living dog and an AIBO robotic dog in reducing loneliness.[5] Shibata *et al.* applied a mental commit robot, Paro, to RAT, and they verified that the interaction with Paro has psychological, physiological and social effects on people.[6] In these applications, it is important that the robot imitates the motions of a living animal, especially essential motions, such as walking, running, *etc.*

However, it is difficult for the robot to walk and run like an animal because it is affected by various types of dynamic noise in the real world, in contrast to the ideal world. In recent years, many researchers have studied gait generation methods for various types of robots.[7] A legged robot in the real world will have n-DOF (degrees of freedom) for movement, and it is difficult to solve the optimization problem in n-dimensional continuous state/action space to generate an adequate gait.[8] Therefore, evolutionary approaches, such as use of fuzzy logic, genetic algorithms, neural networks, or various hybrid systems, are employed for gait learning and parameter optimization.[9] For example, Chernova *et al.* generated fast forward gaits using an evolutionary approach for quadruped robots.[10] However, these gait generation methods for legged robots did not evaluate the degree to which the robot's movement approximated that

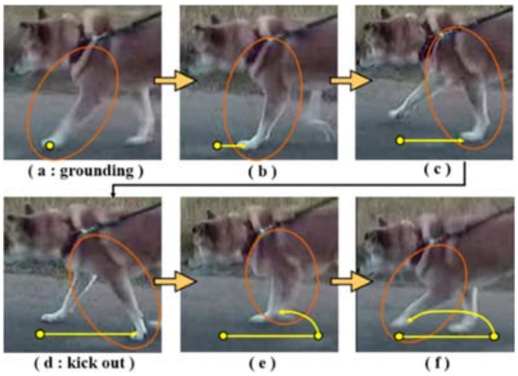


Fig. 1. Normal gait of dog

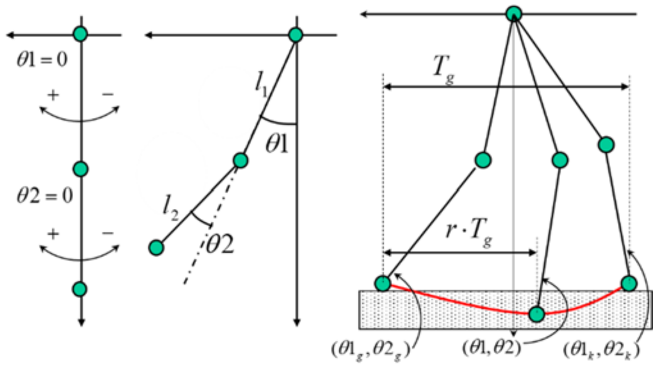


Fig. 2. $\theta 1$, $\theta 2$, r and T_g

of a living animal, because they were not designed for enhancement of the effects of RAT.

In this report, we proposed generation method of stable animal gait for quadrupedal robot. Moreover, we modified the generated gait by GA optimization using chaotic time series analysis to adapt to the dynamical interference by ground reaction.

2. Quadrupedul Gait Generation

Usually, when generating a gait for a robot, we often construct a robot model on the basis of dynamics. However, the dynamics of AIBO change in a complex manner depending on the situation in a real environment, and therefore strict modeling is difficult. Moreover, it may be even more difficult to define subjective human feelings for animals based on a model. Therefore, we generated a gait for AIBO on the basis of that of a living animal and subjective human feelings.

Figure 1 shows the normal gait of a walking dog. In this figure, (a) represents the dog's leg that is in contact with the ground and (d) represents the leg shape, which kicks out.

Here, by studying the moving image of a dog's gait, we noted that there is a turning point that changes the velocity of the leg in front and behind. It appears that the function of the leg changes from providing support to driving. The three parameters ($\theta 1$, $\theta 2$, r) represent the leg shape at this turning point, as shown in Fig. 2. And T_g, θi_g and θi_k represent the grounding time [ms], the initial shape at grounding and kicking, respectively. Hence, the intermediary orbit is uniquely decided by the parameters ($\theta 1$, $\theta 2$, r). Briefly, the problem of generating an adequate orbit for AIBO's leg is changed to the problem of optimizing the parameters ($\theta 1$, $\theta 2$, r).

In the previous study, we attempted to generate an animal gait for a quadrupedal robot using a genetic algorithm (GA)[11] and gait patterns based on zoological characteristics.[12] Moreover, a questionnaire study was performed to determine an adequate mix of several combinations of gait velocity and duty ratio for generated gait, and thus a more natural animal-like gait for the AIBO was chosen based on subjective human feelings from among the various gaits.[13] in this gait, T_g is 510 [ms] and the duty ratio is 0.63. Moreover, we have confirmed on questionnaire study that many participants feel that this AIBO gait is fairly similar to that of a living animal.

3. Modification for Dynamical Interference

We performed further experiment using the generated gait to check the adaptability to interaction with ground like that shown in Fig. 3. AIBO walked forward unsteadily, and the motion did not resemble the gait of a living animal. It seems that the generated gait cannot adapt to the dynamical interference by ground reaction.

Thus, we tried the additional optimization for generated gait to correct minor deviation of angle and timing for each joint. In this optimization experiment, the parameters ($\theta 1$, $\theta 2$, r) of each leg generated by above-mentioned process are modified slightly by GA.

The GA is viewed as an optimization method as the iterative process of evolution toward better search solutions is equivalent to the process of optimizing the fitness function.

The GA process is shown in Fig. 4. A population comprising a set of s individuals is used by the GA process to search for the target gait in the real world. As the elitist model of the GA is adopted, the best sorted individual in the N-th population, designated as a vector

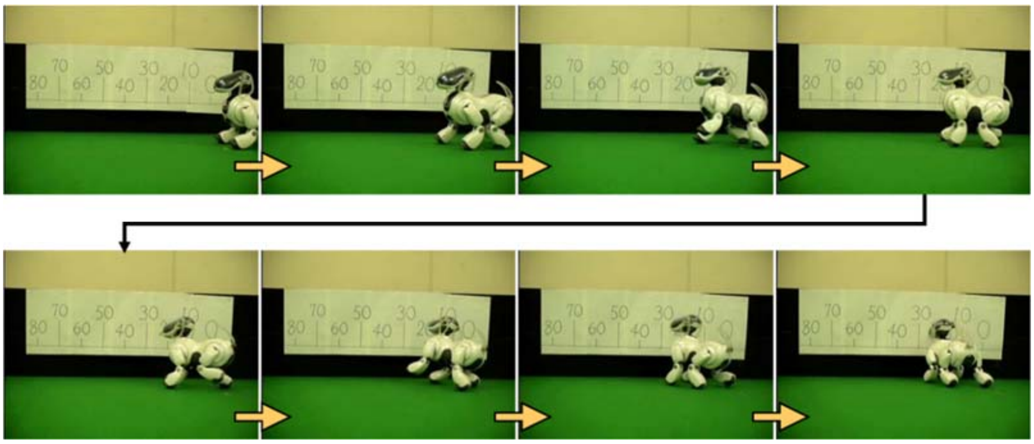


Fig. 3. Walking on ground

ϕ_l^N and possibly representing the leg's parameter, which can adapt to the dynamical interference by ground reaction is selected to survive. Let us denote the components of ϕ_l^N expressing the gait of the *l*-th individual in the *N*-th generation by $\theta1_l^N$, $\theta2_l^N$, and r_l^N. Here, ϕ_l^N is represented as following equation.

$$\begin{aligned} \phi_l^N &= (\theta1_l^N, \theta2_l^N, r_l^N) \\ &= (\theta1 + \Delta\theta1_l^N, \theta2 + \Delta\theta2_l^N, r + \Delta r_l^N) \end{aligned} \quad (1)$$

In this equation, the range of $\Delta\theta1_l^N$, $\Delta\theta2_l^N$and $\Delta r_l^N$are ± 8[∘], ±8[∘]and ±0.15, respectively.

In the optimizing process of GA, the evaluation value of recurrence plot generated by chaotic time series analysis using acceleration sensor at 5 cycles of gait is adopted as the fitness value of the GA. In the recurrence plot, the stability of gait is represented at evaluation value E (0.0 ~ 1.0). The obtained fitness values $E_1^N(\phi_1^N)$, $E_2^N(\phi_2^N)$, ⋯, $E_s^N(\phi_s^N)$ are sorted. Based on the ranking and a selection rate to die, the weakest individuals in terms of poor fitness values are replaced by newly created individuals. In creating the new individuals, random selection and random crossover are first performed. In this process, paired mates and two-point crossover are used. Next, a random bit-by-bit mutation (exchange of 1 by 0 or *vice versa*) is performed on the individuals obtained after the crossover. This ends the *N*-th generation and the population obtained after these operations constitutes the population at the starting point of the $(N+1)$-th generation. The preceding steps are then repeated with the individuals in population $N+1$ to evolve the population toward the solution.

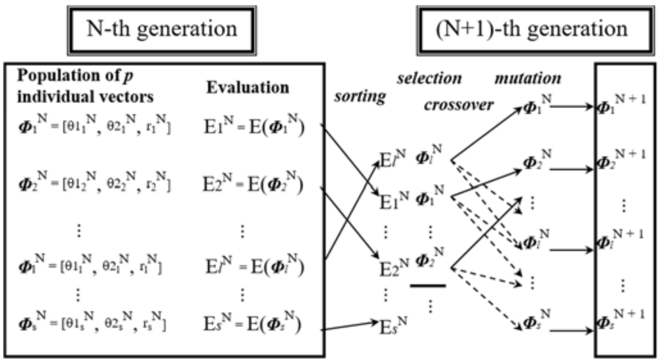


Fig. 4. Elitist model searching of a GA

Figure 5 shows the gaits of each generation in this experiment. In the first half of the optimization, AIBO walked unsteadily and diagonally. However, at the 15-th generation, AIBO walked straight ahead stably and its gait resembled that of a living dog.

4. Conclusion

In this report, we proposed generation method of stable gait for quadrupedal robot. Moreover, we modified the generated gait by additional GA optimization using chaotic time series analysis to adapt to the dynamical interference by ground reaction.

References

1. M. M. Baun, N. Bergstrom, N. Langston and L. Thoma, Physiological Effects of Human/Companion Animal Bonding, *Nursing Research*, Vol. 33, No. 3, pp. 126-129, 1984.
2. A. H. Fine, *Handbook on Animal-assisted Therapy: Theoretical Foundations and Gudelines for Practice*, Academic Pr., 2006.
3. T. Shibata, K. Wada, T. Saito and K. Tanie, Human Interactive Robot for Psychological Enrichment and

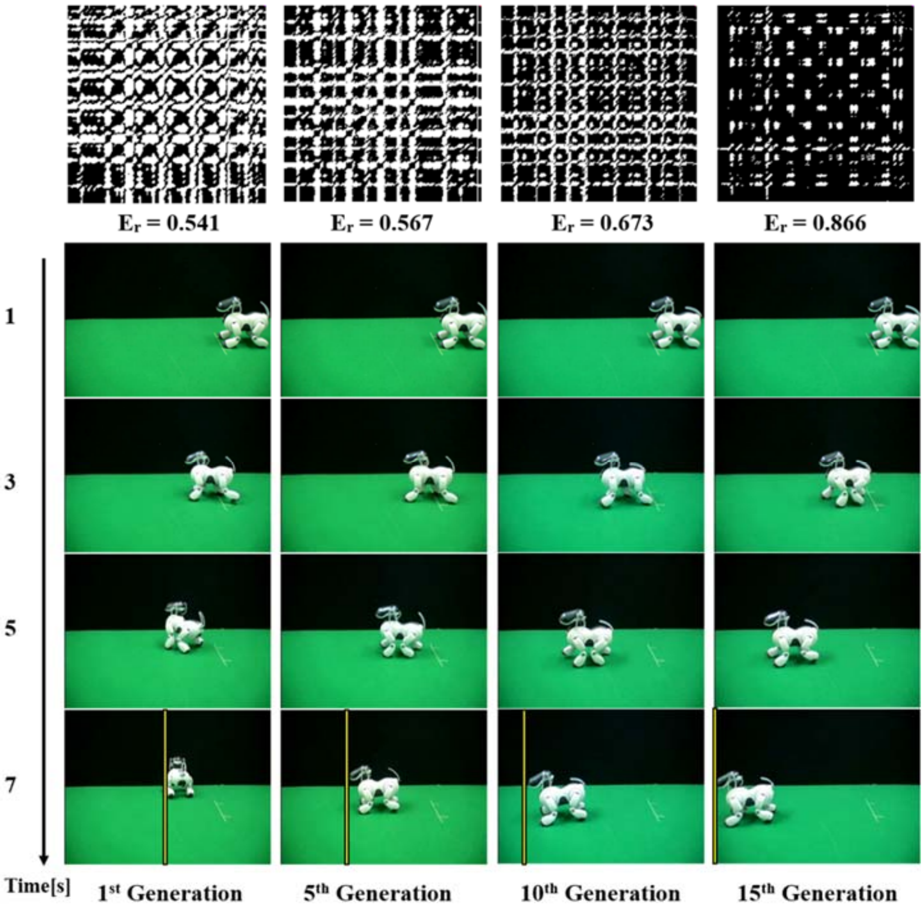


Fig. 5. Gait at Each Generations

Therapy, *Proceedings of the Symposium on Robot Companions: Hard Problems and Open Challenges in Robot-Human Interaction*, pp. 98-109, 2005.

4. http://paro.jp/english/index.html
5. M. R. Banks, L. M. Willoughby and W. A. Banks, Animal-Assisted Therapy and Loneliness in Nursing Homes: Use of Robotic versus Living Dogs, *Journal of the American Medical Directors Association*, Vol. 9, No. 3, pp. 173-177, 2008.
6. K. Wada, T. Shibata, T. Saito and K. Tanie, Effects of Robot-Assisted Activity for Elderly People and Nurses at a Day Service Center, *Proceedings of the IEEE*, Vol. 92, No. 11, pp. 1780-1788, 2004.
7. H. Kimura, Y. Fukuoka and H. Katabuti, Mechanical Design of a Quadruped "Tekken3\&4" and Navigation System Using Laser Range Sensor, *Proceedings of International Symposium on Robotics*, 2005.
8. H. Kimura, T. Yamashita and S. Kobayashi, Reinforcement Learning of Walking Behavior for a Four-legged Robot, *40th IEEE Conference on Decision and Control*, pp. 411-416, 2001.
9. H. Inada and K. Ishii, Behavior Generation of Bipedal Robot Using Central Pattern Generator (CPG) -1st Report: CPG Parameters Searching Method by Genetic Algorithm, *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2179-2184, 2003.
10. S. Chernova and M. Velosa,An Evolutionary Approach to Gait Learning for Four-Legged Robots, *Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2562-2567, 2004.
11. D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, 1989.
12. R. McN. Alexander, A. S. Jayes and R. F. Ker, Estimation of energy cost for quadrupedal running gaits, *Journal of Zoology*, vol. 190, pp. 155-192, 1980.
13. H. Suzuki and H. Nishi, Animal Gait Generation Based on Human Feeling for Quadrupedal Robot, *International Journal of Innovative Computing, Information and Control*, Vol. 4, No. 12, pp. 3341-3348, 2008.

Development of Antagonistic High Power Joint Mechanism with Cams

Katsuaki Suzuki
Life Science and Systems Engineering, Kyushu Institute of Technology
2-4 Hibikino,Wakamatsu-ku, Kitakyushu 808-0196, Fukuoka, Japan
E-mail: suzuki.katsuaki448@mail.kyuteck.jp

Yuya Nishida
Life Science and Systems Engineering, Kyushu Institute of Technology
2-4 Hibikino,Wakamatsu-ku, Kitakyushu 808-0196, Fukuoka, Japan
E-mail: ynishida@lsse.kyutech.ac.jp

Takashi Sonoda
Integrated Systems Engineering, Nishinippon Institute of Technology
1-11 Niitsu, Kamata-cho, Kyoto-gun 800-0344, Fukuoka, Japan
E-mail: sonoda@nishitech.ac.jp

Kazuo Ishii
Life Science and Systems Engineering, Kyushu Institute of Technology
2-4 Hibikino,Wakamatsu-ku, Kitakyushu 808-0196, Fukuoka, Japan
E-mail: ishii@brain.kyutech.ac.jp

Abstract

Acquiring flexible and agile behaviors as seen in biological systems, the robot can achieve acrobatic movements such as jumping and throwing. These acrobatic movements are expected to extend the range of robot activity. In this research, we propose a special mechanism using a pair of motors, springs and cams, which has three functions: normal operation, instantaneous operation and variable rigidity. Then, we derived a mathematical model of the mechanism and discussed the input-output characteristics of the mechanism by changing the design parameters, and analyzed the difference between the theoretical and measured results.

Keywords: high power joint mechanism, Link mechanism, Cam mechanism

1. Introduction

In order to perform high-speed operations such as jumping and throwing, an actuator that can realize two characteristics, strong force and high speed, is required. As a means to realize instantaneous operation, a mechanism that stores energy in a relatively lightweight mechanical element such as a spring and releases it to extract instantaneous force is often studied. Yamada has developed a closed-loop flexible catapult using jump buckling, and has realized a robot capable of jumping and swimming [1]. Although instantaneous force is generated by jump buckling, the output value is constant, and it is difficult to realize normal operation with low speed and high torque. To create an actuator that can generate instantaneous forces, the authors have developed a high-power joint mechanism that mimics the hind limb of locusts [2]. The developed mechanism can realize not only instantaneous operation but also normal operation that can be applied to walking etc. However, instantaneous operation is possible, but viscous friction loss related to the actuator reducer is large due to the structure of the mechanism. In this study, we propose a mechanism that uses a cam to eliminate this viscous loss and further enhance the function of the mechanism. In the next chapter, we describe the newly developed high-power joint mechanism.

2. Overview of the proposed mechanism

Fig.1 shows the proposed mechanism. The mechanism consists of a pair of motors, cams, cam follower, wire, spring, a slider, a linear guide, and an output link. The feature of the mechanism is that it has three functions: normal operation, instantaneous operation, and change in joint stiffness. In normal motion, the output link moves at the same output as the mounted motor. By rotating two cams on the same axis as the motor in the same direction, the output link moves. In the rapid motion, the output link can instantaneously exhibit more output than the mounted motor. By rotating the two cams in different directions, the spring can be charged with energy. Then, after the cam position reaches the release point, the cam follower disengages from the cam, releasing the stored energy with theoretically no loss. A Cam Charger developed by the authors is applied to this release mechanism[3]. After release, the slider translates using the force stored in the spring as the thrust. The translation of the slider is translated into rotational movement of the output link through the wire. This allows for agile rotational movement. For variable stiffness, it is possible to change the stiffness of the output link main joint. The authors have developed a variable stiffness joint mechanism that utilizes the characteristics of the cam's reduction ratio[4]. The proposed mechanism has a structure in which the rigidity of the joint changes according to the amount of displacement of the spring. This is possible by using two cams to compress the spring into place and keep the cam angle constant.

3. Torque and angular velocity of the output link

From the geometric model, derive an expression for the output value of the mechanism. First, from a static viewpoint, consider the torque generated in the output link. Fig. 1 (d) shows the link vector of the proposed mechanism during rapid motion. The spring force $\boldsymbol{F_{sp}}$ generated in the slider is determined by the spring change amount $\boldsymbol{\delta_{sp}}$ and the spring constant k_{sp}, and is expressed by the following equation.

$$\boldsymbol{\delta_{sp}} = \boldsymbol{l_{sp\_max}} - \boldsymbol{l_{sp}} \tag{1}$$

$$\boldsymbol{F_{sp}} = k_{sp}\boldsymbol{\delta_{sp}} \tag{2}$$

The force F_w generated on the wire is expressed by the following equation, when the angle of the wire is defined as θ_w.

$$F_w = \frac{|\boldsymbol{F_{sp}}|}{cos\theta_w} \tag{3}$$

By using the Jacobi matrix $\boldsymbol{J_{sp}} = (\boldsymbol{l_{l1}} \times \boldsymbol{e_z})$ derived by the cross product of the force $\boldsymbol{F_w}$ and the rotation direction vector $\boldsymbol{e_z}$, the torque T_o generated in the output link is expressed by the following equation.

$$T_o = \boldsymbol{J_{sp}^T F_w} \tag{4}$$

The reduction ratio, which is the input - output relationship when converting the spring force $\boldsymbol{F_{sp}}$ to the output link torque T_o, is expressed by the following equation.

$$G = \frac{T_o}{|\boldsymbol{F_{sp}}|} \tag{5}$$

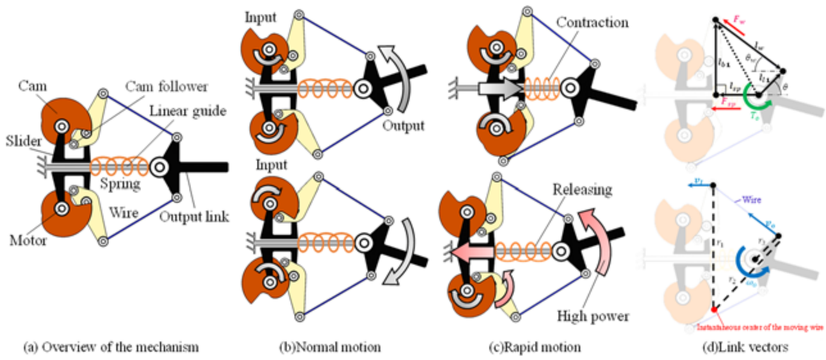


Fig.1 Cam competitive mechanism

In rapid motion,the moving mechanical elements are a slider, two wires, and an output link. Among the moving mechanical elements, the mass and inertia of the slider and output link are so large that they cannot be ignored. Therefore, it is necessary to consider the loss of kinetic energy and derive the velocity that occurs on the output link. If the mass of the slider is m_l, the velocity is v_l, the inertia of the output link is I_o, the angular velocity is ω_o, and the energy completely stored in the spring is E, the instantaneous energy change during operation is as follows.

$$\frac{1}{2}m_l {v_l}^2 + \frac{1}{2}I_o {\omega_o}^2 = E - \frac{1}{2}k_{sp}\left|\boldsymbol{\delta_{sp}}\right|^2 \quad (6)$$

Considering the instant center of rotation of wire, it is possible to define the position vectorsl_1, l_2, l_3 shown in Fig. 1 (d), and the relationship between the angular velocity ω_o of the output link and the velocity v_l of the slider is Expression.

$$v_l = \frac{r_1 r_3}{r_2}\omega_o \quad (7)$$

From (6) and (7), the angular velocity of the output link can be expressed by the following equation.

$$\omega_o = \sqrt{\frac{2E - kx^2}{m_l\left(\frac{r_1 r_3}{r_2}\right)^2 + I_o}} \quad (8)$$

4. Analysis of output characteristics

We analyzed how the output characteristics of the mechanism change when the design parameters of the mechanism change. The output characteristics were analyzed using equations (1) and (2). Fig.2 shows the design parameters changed during the analysis and the range of the numerical values. As a result of analysis, several reduction ratios and output torques were classified into seven patterns. For Pattern 1, the reduction ratio is the largest at startup, and gradually decreases with the displacement of the output link. This property is useful when the proposed mechanism is applied to arm joints for throwing. This is because when throwing, the most torque is required during acceleration. Patterns 4, 5, and 6 have a characteristic that the reduction ratio gradually increases, so that continuous torque can be output. These properties can be useful for instantaneous movements in various underwater environments. Fluid resistance may be dominated by inertial force depending on the shape of the output rigid body. In this case, since the fluid resistance approaches a magnitude proportional to the square of the speed, the output unit needs torque even outside the acceleration range.

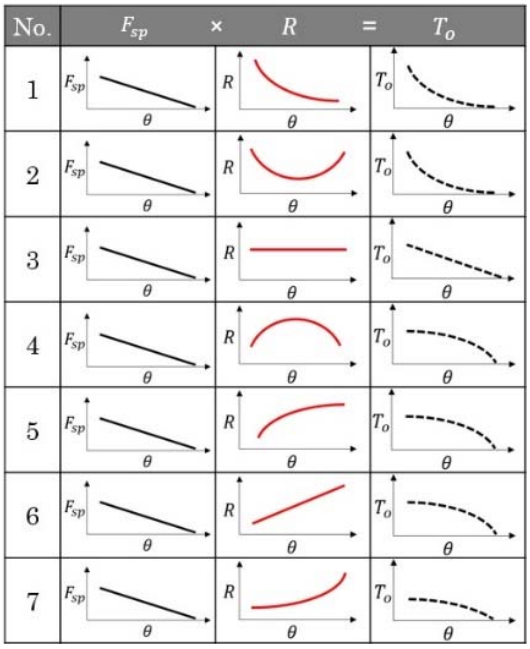


Fig.2 Pattern of the reduction ratio

5. Experimental device

An experimental device was fabricated to evaluate the proposed mechanism on a real machine. Fig. 3 shows the external view of the experimental device, Fig. 4 shows the link vector defined, and Table 1 shows the specifications of the device. k_{sp} can be changed by changing the number of mounted springs.

Table1. Spec of the device

Symbol	Value	Unit
$\boldsymbol{l_{l1}}$	59.55	[mm]
$\boldsymbol{l_{w1}}$	215	[mm]
$\boldsymbol{l_{a1}}$	80	[mm]
$\boldsymbol{l_{b1}}$	107.6	[mm]
$\boldsymbol{l_{c1}}$	28.5	[mm]
$\boldsymbol{l_{d1}}$	11.96	[mm]
$\boldsymbol{l_{sp\_min}}$	164.45	[mm]
$\boldsymbol{l_{sp\_max}}$	211.09	[mm]
$\boldsymbol{\delta_{sp}}$	46.64	[mm]
$\emptyset$	57.11	[deg]
θ_{max}	+23	[deg]
θ_{min}	-23	[deg]
K_{sp}	1.905 or 3.81	[N/mm]
m_l	1.026	[kg]
I_o	0.00598	[kgm^2]

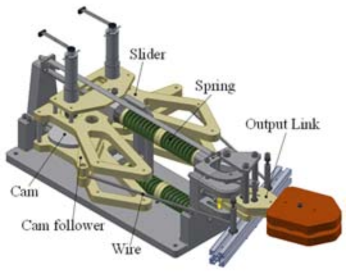


Fig.3 Experimental device

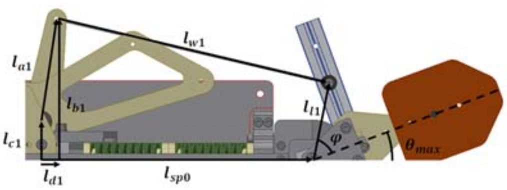


Fig.4 Link vectors of device

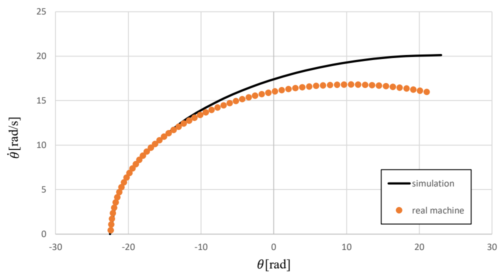


Fig.5 Angular velocity of the output link ($k_{sp} = 1.905$)

6. Performance evaluation test

In order to compare the simulation and the real machine of the angular velocity, the performance was evaluated using an experimental device. The actually measured angular velocity was calculated based on a photographed image of the operation of the output unit in the experimental apparatus taken from above. In the case of $k_{sp} = 1.905$, it can be confirmed from Fig. 5 that the angular velocity between the theoretical value and the measured value is almost the same when θ is around -25 [deg] to 0 [deg]. When θ is 0 [deg] to 25 [deg], the value of the real machine is slightly smaller than the value of the simulation. This factor is considered to be a loss due to viscous friction in the actuator, which is not reflected in the simulation.

In the case of $k_{sp} = 3.81$, from Fig. 6, it can be confirmed that when θ is 0 [deg] to 25 [deg], value of the simulation is greatly reduced with respect to the real machine. This is considered to be because the increase in the speed of the mechanism increased the viscous friction resistance as compared with the case of k_sp=1.905.

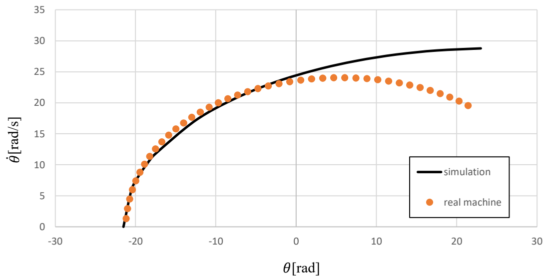


Fig.6 Angular velocity of the output link ($k_{sp} = 3.81$)

7. Conclusion

In this study, we proposed a special mechanism with three functions and discussed the mechanisms to realize these three functions. Next, we constructed a mathematical model for torque and angular velocity during instantaneous operation, and showed that the mechanism can realize some output characteristics by changing design parameters. In addition, we fabricated an experimental device for the mechanism and compared theoretical and measured values of angular velocity during instantaneous operation. The error between the theoretical and measured values was small, confirming the validity of the mathematical model and the actual machine.

As future work, we will discuss normal operation and the function to change the rigidity of the output link. Furthermore, we plan to develop a high-performance hand as one of the applications in the mechanism.

References

1. GENERATOR, ROBOT POWERED BY IMPULSE FORCE. "Basic Performance of Flat Kick-swimming Robot Powered by Impulse Force Generator with Spatially-deformable V-shape Elastica." Adaptive Mobile Robotics: Proceedings of the 15th International Conference on Climbing and Walking Robots and the Support Technologies for Mobile Machines, Baltimore, USA, 23-26 July, 2012. World Scientific, 2012.
2. Nishida, Yuya, Takashi Sonoda, and Kazuo Ishii. "Jacobian Matrix Derived from Cross Product and its Application into High Power Joint Mechanism Analysis." Journal of Bionic Engineering 7 (2010): S218-S223.
3. Sonoda, Takashi, et al. "Development of antagonistic wire-driven joint employing kinematic transmission mechanism." Journal of Automation Mobile Robotics and Intelligent Systems 4 (2010): 62-70.
4. A.A.F Nassiraei, et al., Realization of the Rapid Movements for the Entainment Robots by Using Two New Actuators"Inertia Actuator" &"Cam Charger", Proc. ASME IMECE2006, IMECE2006-14257, 2006

Towards fully automated driving in urban areas

Naoki Suganuma

Institute for Frontier Science Initiative, Kanazawa University
Kakuma-machi, Kanazawa, Ishikawa, 9201192, Japan[1]
E-mail: suganuma@staff.kanazawa-u.ac.jp

Abstract

In recent years, research on autonomous driving technology has been actively conducted. In our laboratory, research and development on autonomous vehicle started from 1998 aiming for autonomous driving in urban areas. Additionally, our laboratory started the first public road test among Japanese University in Japan from 2015, and it has already been conducted public road experiment over 14,000 km. In this presentation, we will outline the technologies required for autonomous driving based on the knowledge obtained from these achievements and describe the current state and issues of autonomous driving technologies.

Keywords: Autonomous vehicle, public road experiment, perception, motion planning.

1. Introduction

In recent years, technology related to the autonomous vehicles attracted many attentions from all over the world. Especially, after successful development [1] at DARPA Urban Challenge held at USA in 2007, many technologies have been developed and some companies and universities have been conducted public road testing of their technologies. It is believed that the technologies of autonomous vehicle will solve traffic problems such as traffic accidents, lack of public transportation, and so on.
In our laboratory, we started researches related to the autonomous vehicle for urban driving at 1998, we have already conducted the researches for more than 20 years. Moreover, from February 2015, we have conducted public road testing of autonomous vehicle technologies for the first time among Japanese universities. Moreover, our autonomous vehicle has already been driven more than 1,000km at public road. In this paper, overview of the autonomous vehicle developed by our laboratory is introduced, and current status of our activities is reported.

2. Experimental vehicle

Figure 1 shows our experimental vehicle used for public road testing. In this vehicle, in order to percept under complex environment like urban road, many sensors are equipped. On the roof, LIDARs (One velodyne VLS-128 and two velodyne VLP-16), GNSS/INS (Applanix

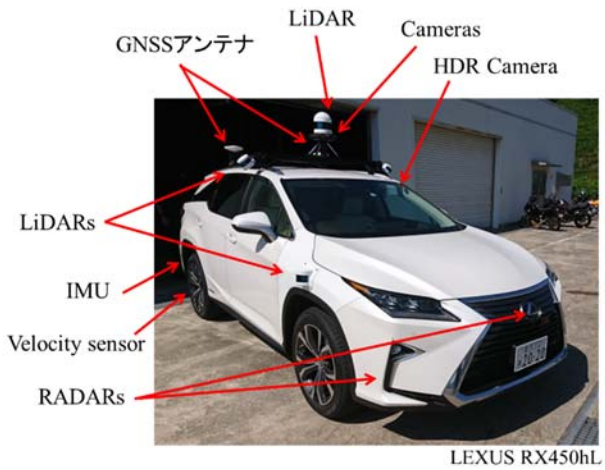


Figure 1 Sensors equipped with test vehicle

POS-LV220), and cameras (six sekonix AR0231) are installed. These sensors are used to not only detect surrounding objects [2] and obstacles as shown in figure 2 but also localize ego-vehicle position and orientation precisely [3] as shown in figure 3. Inside the bumber of the vehicle, two LiDARs (two Velodyne VLP-32) and nine millimeter-wavelength RADARs (DENSO-TEN) are installed and these sensors are used to mainly detect blind spot object and distant objects, respectively. Moreover, inside the vehicle, a color high dynamic range cameras are installed and used to recognizing traffic lights [4] from distant as shown in figure 4.

3. Public road testing of autonomous vehicle

We started public road testing of autonomous vehicle from February 24th, 2015 and it has passed for about five years at this moment. Currently, our public road testing is conducted mainly at Odaiba area at Tokyo metropolitan, Kanazawa and Suzu city area at Ishilawa prefecture, and an area around Abashiri city at Hokkaido prefecture in Japan. In Suzu city, the problem of driving situation of country road is tested, and use as a public transport is considered. In Odaiba and Kanazawa city areas, as shown in figure 5, problems of driving at urban area where many vehicles are driving is considered. At abashiri area, snow condition at winter is considered.

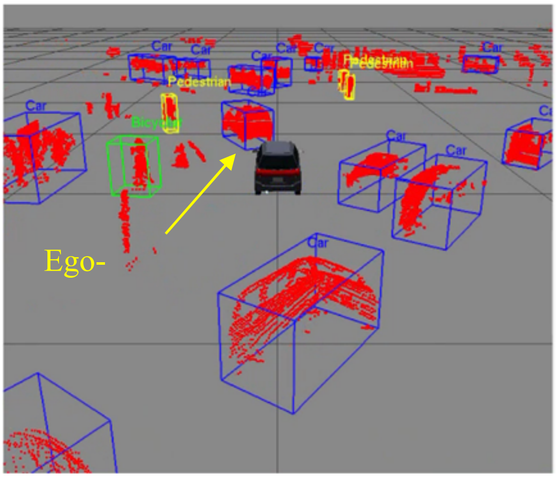


Figure 2 Recognition of surrounding objects

The total driving distance of autonomous driving at public road reaches more than 14,000. We are planning to experiment on a wide variety of roads in Japan aiming at solving various technical problems using five test vehicles.

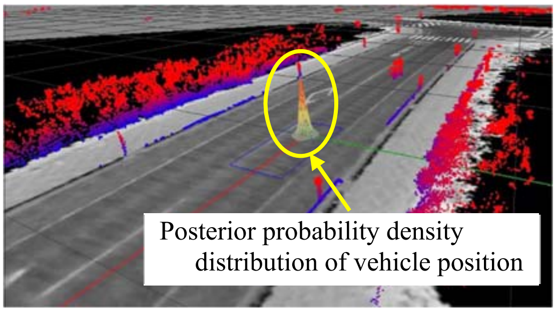


Figure 3 Localization of ego-vehicle position

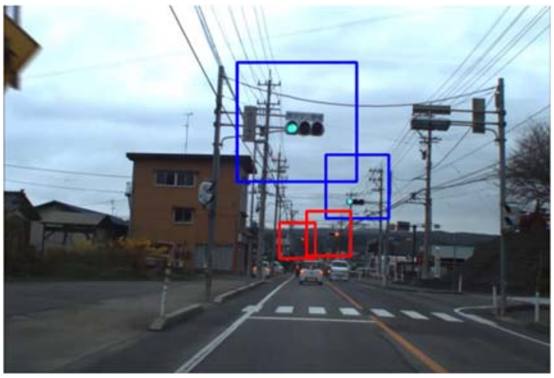


Figure 4 Recognition of traffic lights

(a) In case of driving at multiple lane

(b) In case of making right turn at intersection

Figure 5 Autonomous driving at urban road in Kanazawa city

4. Conclusions

In this paper, overview of the autonomous vehicle developed by our laboratory was introduced, and current status of our activities was reported. We do believe that the technology of autonomous vehicle will solve many traffic problems. Especially, in Japan, aging society and lack of public transportation have been become real problems. Therefore, the author hopes that technologies of autonomous vehicle developed by many researchers will be used at real society in near future and many traffic problems will be solved.

References

1. M.Montemerlo, et al., Junior: The Stanford Entry in the Urban Challenge, *Journal of Field Robotics*, Vol.25, No.9, pp.569-597, 2008
2. N.Suganuma, et al., LIDAR-based Object Classification for Autonomous Driving on Urban Roads, *Journal of Advanced Control, Automation and Robotics*, Vol.3, No.2, pp.92-95, 2017
3. N.Suganuma, et al., Localization for Autonomous Vehicle on Urban Roads, *Journal of Advanced Control, Automation and Robotics*, Vol.1, No.1, pp.47-53, 2015
4. K.Yoneda, et al., Traffic Signal and Arrow Detection using Digital Map and Luminance Enhancing, *Proc. of the 1st International Conference on Digital Practice for Science, Technology, Education, and Management*, pp124-129, 2018

A Structural Language for Multilevel Dynamics in the Design of Robot Soccer Systems

Jeffrey Johnson
Engineering and Innovation, The Open University, Walton Hall,
Milton Keynes, MK7 6AA, UK

Ruggero Rossi
Engineering and Innovation, The Open University, Walton Hall,
Milton Keynes, MK7 6AA, UK
E-mail: jeff.johnson@open.ac.uk, ruggero.rossi@open.ac.uk.
http://www.complexityanddesign.org

Abstract

Relational structure is ubiquitous in complex systems but very hard to implement in machines. Traditionally relational structures were hand-crafted using logic-based methods including various relational approaches to pattern recognition. Today the hope is that machines will find relational structures automatically by techniques of deep learning. Both approaches require new methods for representing relational structure for dynamic complex multilevel systems. We use the platform of robot soccer to investigate these ideas. This paper follows a previous paper which presented new dynamic structures for evolving tactics and strategies in team robotics. Here the notation is extended to include structures of structures of structures. For example a red defender robot r_1 may closely mark a blue attacker robot b_2 to create a structure $< r1, b_2; R_{closely\_mark}>$. This may be part of another structure $< < r_1, b_2; R_{closely\_mark}>, b_3; R_{defenders\_dilemma} >$ as another robot b_3 joins in to change the relational structure. This approach is illustrated by a RoboCup simulation game. Our next step is to build a competitive player to show that the ideas are operational and may give tactical and strategic advantages.

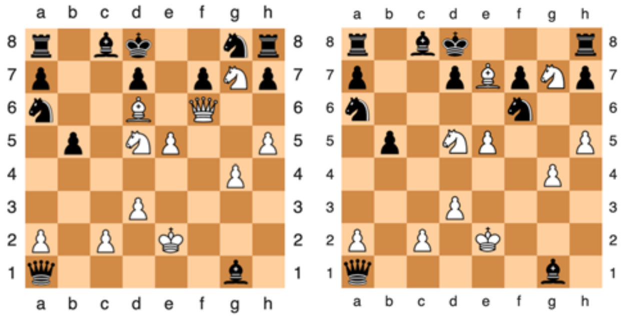
Keywords: hypernetworks, hypergraphs, connectivity, robot soccer, design, multilevel dynamics.

1. Introduction

A previous paper presented new dynamic structures for evolving tactics and strategies in team robotics[1]. The motivation for this research is to develop a coherent methodology for the planning, design, management and control of complex socio-technical systems such as cities, hospitals, airlines and banks, and to formulate socio-economic policy at local, national and international levels.
Team robotics provides an excellent laboratory subject for complex systems research since agent interaction can be studied 'from the outside' which avoids the complication of reflexivity when humans study human systems 'from the inside'.

The challenge of robot soccer can be simply stated as "By the middle of the 21st century, a team of fully autonomous humanoid robot soccer players shall win a soccer game, complying with the official rules of FIFA, against the winner of the most recent World Cup.[2] ". The RoboCup Simulation League used for illustration in this paper provides an excellent international platform for complex systems research.

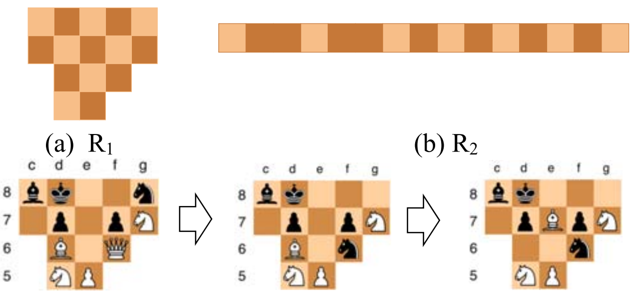
Our approach to robot soccer follows Atkin's method of analysing chess[3]. For example, Figure 1 shows the final moves in the 'Immortal Game' between Anderssen and Kierseritzky held on 21st June 1851. In this remarkable game white sacrifices most of it major pieces including the Queen on square f_6 (Fig. 1(a)). However when the black knight takes the queen the white bishop moves to square e_7 for checkmate.



(a) after Qf6+ (b) after Nxf6 and Be7#

Fig. 1. The final moves in the 1851 Immortal Game (source: https://en.wikipedia.org/wiki/Immortal_Game)

Although there are sixty-four squares on the board, only <d_5, e_5, d_6, e_6, f_6, c_7, d_7, e_7, f_7, g_7, c_8, d_8, e_8, f_8, g_8> play a part in the checkmate. We enclose them in the angular brackets < and > to show that they form a *structure*. This can be made more precise by making explicit the relation R_1 that assembles them (Fig. 2(a)). Figure 2(b) shows the squares assembled by a hypothetical relation, R_2, to form a different structure – a *row* of squares.



(a) R_1 (b) R_2

(c) R_3 R_4 R_5

Fig. 2. (a) Squares d_5, e_5, d_6, e_6, f_6, c_7, d_7, e_7, f_7, g_7, c_8, d_8, e_8, f_8, g_8 assembled by relation R_1, (b) these squares assembled by R2 into a linear structure, (c) The squares and pieces assembled into structures by the relations R_3, R_4 and R_5.

The two structures in Figure 2(a) and 2(b) can be represented symbolically as:

< d_5, e_5, d_6, e_6, f_6, c_7, d_7, e_7, f_7, g_7, c_8, d_8, e_8, f_8, g_8; R_1> and
< d_5, e_5, d_6, e_6, f_6, c_7, d_7, e_7, f_7, g_7, c_8, d_8, e_8, f_8, g_8; R_2>

As a technicality an expression of the form <a, b, c, d> is called a *simplex* and the elements a, b, c, and d are said to be its *vertices*. The expression <a, b, c, d; R> is called a *hypersimplex* since it gives not just a list of vertices but also specifies the way these elements are to be assembled into a structure – *the relational structure* made explicit by the symbol R. A collection of hypersimplices is called a *hypernetwork* [4].

In much network theory the relational structure is implicit. A major methodological requirement of hypernetwork theory is that relations must be explicit [4]. This is done by listing the vertices followed by the semicolon symbol and one or more symbols to represent the relation. E.g. Figure 2(c) includes the structure <<kNight, g_8; $R_{occupies}$>, <Queen, f_6; $R_{occupies}$>; $R_{attacks}$> meaning that the black knight can take the white queen. Black has no choice but to accept this sacrifice (Fig. 2(b)) enabling the white bishop to move to f_6 to form the structure

< <K, d_8; $R_{occupies}$>, <B, f_7; $R_{occupies}$>, <N, g_7; $R_{occupies}$>, <N, d_5; $R_{occupies}$>; $R_{checkmate}$>.

This example illustrates the multilevel nature of the representation. <kNight, g_8; $R_{occupies}$> means the structure formed by combining kNight and g_8 by the occupation relation. This structure exists independently of anything else. As a whole, <kNight, g_8; $R_{occupies}$> exists at a higher level of assembly to its parts. If the parts are said to exist at *Level N*, then the assembly <kNight, g_8; $R_{occupies}$> exists at *Level N+1*. Similarly the expression

<<kNight, g_8; $R_{occupies}$>, <Queen, f_6; $R_{occupies}$>; $R_{attacks}$>

exists at level N+2 since it assembles two *Level N+1* structures.

Even systems with small number of elements can have astronomic numbers of relational combinations. The challenge is formulate a way to represent the system in a parsimonious way. One way to do this is to *name* structures, *e.g.* let <kNight, g_8; $R_{occupies}$> = BNg8. Then BNg8 is a *Level N+*1 structure formed from the *Level N* black knight and the square g_8. <BNg8, WQf6; $R_{attacks}$> is a Level N+2 structure formed from the two *Level N+1* structures BNg8 and WQf6. This way of forming structures provides a rich structural language for the multidimensional dynamics of complex systems.

2. Fundamental elements and relationships

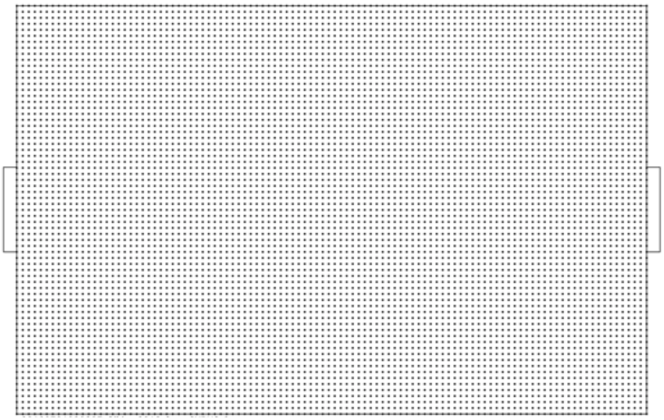
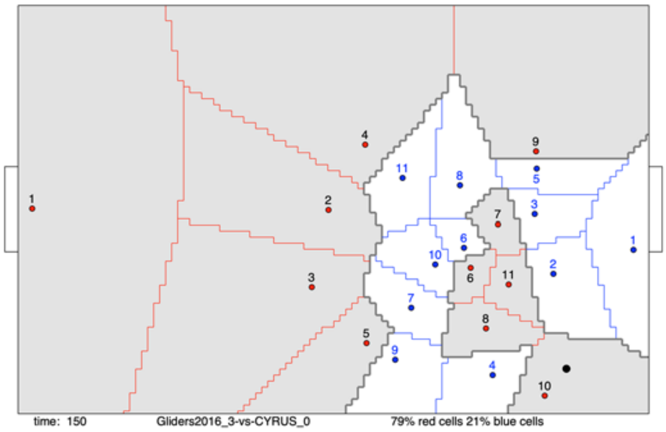
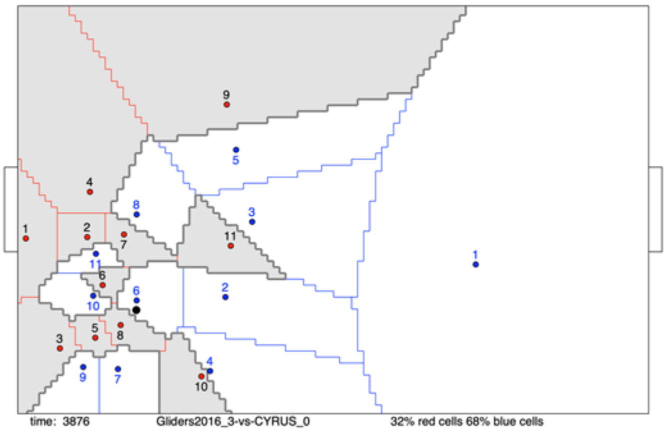


Fig. 3. The centres of the 105 x 68 soccer pitch cells



(a) the red team occupies 79% of the pitch



(b) the blue team occupies 68% of the pitch.

Fig. 4. Areas of the pitch as emergent features

The soccer pitch is represented by a set of discrete square cells. The centres of the cells are shown in Fig. 3. The cells are *fundamental elements* of our soccer system. Others include the ball and the sets of eleven blue and eleven red soccer players.

The *fundamental relationships* include the spatial relationships between the players and the ball, and the closeness of the players and the ball to the centres of the cells. In contrast to the static ranks and files of chess, the salient areas of the pitch emerge and change rapidly in robot soccer as the players and the ball move.

The cells closest to the players form polygons (Figs 4(a) and (b). The red team plays from left to right. The cells closest to each red player form a grey polygon with a red boundary. The polygons for the blue players are white bounded by blue lines. The ball is shown in black.

The occupation of the pitch depends on the positions of the players, *e.g.* in Figure 4(a) the red team occupies 79% of the pitch while in Figure 4(b) the blue team occupies 68% of the pitch. In general it is better for a team to occupy as much of the pitch as possible, and winning teams usually occupy the majority of the pitch for the majority of the game[5].

Figure 5 shows that in this game the red team occupied the majority of the pitch with a much greater frequency that the blue team. Red won by three goals to nil.

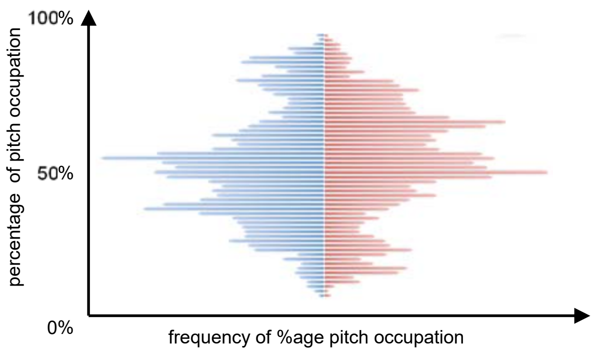


Fig. 5. During the game the red winning team occupies the majority of the pitch cells significantly more frequently than the losing blue team.

3. Analysing a robot soccer game.

Of course, it is not just pitch occupancy that matters – it is the way the players position themselves to form favourable structures. It also depends on the position and motion of the ball.

This section concerns a robot soccer game between the Gliders2016 and CYRUS teams. The data set for this consists of the x-y positions of all the players and the ball for 6000 one-tenth second ticks of the clock – games last for ten minutes. Figures 6 to 14 show a remarkable sequence of structural development that is a precursor to the red team scoring a goal. It is characterised by the creation of 'islands' of ownership and passes between the islands creating an irresistible structure from which the red team scores a goal.

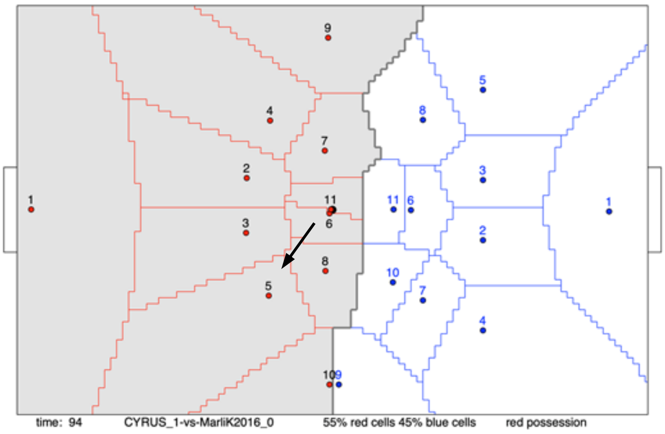


Fig. 6. time 94: < <r11, ball; ownership>, r_5; pass>

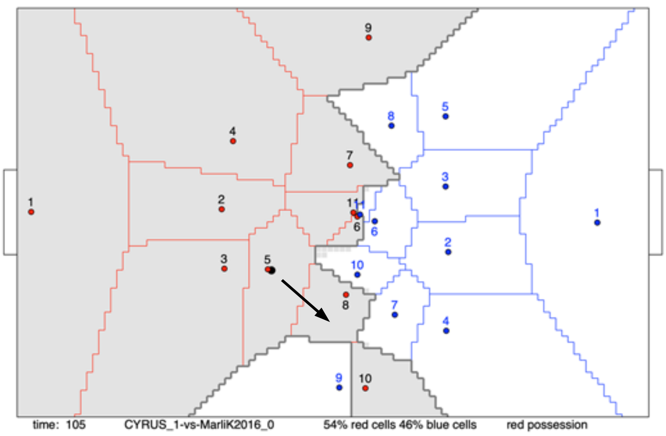


Fig 7. time 105: $<<r_5, \text{ball}; R_{\text{ownership}}>, r_8; R_{\text{pass}}>$

In Figure 6 red player number 6, denoted r_6, is at the centre for the kick-off. It possesses the ball to form the structure $<r_6, \text{ball}; R_{\text{possession}}>$. It passes to r_{11} to form the structure $< <r_6, \text{ball}; R_{\text{possession}}>, r_{11}; R_{\text{pass}}>$. This is followed by $< <r_{11}, \text{ball}; R_{\text{possession}}>, r_5; R_{\text{pass}}>$.

Two important changes occur between times 94 and 105 when r_5 receives the ball. The first is that r_6 and r_{11} move forward towards b_6 and b_{11} to form the hypersimplex $<r_6, r_{11}, b_6, b_{11}; R_{\text{close\_together}} >$ on the forward edge of red's space (Figs 6 and 7). The second is that r_{10} moves past b_9 creating an island around b_9 in red's half but also giving r_{10} a significant near-island in blue's half. At first sight this might seem to be even trade-off, but it is not since red has the initiative and the ball is moving into blue's half. Already the blue team has suffered a serious structural weakness (Figs 7 & 8).

Player r_8 receives the ball at time 114 to form the hypersimplex $<r_8, \text{ball}; R_{\text{ownership}}>$. r_{10} has moved forward to create a detached red island in blue's half, enabling the structure $<<r_8, \text{ball}; R_{\text{ownership}}>, r10; R_{\text{pass}}>$. This is a bad position with poor prospects for blue.

What happens off the ball between times 114 and 127 at the centre of the pitch is quite remarkable – b_{11} moves *towards* the red goal. This inexplicably bad move enables red to form the structure $<r_6, r_7, r_{11}; R_{\text{island}}>$ which is very strong for red and very weak for blue. At the same time b_6 hardly moves as the red island moves past it towards the blue goal (Figs 8 and 9).

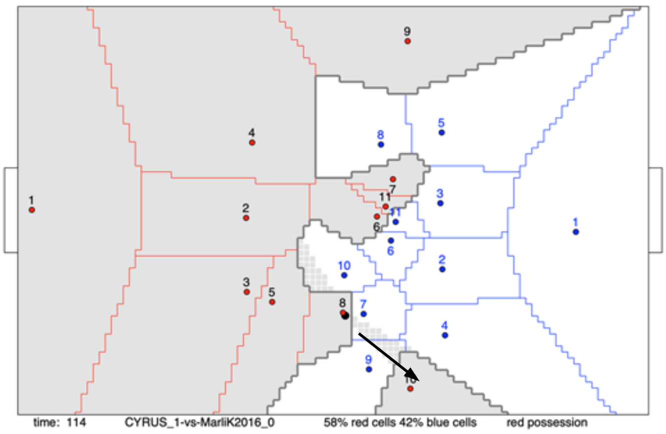


Fig. 8. time 114: $<<r_8, \text{ball}; R_{\text{ownership}}>, r_{10}; R_{\text{pass}}>$

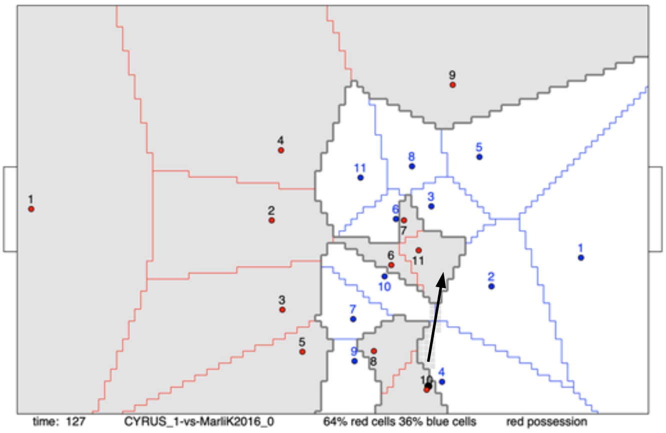


Fig. 9. time 127: $<<r10, ball; R_{ownership}>, r11; R_{pass}>$

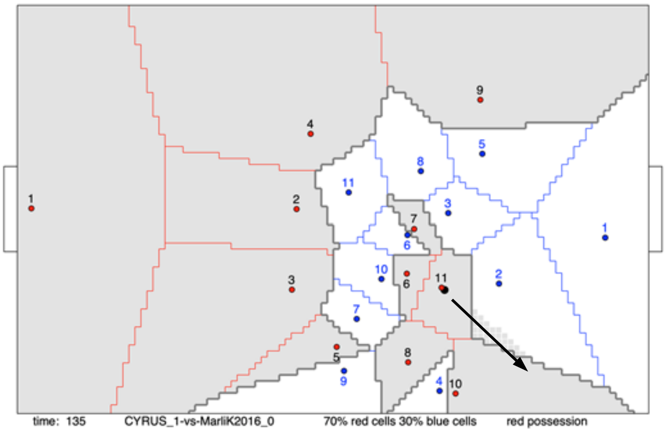


Figure 10. time135: $<<r11, ball; R_{ownership}>, r10; R_{pass}>$

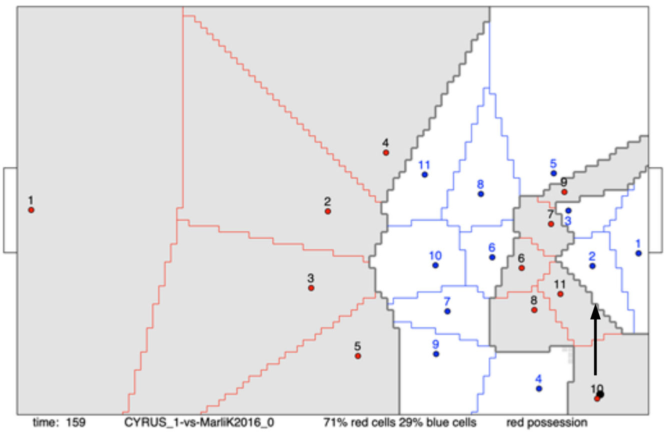


Fig.11. time 159: $<<r11, ball; R_{ownership}>, r10; R_{pass}>$

Although red is playing a very strong attacking game, between times 114 and 127, r5 moves down to deny b9 the space it enjoyed – a strong defensive move (Fig. 9).

Red presses home its attack between times 114 and 127. $<r_6, r_7, r_{11}; R_{island}>$ occupies more of the pitch and r_8 moves fast towards r_{10} to create $<r_8, r_{10}; R_{island}>$ on its right wing (Fig. 9).

At time 127 (Fig. 9) r_{10} receives the ball to form the structure $<r_{10}, ball; R_{ownership}>$, creating the possibility of a pass structure $<<r_{10}, ball; R_{ownership}>, r_{11}; R_{pass}>$.

Between times 114 and 127 b_4 moves towards r_{10} to try to form $<r_{10}, b_4; tackle>$. However (Fig. 9) b_4 encounters the *defender's dilemma*: b_4 has two choices (i) tackle r_{10} to try to gain the ball, giving r_{10} the possibility of passing to r_8 within the island $<r_6, r_8; R_{island}>$ or passing the ball into the newly formed $<r_6, r_7, r_{11}; R_{island}>$, or (ii) moving away from r_{10} to try stop these passes. Here b_4 chooses to try to tackle r_{10}, but r_{10} passes the ball into the island $<r_6, r_7, r_{11}; R_{island}>$.

There are two defenders dilemma structures in Figure 9: $<b_4, r_{10}, r_8; R_{dd}>$ and $<b_4, r_{10}, r_{11}; R_{dd}>$. The first is weak because $<b_9, r_8; R_{close}>$ and a pass to r_8 could lose the ball. $<b_4, r_{10}, r_{11}; R_{dd}>$ is safer and r_{10} passes to r_{11}.

Between times 127 and 135 b_4 makes the error of letting r_{10} move pass it towards the goal (Fig. 10). This gives r_{10} control of a large part of its left wing, and lets it connect with $<r_6, r_7, r_8, r_{11}; R_{island}>$ to form $<r_6, r_7, r_8, r_{11}, r_{10}; R_{island}>$. Although b_4 subsequently moves back to try to improve the defensive position on red's right wing it has lost a tempo and cannot stop the advance of r_{10}.

Red's dominant structure at time 135 within blue's half is a great weakness and enables r_{11} to consider a pass to r_{10} deep in the blue half (Fig. 10).

The pass is made and r_{10} makes a fast run down the wing to receive it at time 159. By now red's position has become very strong (Fig 11).

The red island $<r_6, r_7, r_8, r_{11}, r_{10}; R_{island}>$ is moving towards the blue goal and by time 159 has surrounded it to create a much weaker defensive island $<b_1, b_2, b_3; R_{island}>$ with attackers outnumbering defenders by 6 to 2 (excluding the goalkeeper).

By time 170 red has an immensely superior position and following a sequence of thrilling short passes within the red attacking island, $<r_{10}, r_{11}; R_{pass}>$ between times 159 and 170, $<r_{11}, r_9; R_{pass}>$ between times 170 and 172, $<r_9, r_7; R_{pass}>$ between times 172 and 174, with r_9 shooting at time 174 to create the structure $<r_9, ball, blue\_goal; R_{goal\_scored}>$ (Figs 12, 13, 14) at time 180.

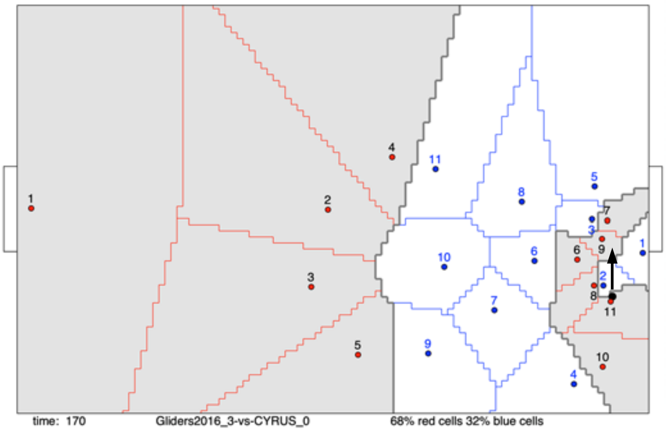


Fig. 12. time 170: <<r_{10}, ball; $R_{ownership}$>, r_{11}; R_{pass}>

Discussion

This is a sad story for blue, but could the outcome have been avoided?

Figure 15 shows a critical state in red's development at time 105. Figure 16 shows, counterfactually, that if blue players b_3, b_4, b_5, b_9 and b_{11} had moved differently between times 94 and 105 they could have denied red the structure from which it launched its attack.

In particular the movements of b_9 and b_{10} disconnect red's structure isolating r_{10} from r_5 so that the pass <<r_5, ball; $R_{ownership}$>, r_8; R_{pass}> is no longer a feasible option. Nor is the alternative <<r_5, ball; $R_{ownership}$>, r_{10}; R_{pass}>. See the green circles in Figure 16.

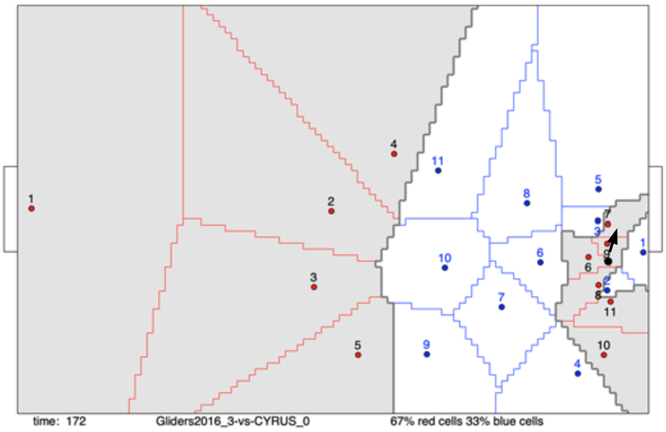


Fig. 13. time 172. : <<r_{11}, ball; $R_{ownership}$>, r_9; R_{pass}>

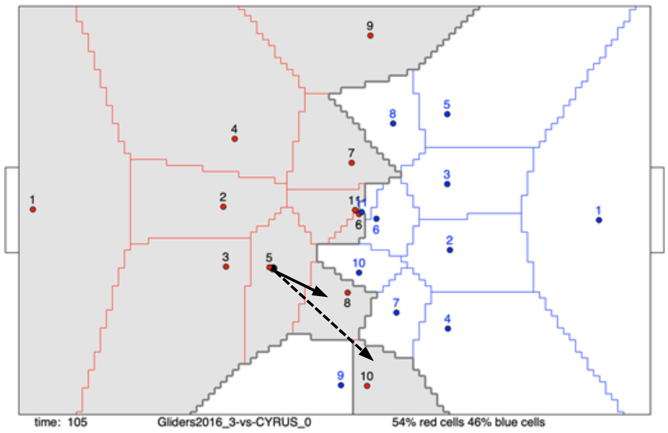


Fig. 15. The position at time 105 given in Figure 7.

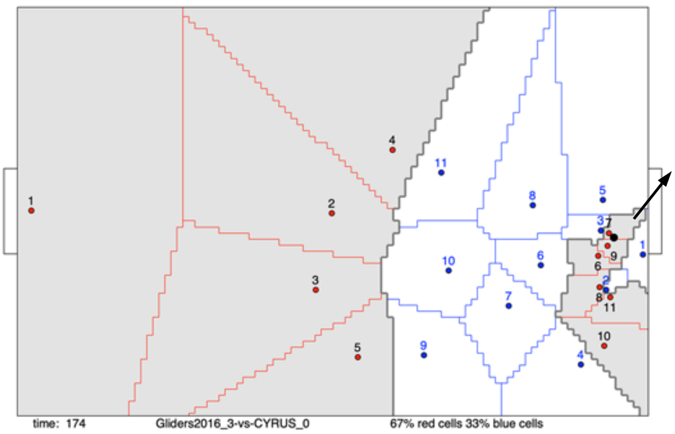


Fig. 14. time 174. : <<r_9, ball; $R_{ownership}$>, r_7; R_{pass}>

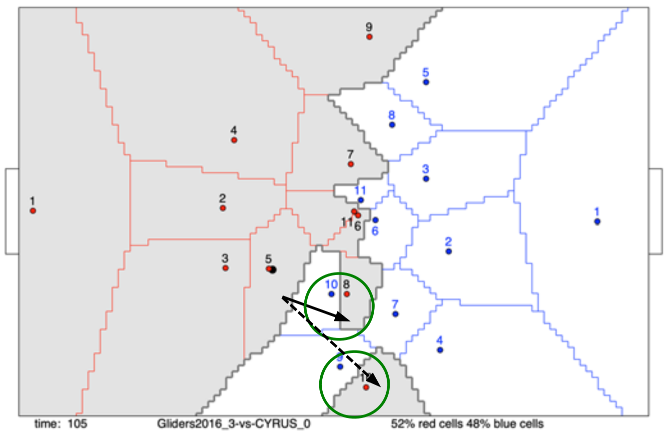


Fig.16. time 105: Counterfactual movements of b_9 and b_{10} could have disconnected the red structure and denied red the pass structure <<r_5, ball; $R_{possession}$>, r_8; R_{pass}> and the pass structure <<r_5, ball; $R_{possession}$>, r_{10}; R_{pass}>

Figures 17 and 18 show the counterfactual outcome if blue players b_6, b_9 and b_{11} had moved differently between times 105 and 114. Then the structure would again be less favourable for the red team and deny it the pass $\langle\langle r_8, \text{ball}; R_{\text{ownership}}\rangle, r_{10}; R_{\text{pass}}\rangle$.

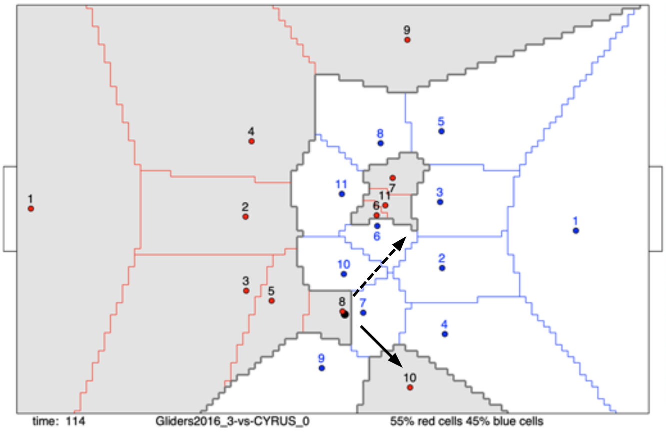


Fig. 17. The position at time 114 given in Figure 8.

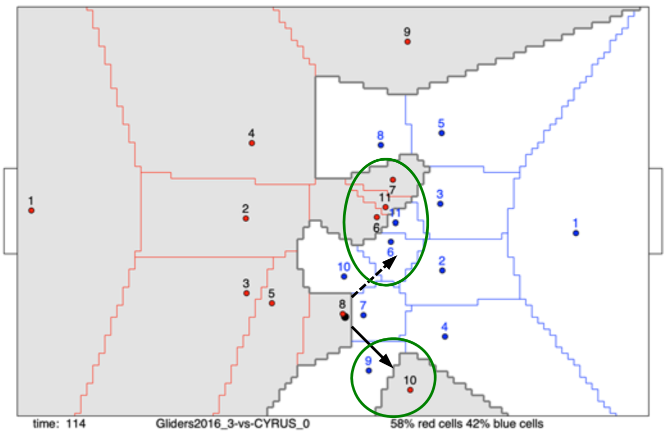


Figure 18. Counterfactually, blue can achieve a better position at time 114 if b_9, b_{10} and b_{11} move differently.

4. Conclusions

In [2] we give three research questions concerning our approach to robot soccer. The first was whether some hypersimplices are particularly disposed to scoring goals? Our recent research suggests that the ability to force a defender's dilemma indicates that teams are disposed to win [5].

In this paper we have focused on the hypersimplex notation and the requirement to make relations explicit. What is new about this is that have shown how to form higher level relational structure between lower level relational structures of the form

$$\langle \langle a, b, c; R_1\rangle, \langle\langle x, y; R_2\rangle, z; R_3\rangle\rangle, d, e, f; R_4 \rangle$$

Also it has been shown that using this notation can be a bridge between free format speaking to analyse multilevel systems and the precise notation required for machines to manipulate the structures. This relates to the second question in [2] which concerned machine-learning the relational structure of hypersimplices. For machine learning it is necessary to have a way of representing structures that can be implemented in computers.

Our approach to robot soccer is explicitly relational and follows the approach taken by Atkin for computer chess. It has been illustrated that the highly dynamic connectivity of the relationships between the cells of the pitch, the players and the ball has an important role in robot soccer. This provides a tentative answer the third question in [2] about exploiting the multidimensional connectivity of hypersimplices to develop tactics and strategies in robot soccer.

5. Further work

Having spent a considerable effort developing the ideas presented in this and previous papers, the next step is to develop a competitive RoboCup robot soccer player. We expect to report progress on this at ICAROB 2021.

6. References

1. R.H. Atkin, *Mathematical structure in human affairs*, Heinemann Educational Books, London, 1974.
2. J. H. Johnson, and R. Rossi, Dynamic Structures for Evolving Tactics and Strategies in Team Robotics. *Journal of Robotics, Networking and Artificial Life*, Dec 2019.
3. H. Kitano, M. Asada, Y. Kuniyoshi, I. Noda, E. Osawa, Robocup: the robot world cup initiative. in: W.L. Johnson, B. Hayes-Roth, (Eds.), *Proceedings of the first international conference on autonomous agents*, ACM Press, New York, 1997, pp. 340–347.
4. J. H. Johnson, *Hypernetworks in the science of complex systems*, Imperial College Press, 2014.
5. R. Rossi, J. H. Johnson. 'The defender's dilemma in robot soccer and other structural statistics', Mimeo, The Open University, 2020.

User Interface and Motion Planner for Task Database

Tokuo Tsuji[1], Kensuke Harada[2,3], Natsuki Yamanobe[3]

[1]Institute of Science and Technology, Kanazawa University, Kadoma-cho, Kanazawa 920-1192, Japan

[2]Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama, Toyonaka 560-8531, Japan

[3]Intelligent Systems Research Institute, National Institute of Advanced Industrial Science and Technology(AIST), 1-1-1 Umezono, Tsukuba 305-8568, Japan

E-mail: tokuo-tsuji@se.kanazawa-u.ac.jp, harada@sys.es.osaka-u.ac.jp, n-yamanobe@aist.go.jp

Abstract

In this paper, we present user interface modules for handing task motion data on robotic manipulation. The database covers robot motion, human motion and object information. The authors especially present its data structure, data registration, data search, GUI, and API. The task motion data is registered using vision sensor and GUI. The robot work motion is generated automatically with planning module and modification module.
Keywords: Robotic assembly, Motion planning, Task planning, Database.

1. Introduction

Most robots repeat a determined motion in mass production. However, in the case of variable production, it is necessary to teach a new motion to robot according to changes in the product. It is said that robot teaching and device setting costs are much larger than the cost of the robot body. Therefore, frequently teaching robot motions and changing devices according to product changes will lead to a significant increase in cost.

Human flexibly realize a wide variety of tasks. It is considered that the experiences of various tasks are abstracted, conceptualized, and accumulated as skills and knowledge. Then the skills and knowledge are applied to new tasks.

The authors have constructed a cloud database [1][2][3] that accumulates various information related to motions and shares among robots for motion generation. The database accumulates human work motions and teaching data for robots, work procedure, and object information such as the shape of the object, motion trajectories during work.

This research aims to construct the database for a robot adaptively performing a task. Fig. 1 shows the overview of its framework. The database is composed of

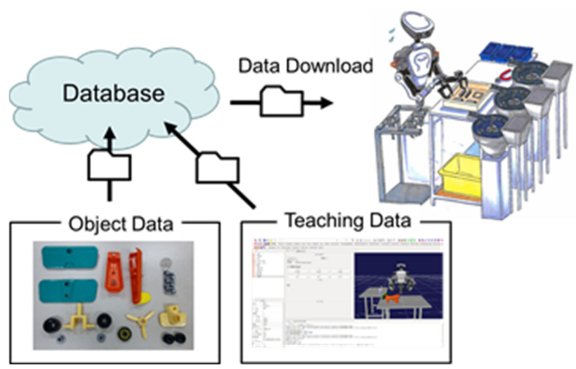


Fig. 1. Framework on motion generation of industrial robot from database where the database includes information on object data and teaching motion data

the object data and teaching motion data. When a robot performs a novel task, a robot download and modify the data so as to fit the novel task.

2. The structure of the database

The structure of database is designed by considering versatility and reusability. The motion is described by the sequence of coordinates of the object and is independent of the configuration of robot. A motion can be shared in general robots. The hierarchical description makes it possible for robots to reuse appropriate level of motion.

An assembly of a product is composed of multiple task elements where each task element is composed of a set of motion sequence. Therefore, the first, the second and the third layers are respectively named as "work", "task" and "action" layers. Since multiple objects are related in a single action, two objects are named as follows. The object held by a robot is named as the "main object" and the object assembled to the main object is named as the "sub object". The action data include the trajectory which is the time series of relative posture of the main object and the sub object. If the position of the sub object changes, the motion is adapted using the relative postures in the action data. The 3 layers are effective for reusing motion. For example, the action "tightening screw" can be used in another task data. The task data can also be used in another work.

Fig. 2 shows an example of assembly of a toy airplane. The task is "fixing the wing to the body" and composed of two actions: "put the wing on the body", and "put and rotate the screw". At the first action, the main object is the upper wing and the sub object is the upper body. At the second action, the main object is the screw and the sub object is the upper wing.

The following is a set of information included in the database:

・Motion strategy (A sequence of task/motion)
・Human (Identified name etc.)
・Robot (Name, type, hand)
・Object Name
・Object category (e.g., JICFS code)
・Object shape

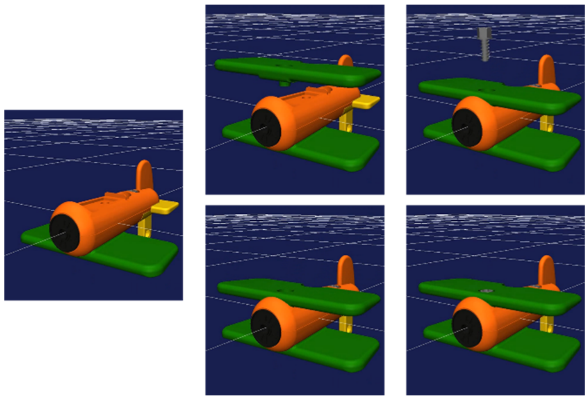


Fig. 2. An example of task for toy assembly.

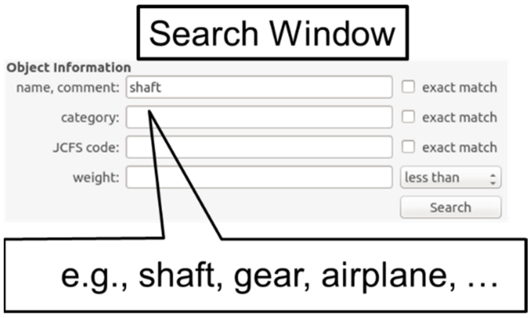


Fig. 3. Search window for the database.

・Object parameters (Mass, CoG, material, stiffness, etc.)
・Grasping configuration (Main object)
・Fixture configuration (Sub object)
・Time series of the main object's pose w.r.t the sub object
・Control method and its parameters
・Force information applied by the hand

The data can be obtained by several methods such as a human motion capture and a motion generation tool. The data is stored in the YAML format where some examples are shown in Fig. 2.

3. Interfaces of databases

3.1. Motion searching function

Fig. 3 shows the searching function of the motion database plugin. As keywords used for the search, we set

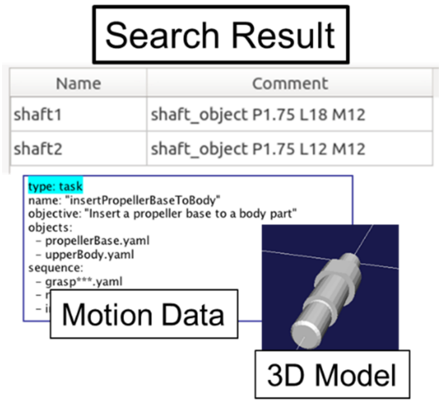


Fig. 4. Search result of task database.

the name of parts, the category of parts, the feature of

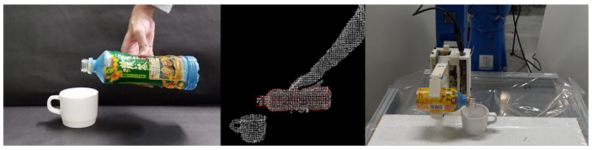


Fig. 5. Vision system for measuring object motion grasped by human and robot motion generation using the motion data.

object data, object's JCFS code, and object's weight. Inputting keywords in the search window, we obtain the candidates of objects and motion data. The Fig. 3 and Fig.4 show an example where we search the keyword "shaft" and obtained the object model "shaft1" and "shaft2" and corresponding motion data.

3.2. Object trajectory of human motion

Vision sensor is used generating action data. The trajectory of an object is measured by vision and stored to the database. In Fig. 5, an example is shown. A depth camera is used for measuring the position of a plastic bottle. The trajectory of a plastic bottle grasped by human is stored to the database as an action data. After a robot plans the grasping position on the plastic bottle, the robot moves the same trajectory of human.

3.3. Graphic user interface

Graphic user interface of Choreonoid is implemented for registering the action data in the database. The main object and the sub object can be selected by clicking a button on the window. The object position of the main

Fig. 6. Definition of features.

object can be operated by mouse motion and be set with the value. The trajectory of the main object is saved as yaml file.

The grasp position can be registered manually using the graphic interface. The robot motion can be generated according to the motion of the object.

4. Motion generation using database

4.1. Motion planning for generating task motion

The authors have implemented motion planner. Robot motion of the task data is generated by using grasp planning and path planning [4]. The hand position on the object position is determined by grasp planning. The path planning is used for connecting the ends of multiple actions. After once grasp position of the main object is determined, the robot motion is generated automatically according to the motion of the main object. If the hand collides the sub object during an action, grasp position is planned again.

4.2. Motion modification

Let us explain the feature of objects defined in this research. Fig.6 shows three examples of features. In case of a bolt, diameter, pitch and length are the features. In case of a gear, diameter and the number of teeth are the features. In case of a L shaped object, the dimension of the object is the feature.

Fig. 7 shows the modification module of motion data. The inputs to this module are the features of the new main and sub objects and the existing motion data of similar but different object. An example of the main and sub objects, we can consider the bolt and nut. As an example of new objects, we can consider the new bolt with different pitch and length. By considering the difference of feature between objects, we modify the motion data and make it fit to the newly appeared objects.

In our experiment, we used two screwing units with different features as shown in Fig. 8. The first unit has the screw length 12mm and base thickness 15mm. The second unit has the screw length 18mm and base thickness 25mm. We prepared the motion data to assemble the first unit and modify this motion data to fit to the second unit.

Fig. 9 shows the experimental results. As shown in this figure, the robot adaptively generates the motion data and generates the screwing motion of different screwing unit.

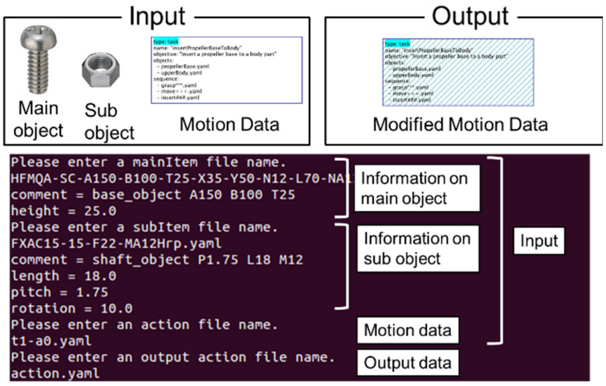


Fig. 7. Modification module of motion data.

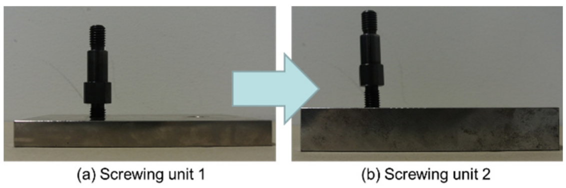


Fig. 8. Screwing units used in the experiment.

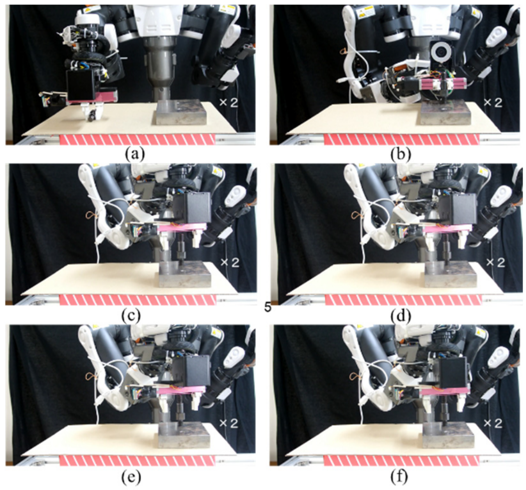


Fig. 9. Experimental result of motion modification.

5. Conclusion

Task database and the interfaces are implemented for robotic motion generation. The task data is independent with the configuration of robots. We designed three layers for task data. The data of each layer can be reused. The GUI and vision sensor is used for storing task data. The robot motion is generated with search module, planner module, and modification module.

References

1. N. Yamanobe, T. Tsuji, K. Harada, K. Nagata, R. Hanai, and W. Wan, and I.G. Ramirez-Alpizar, "Construction of Cloud Database for Task Motion Generation", *Annual Conf. on the Robotics Society of Japan*, 1B2-07, 2016.
2. K. Harada, K. Nakayama, W. Wan, K. Nagata, N.Yamanobe, T. Tsuji, and I.G. Ramirez-Alpizar, "Tool Exchangeable Grasp/Assembly Planner", *Proc. of Int. Conf. on Intelligent Autonomous Systems*, #34, 2018, pp. 1-13.
3. W. Wan and K. Harada, "Developing and Comparing Single-arm and Dual-arm Regrasp," *Robotics and Automation Letter*, vol. 1, no. 1, 2016, pp. 243-250.
4. T. Siméon, J.P. Laumond, J, Cortés, and A. Aahbani", Manipulation Planning with Probabilistic Roadmaps", *Int. J. Robotics Research,* vol.23,no.7-8, 2004, pp.729-746.

POMDP-based action planning for the recognition of occluded objects with Humanoid robots

Masato Tsuru
Graduate School of Engineering Science, Osaka University, 1-3, Machikaneyama, Osaka, Japan
E-mail: tsuru@hlab.sys.es.osaka-u.ac.jp

Pierre Gergondet
CNRS-JRL, AIST 1-1-1, Umezono, Tsukuba, Japan

Tomohiro Motoda
Graduate School of Engineering Science, Osaka University

Adrien Escande
CNRS-JRL, AIST

Eiichi Yoshida
Intelligent Systems Research Institute, AIST, 1-1-1, Umezono, Tsukuba, Japan

Ixchel G. Ramirez-Alpizar
AIRC, AIST, 2-3-26, Aomi, Koto-ku, Tokyo, Japane

Weiwei Wan
Graduate School of Engineering Science, Osaka University

Kensuke Harada
Graduate School of Engineering Science, Osaka University, AIRC, AIST

Abstract

In this paper, we present a high-layer motion planner which plans humanoid robot actions to search for object models in the robot workspace. To overcome the occlusion problem, our proposed method plans to get different perspectives of the object. POMDP (Partially Observable Markov Decision Process) is used to determine the observation pose of a robot. The planner then builds a comprehensive point cloud of by merging the point clouds gathered from different positions. The point cloud is compared to a 3D model of the object to estimate the pose of the real object.

Keywords: *humanoid robot, POMDP, planning, point cloud, occlusion, 6DoF registration*

1. Introduction

Recently, the necessity for humanoid robots, to assist us in our daily life, has been increasing. In such situations, since robots usually have to detect specific objects by themselves, object search and recognition for humanoids is of great importance for robotics researchers.

We propose a method to improve object recognition and localization capabilities for humanoid robots. Our method uses POMDP (Partially Observable Markov Decision Process) to plan the view pose of a robot to scan

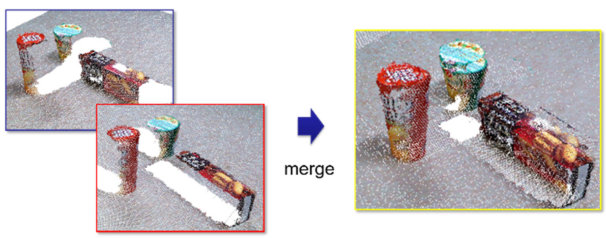


Fig. 1. By merging two point clouds which obtained from different directions, our robot relaxes the problem coming from occlusion.

the object and its environment by using a 3D depth sensor, and create a merged point cloud derived from different points of view. To do it, we set one task, searching an object. We give a target object as a 3D model to the robot, then the robot try to find it from a certain region. The robot obtains 3D point cloud data through its head camera, Xtion, and detects the target object by point cloud matching. However, in the event of occlusion, the robot must recognize the object using incomplete information, which is not always possible. To avoid said problem, our robot moves to another position to get additional point cloud (Fig.1).

It is important to choose a second robot position or viewpoint, which maximizes the information gathered from point cloud data gathering. To minimize the number of times the robot must change position to gather a satisfactory cloud point of the object and its environment. Therefore, we propose a high-layer system to compute the necessary viewpoints and robot positions required to gather a complete point cloud representation of the object. The system analyzes gathered point cloud and estimates the direction which can fill the missing point cloud. Thanks to POMDP controller, our robot doesn't choose a same view direction many time. After changing its position several times and finding the object, the robot tries to grasp the target.

2. Related Works

Saidi et al.[1] searched a target object with autonomous humanoid robot. They also point out the difficulty of object searching task, and solved it with a walking strategy. They used heuristics to plan the next action in unknown environments. We improves the strategy of search with POMDP. Our work also deals with object grasping after a successful recognition is performed.

Kim and Likhachev[2] implemented POMDP with PR2, a mobile robot. They tackled the occlusion problem in 3D point cloud. Their robot detected objects with changing its position and height of torso. The target object was estimated from a small partial point cloud that was hidden by other objects. Finally, they set Breadth-First Search as a baseline and proved the efficiency of POMDP through several experiment. In this research, we apply this method to a humanoid robot HRP-2 that performs stability problems or complex posture problems.

Foissotte et al.[3] proposed a method which selects the best posture to observe a target object, within humanoid stability constraints. In that research, the target location is fixed, and humanoid robot is not allowed to change its position dynamically. The robot can only take one step, change its head position and orientation properly, and fold out its arms to keep stability. The viewpoint is selected by the amount of information. Their method was tested in a simulation environment with OpenGL, and evaluate the observable area with considering occlusions. They proved the efficiency of the planned action for humanoid; however, they didn't generate joint trajectories and consider the dynamical stability in changing posture. In our research, we extend this research and allow the robot to change its standing position and posture. Thanks to it, less difficult postures are generated and it is easy to realize.

3. Proposed Method

3.1. *Overview*

Our system overview is shown in Fig. 2. In this research. We discrete the searching space to simplify the POMDP formulation, and list up all of feasible viewpoints. From that simplification, the robot can search only in a pre-limited area. First, the robot gets point cloud from a randomly selected position in the list. Then, each viewpoint is evaluated by its probability to get new information, depending on the obtained point cloud. The robot executes the highest scored observation action and gets a new point cloud from another direction. The system will merge the new point cloud with the previously obtained one by using the ICP (Iterative Closest Point) algorithm, based on the robot position and orientation. After every observation, the POMDP controller re-calculates every viewpoint score, judging from known point cloud. The robot continues its search until the object pose is fully recognized.

In every step, the robot estimates the object's position and checks that the ICP error. In the case of small errors, the object pose is considered to be known and a grasp of the object is attempted. In grasp planning, robot solves Inverse-Kinematics of right arm, from its shoulder to the hand and checks the stability of the posture. And an approaching motion of the right arm is also generated automatically.

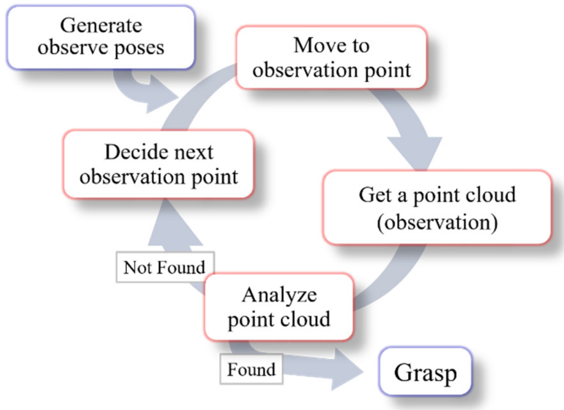


Fig. 2. Basic structure of proposed system

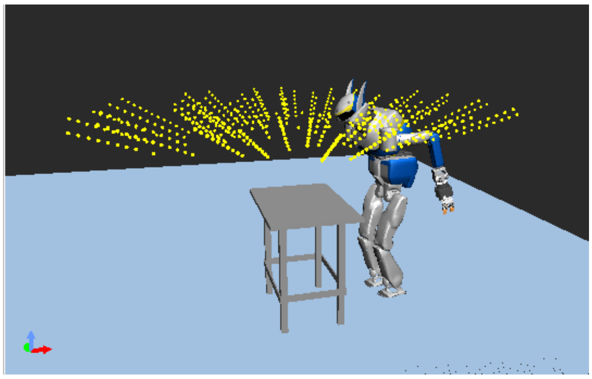


Fig. 3. Generated viewpoint candidates and one example of the postures. Every yellow viewpoint has its observation posture, which are confirmed stability and joint angle limits.

3.2. *Posture Generation*

Because our method needs a lot of posture candidates, we generate a lot of observation postures at first step. As a prime of this research, the searching environment is known roughly. So, we set the target area on the table. Then, we define view directions toward the area. After that, viewpoints, which corresponds to robot head positions and orientations, are created for each 5 cm on the view direction lines. Our planner calculates the whole-body Inverse Kinematics to place the robot head to each viewpoint. If the solution is stable, we add the new one position to a feasible action list. Doing this process, the robot obtained several observation postures like Fig. 3.

3.3. *Point Cloud Processing*

In this research, we use a feature-based point cloud recognition for detecting the target object by merging point clouds. We use CVFH (Clustered Viewpoint Feature Histogram) descriptor[3], which can be used to recognize an object from even partially observed point cloud. We give the 3D model of the object to CVFH before starting experiment. CVFH allows the robot to initially estimate the object location. Then, our algorithm calculates a new robot pose to obtain the missing point cloud data. If there are no hints, even POMDP must decide the next action almost randomly and it is not desirable for us. The robot then proceeds to change its position across its workspace, gathering several point clouds from different directions, finally merging the gathered data into a single large point cloud. Finally, we use the ICP algorithm to fit the object 3D model to the multidirectional point cloud. After acquiring a new set of point clouds from a given viewpoint, our algorithm merges it with the previous sets and uses ICP to try to match the 3D model of the searched object with the merged set of point clouds. ICP returns a score that quantifies the matching between the point cloud and the 3D model. If the score is high enough, the algorithm considers the object pose to be known and uses the position and orientation given by ICP to plan a grasp motion of the object.

3.4. *POMDP*

The POMDP represents the main contribution of this paper. POMDP (Partially Observable Markov Decision Process) is extended and generalized from MDP (Markov Decision Process), and is used in partially observable environments. In POMDP, the position and orientation of an object is represented by a probability distribution. The system updates the probability distribution (it is often called as belief) by every ICP result depending on its accuracy score. If there are some similar objects, the CVFH + ICP usually propose multiple candidates. In such case, the POMDP changes its probability distribution for each candidate. And then, the POMDP evaluates every action by virtual action. In the virtual action process, the POMDP system simulates all of possible observations and evaluates these qualities and quantities. That means, for every feasible viewpoint, by assuming all of the candidates are true and observing them virtually, the system calculates unknown area that the camera may observe. These virtual observation scores are multiplied with the probability distribution, and then, the robot executes observation from the highest scored viewpoint.

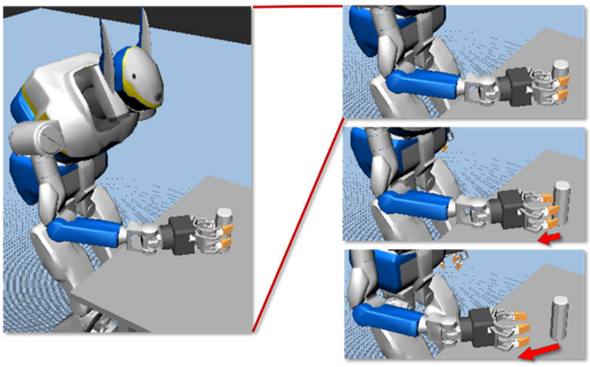


Fig. 4. Generation of grasping motion. The final grasping pose is first generated, and then the approaching trajectory is generated.

3.5. *Grasp Planning*

After the algorithm detects the object pose, we use grasp planning to make the robot pick-up the object. The 3D model is given for searching, so we use a model-based grasping method. Harada et al.[4] realized a method of general grasping motion generation with approximating the target object by cylinders. They also discussed about an approach direction in grasping with considering open and close of hand. Our system implements said method to realize automatic grasp motion generation. The system approximates the target 3D model by cylinders or boxes. Then it generates grasping hand poses and checks each stability. And an approaching motion is planned automatically by withdrawing the hand, with relying to the robot hand's coordinate system (Fig. 4).

4. Simulation Experiment

Experiment was done with Kinect, and we moved it to viewpoints planned by POMDP. The HRP-2 robot in simulator knows the 3D model of the target object. The target of this experiment is to identify the position of a can in a cluttered environment. Our proposed system generated about 250 observation postures surrounding the table. The first posture was selected randomly since there are no prior information about the object position. After the first observation, the planner decided a second viewpoint for the robot in order to get additional cloud point data from a different perspective. In our simulations, the robot changed its position three times before identifying the object. Finally, the robot grasped the can, as seen in Fig. 5.

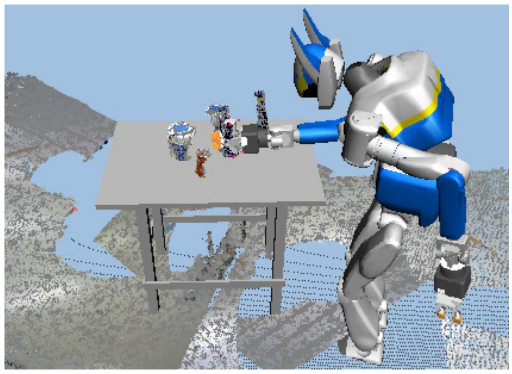


Fig. 5. The robot finally identified the can and grasped it. To disturb the recognition of the can, a flashlight and a case of wet wipe were also placed on the table.

5. Conclusions and Future Works

We proposed an observation motion planning framework for a biped humanoid robot based on POMDP to compensate missing cloud point information, in cluttered environments.

However, our algorithm does not take apply the cost to the robot movement into account when planning the viewpoints. In future work we will implement the costs to robot movement to reduce the walking distance of a robot.

6. References

1. F. Saidi, O. Stasse, K. Yokoi and F. Kanehirot, "Online object search with a humanoid robot," 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems, San Diego, CA, 2007, pp. 1677-1682.
2. Sung-Kyun Kim and M. Likhachev, "Planning for grasp selection of partially occluded objects," 2016 IEEE International Conference on Robotics and Automation (ICRA), Stockholm, 2016, pp. 3971-3978.
3. T. Foissotte, O. Stasse, A. Escande and A. Kheddar, "A next-best-view algorithm for autonomous 3D object modeling by a humanoid robot," Humanoids 2008 - 8th IEEE-RAS International Conference on Humanoid Robots, Daejeon, 2008, pp. 333-338.
4. A. Aldoma et al., "CAD-model recognition and 6DOF pose estimation using 3D cues," 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops), Barcelona, 2011, pp. 585-592.
5. K. Harada, K. Nagata, T. Tsuji, N. Yamanobe, A. Nakamura, and Y. Kawai, "Probabilistic approach for object bin picking approximated by cylinders," in Proc. IEEE Int'l Conf. Robotic and Automation (ICRA), May 2013, pp. 3742–3747

Cost-oriented Planning for Error Recovery in an Automation Plant

Akira Nakamura[*1], Natsuki Yamanobe[*1], Ixchel Ramirez Alpizar[*1], Kensuke Harada[*1, *2] and Yukiyasu Domae[*1]

*[*1]Artificial Intelligence Research Center*
National Institute of Advanced Industrial Science and Technology (AIST)
Second Annex, AIST Tokyo Waterfront, 2-4-7 Aomi, Koto-ku, Tokyo 135-0064 Japan

*[*2]Robotic Manipulation Research Group*
Systems Innovation Department
Graduate School of Engineering Science, Osaka University
1-3 Machikaneyama, Toyonaka 560-8531, Japan

E-mail: a-nakamura@aist.go.jp, n-yamanobe@aist.go.jp, ixchel-ramirezalpizar@aist.go.jp,
harada@sys.es.osaka-u.ac.jp, domae.yukiyasu@aist.go.jp
www.aist.go.jp

Abstract

In an automation plant, errors are more likely to occur during complicated tasks. In the case of a major error, the task is commonly re-executed after returning to the previous step. Therefore, deciding both the prior step that should be returned to and the recovery approach after return are important problems to consider. In this paper, cost-oriented planning of error recovery taking these two factors into account is proposed.

Keywords: cost-oriented planning, error recovery, task stratification, error classification, automation plant

1. Introduction

In recent years, the automation of production plants has become a popular phenomenon. As high-mix low volume production is increasing, production methods differ from those in the case of large-scale mass production of the same products, and this has resulted in an increase in the cases of manufacturing process failure. Thus, the development of techniques to recover the processes is considered important.

Various studies on error recovery have been conducted in the manufacturing and robot research fields.[1-6] While several studies have been conducted on techniques to add a correction to the input of plants with a small error and perform recovery, there has been insufficient research on methods to return to the previous step of failure occurrence in order to perform recovery.

In this study, we investigate returning to a previous step for error recovery. We have considered methods of error recovery using task stratification and error classification.[7-9] In particular, we have shown techniques of plant correction which are performed using an error cause derived by the proposed method, and a method of planning for error recovery in consideration of cost has been proposed. When deriving the recovery process, multiple possible courses appear in several cases; in our method, the most suitable course is chosen by considering the cost of error recovery.

The remainder of this paper is organized as follows. The concept of skill is described in Section 2. The technique of failure recovery using error classification is explained in Section 3. A method of cost-oriented

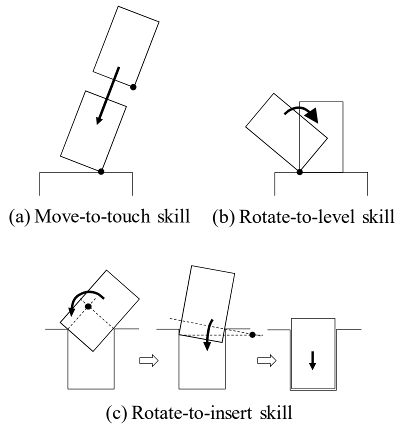


Fig. 1 Three fundamental skills

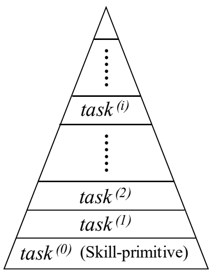


Fig. 2 Hierarchy of tasks

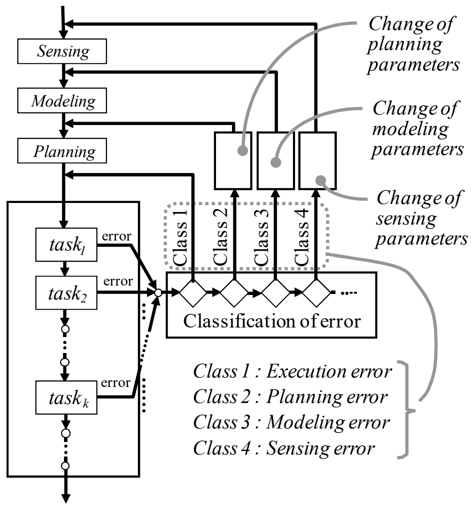


Fig. 3 Fundamental process flow with error recovery

planning for error recovery is proposed in Section 4. Finally, in Section 5, a simple sample is shown using error occurrences in an assembly task.

2. Concept of Skills

In this paper, a "*skill*" implies the unit of motion. This section explains skills which are components of human behavior and machine motion.[10-12]

2.1. *Skill primitives*

We have derived the motion primitives which constitute tasks such as assembly and transport by analyzing a person's behavior. The motion unit is defined as a "skill".[10-12] Three basic skills, "move-to-touch," "rotate-to-level," and "rotate-to-insert," are important (Fig. 1). The representative person's behavior can be composed of these three basic skills and similar ones. The units of machine motion can be considered to be similar to the rudiments of human behavior.

2.2. *Stratification of tasks*

In general, it is possible to regard a task performed by a machine as following a hierarchical structure (Fig. 2).[10-12] One tier above the layer "$task^{(i)}$" is the layer "$task^{(i+1)}$". The "*skill*" layer is located in the lowest layer, "$task^{(0)}$".

3. Error Recovery

In the actual environment, unlike the ideal case, various factors can cause errors in machine performance. This section describes an error classification concept and our error recovery technique.[7-9]

3.1. *Error classification*

Errors can be classified into several groups based on the possible causes. We consider four error groups: execution, planning, modeling, and sensing (Fig. 3).[7-9]

3.2. *Error recovery based on classification*

First, when the error occurs, the cause is estimated. Next, suitable correction of the system is performed based on the tentative cause. The process returns to the previous step, and the task is executed again from this step (Fig. 3).[7-9] The same error is less likely to occur because the corrected process has been executed.

If the scale of the error is small, the process returns to the previous step in the lowest layer of the hierarchy

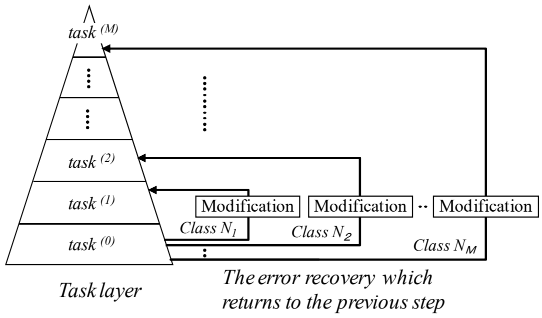


Fig. 4 The expression of task stratification and the process flow of the error recovery

(Fig. 3, Fig. 4). Conversely, if the scale of the error is large, the process returns to the previous step in the highest-ranking layer of the hierarchy, and it is executed again from this step (Fig. 4).

4. Cost-oriented Error Recovery

In this section, the candidate error recovery processes and the selection of the appropriate recovery process are explained.

4.1. *Candidate processes for recovery*

In the previous section, it was shown that the step to which the process returns after failure occurrences alters according to the scale of the error. However, it is possible to return further back than a single step of necessity minimum.

Figure 5 shows various possible error recovery processes for failure occurrences at $Subtask_m$ in Task $\boldsymbol{T}$ [from Start $\boldsymbol{S}$ to Goal $\boldsymbol{G}$], where Task $\boldsymbol{T}$ consists of sequence [$Subtask_1$, $Subtask_2$, $Subtask_3$, $\cdots$, $Subtask_m$, $\cdots$, $Subtask_n$] and $Subtask_k$ (k = 1, 2, 3, $\cdots$, n) consists of sequence [$Skill^k{}_1$, $Skill^k{}_2$, $Skill^k{}_3$, $\cdots$, $Skill^k{}_{n\_k(max)}$]. $Subtask_m$ indicates the minimum traceable unit described in [8], that is, the smallest unit in which it is necessary to return to the first node of a skill primitive sequence if a failure occurs. A skill where a failure occurs is part of the sequence [$Skill^m{}_1$, $Skill^m{}_2$, $Skill^m{}_3$, $\cdots$, $Skill^m{}_{n\_k(max)}$]. When no problems are encountered, there is possibility that the execution of several skills has continued within $Subtask_m$, without returning to the previous step immediately after the failure occurrence.

When the process returns to the node before $Subtask_j$ (j is one of 1, 2, 3, $\cdots$, m), the process [$Subtask_j$ $Subtask_{j+1}$,$\cdots$, $Subtask_n$] is executed in the corrected system (the center bold arrows of Fig. 5). However, re-execution using the same sequence of subtasks is not generally possible owing to several factors such as the transformation of the target object and the change of arrangement of the objects. In such a case, the rerun is performed by using an equivalent task of the original sequence of subtasks (the right-side arrows of Fig. 5). The equivalent task is derived by large-scale re-planning.

4.2. *Selection of the lowest-cost process*

It is necessary to select one out of the multiple possible error recovery processes. Let us consider that it is most suitable to choose the process with minimum practical costs, where costs include material charges, parts charges, electricity bills, and planning expenses. If failure occurs, error recovery is planned, several task sequences for recovery are derived, these costs are calculated, and the most suitable recovery sequence is then chosen.

5. Cost-oriented Recovery in Assembly Tasks

In this section, the cost-oriented error recovery discussed in Section 4 is observed through a simple sample exercise.

5.1. *Candidate processes for recovery*

Let us consider an assembly task in which a nameplate is stuck to a product by four precision screws. Fig. 6 shows a sequence of the task. First, a tacking task is performed (Fig. 6(a) to Fig. 6(d)) in which these four screws are fastened temporarily to the nameplate, placed levelly; secondly, an erecting task is performed (Fig. 6(e) to Fig. 6(f)) in which the nameplate with four screws stands up; thirdly, a touching task is performed (Fig. 6(g) to Fig. 6(h)) in which the nameplate is moved to the fixing location; and finally, an installation task is performed (Fig. 6(i) to Fig. 6(m)) in which the nameplate is fixed on a plane of the product.

Next, we will consider an error, as shown in Fig. 7, in which a precision screw is dropped when a failure

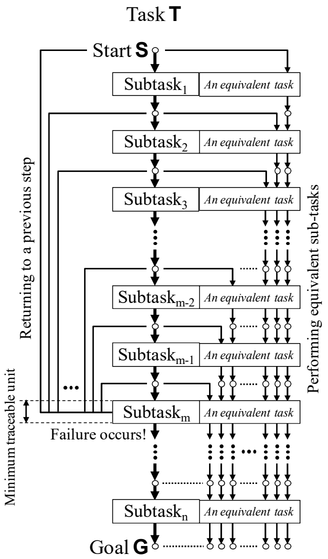


Fig. 5 Various processes of error recovery considered for a failure occurred in Subtaskm

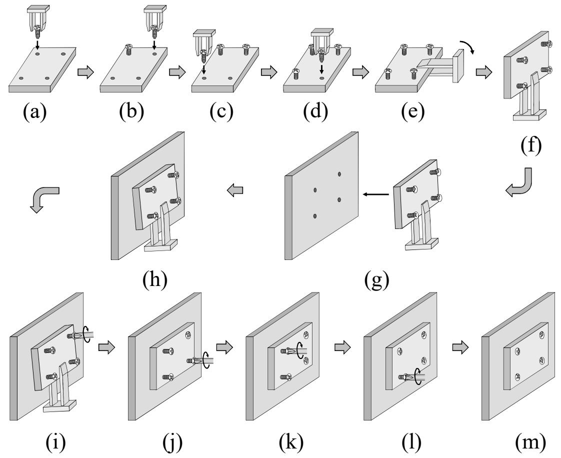


Fig. 6 An assembly task in which a nameplate is stuck to a product by four precision screws

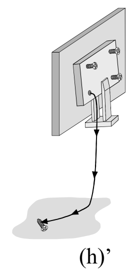


Fig. 7 An error in which a screw is dropped at (h) in Fig. 6

occurs at Fig. 6(h). It is possible to derive various processes for error recovery of this failure. In this example, we consider three types of recovery courses as follows.

- [ER-I] Error Recovery I

  First, we will consider a recovery process that it is easy to understand. The method is rerun from Start ***S***. The nameplate from which one screw in four screws was omitted is discarded, and the re-start is executed using a new nameplate and four new screws, that is, the tacking task (Fig. 6(a) to Fig. 6(d)) is performed using new parts.

- [ER-II] Error Recovery II

  Now, we will consider a recovery process carried out by returning to the previous step of a skill primitive (Fig. 6(d)) in the tacking task. Fig. 8 shows the process with respect to recovery. The sequence from Fig. 8(p) to Fig. 8(r) is an additional process required to return to the previous step. The de-touching task (Fig. 8(h) to Fig. 8(p)) and leveling task (Fig. 8(q) to Fig. 8(r)) are performed. It is noted that the sequence from Fig. 8 (d) to Fig. 8(m) is essentially the same process as that from Fig. 6 (d) to Fig. 6(m).

  There are two types of methods for use of a precision screw at the hole of the fallen screw. One is the use of new screw from the parts box, and another is the reuse of the fallen screw by searching and lifting. The former is shown with [ER-II(N)] and the latter with [ER-II(F)].

- [ER-III] Error Recovery III

  For the third process, we will consider the method for a very small scale of return, that is, a recovery process carried out by returning to a skill primitive in *Subtask$_m$* of the failure occurrence. The method to immediately rerun the tacking task of a screw at the hole for the fallen screw (Fig. 7) in order to regenerate the state of Fig. 6(h) is the easiest to understand. However, we will consider the steady method in which the process of error recovery is carried out after a section *Subtask$_m$* of the installation task (Fig. 6(i) to Fig. 6(k)) is finished. Fig. 9 shows the process with respect to recovery after failure occurrence (Fig. 7). Fig. 9(i), Fig. 9(j), Fig. 9(k), Fig.

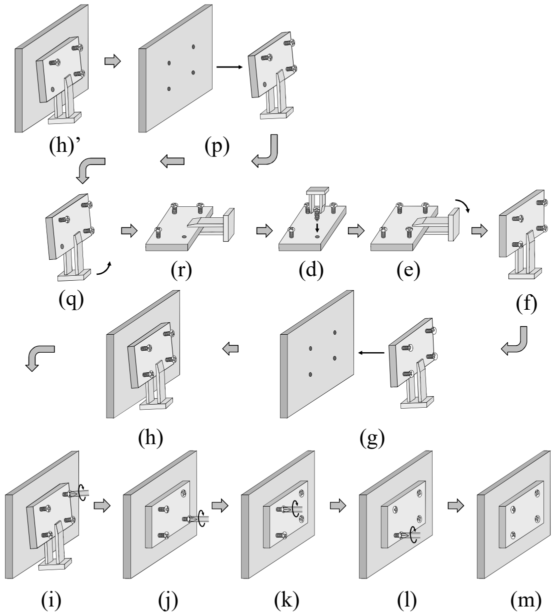


Fig. 8 [ER-II] Error Recovery II

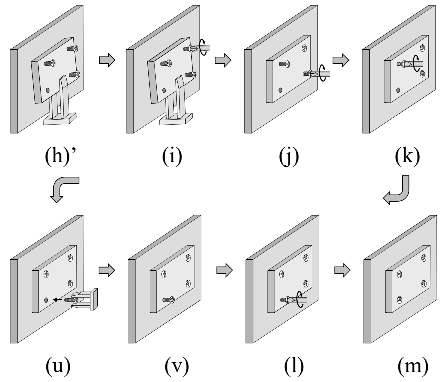


Fig. 9 [ER-III] Error Recovery III

9(l), and Fig. 9(m) correspond to Fig. 7(i), Fig. 7(j), Fig. 7(k), Fig. 7(l) and Fig. 7(m), respectively. The sequence from Fig. 9 (u) to Fig. 9(v) is an additional process to tack a screw into the hole of the nameplate and is performed after the sequence from Fig. 9(i) to Fig. 9(k) because of task reliability. In particular, the task to temporarily fasten a screw to the hole of the nameplate is more difficult when the nameplate is fixed perpendicularly than when it is placed levelly.

There are two types of methods for use of a screw at the hole intended for the fallen screw similar to [ER-II]; [ER-III(N)] and [ER-II(F)] show the methods using a new screw and reusing the fallen screw, respectively.

5.2. *Selection of recovery process based on cost*

Next, the calculation of cost is performed for each recovery process. We consider three types of cost: recovery planning cost *C(R)*, object searching cost *C(S)*, and parts purchasing cost *C(P)*. Here, a unit of cost is introduced for simplification; we assume that the fixed number *"U"* indicates a unit of cost. However, in practical terms, the cost cannot always be related to a fixed value.

First, we consider *C(R)*, which represents the cost generated by the process added with a recovery. The *C(R)* value of [ER-I] is smaller because it follows the same sequence to the original, and conversely, that of [ER-III] is larger because it follows a sequence with a difficult skill primitive. We suppose that *C(R)* of [ER-I] = *U*, *C(R)* of [ER-II] = 2*U*, and *C(R)* of [ER-III] = 3*U*.

Secondly, we consider *C(S)*, which represents the cost generated by the process added for finding a screw. Then, let us assume that [Sch-N] indicates the process of finding a new screw in the parts box, and [Sch-F] indicates the process of finding the fallen screw in the working environment. The *C(S)* value of [Sch-N] is smaller than that of [Sch-F], and thus, we suppose that *C(S)* of [Sch-N] = 2*U*, and *C(S)* of [Sch-F] = 5*U*.

Thirdly, we consider *C(P)*, which represents the cost generated by the purchase of parts. Then, let us assume that [Pch-S] indicates the purchase expense of a new screw, and [Pch-P] indicates the purchase expense of a new nameplate with four new screws. The *C(P)* value of [Pch-P] is significantly larger than that of [Pch-S], and thus, we suppose that *C(P)* of [Pch-S] = *U*, and *C(P)* of [Pch-P] = 8*U*.

We now calculate a cost for each case, considering the total cost $TC = C(R) + C(S) + C(P)$.

$$TC[\text{ER-I}] = C(R) \text{ of } [\text{ER-I}] + 0 + C(P) \text{ of } [\text{Pch-P}] = U + 0 + 8U = 9U \quad (1)$$

$$TC[\text{ER-II(N)}] = C(R) \text{ of } [\text{ER-II}] + C(S) \text{ of } [\text{Sch-N}] + C(P) \text{ of } [\text{Pch-S}] = 2U + 2U + U = 5U \quad (2)$$

$$TC[\text{ER-II(F)}] = C(R) \text{ of } [\text{ER-II}] + C(S) \text{ of } [\text{Sch-F}] + 0 = 2U + 5U + 0 = 7U \quad (3)$$

$$TC[\text{ER-III(N)}] = C(R) \text{ of } [\text{ER-III}] + C(S) \text{ of } [\text{Sch-N}] + C(P) \text{ of } [\text{Pch-S}] = 3U + 2U + U = 6U \quad (4)$$

$$TC[\text{ER-III(F)}] = C(R) \text{ of } [\text{ER-III}] + C(S) \text{ of } [\text{Sch-F}] + 0 = 3U + 5U + 0 = 8U \quad (5)$$

The following relation is derived from Eq. (1) by Eq. (5).

$$TC[\text{ER-II(N)}] < TC[\text{ER-III(N)}] < TC[\text{ER-II(F)}] < TC[\text{ER-III(F)}] < TC[\text{ER-I}]$$

Thus, TC [ER-II (N)] will have the smallest price, indicating that the case of [ER-II] Error Recovery II with use of a new screw instead of the dropped screw has the lowest cost.

6. Conclusion

In this study, a cost-oriented planning for error recovery was proposed. In general, more than one sequence is derived when process of error recovery is considered. This paper explains the method of selection from the multiple possible courses. In our method, the most suitable sequence is chosen based on cost.

In the future, we expect to conduct further research on the derivation of the error recovery process. A selection method besides the cost will be considered, such as working hours and available tools, which are important in automation plants.

References

1. B. R. Donald, Planning multi-step error detection and recovery strategies, Int. J. Robot. Res., 9(1) (1990) 3-60.
2. T. Fukuda, M. Nakaoka, T. Ueyama and Y. Hasegawa, Direct teaching and error recovery method for assembly task based on a transition process of a constraint condition, in Proc. IEEE Int. Conf. Robot. Autom., (Seoul, Korea, 2001), 1518-1523.
3. M. Scheutz and J. Kramer, Reflection and reasoning mechanisms for failure detection and recovery in a distributed robotic architecture for complex robots, in Proc. IEEE Int. Conf. Robot. Autom., (Roma, Italy, 2007), 3699-3704.
4. P. Pastor, M. Kalakrishnan, S. Chitta, E. Theodorou and S. Schaal, Skill learning and task outcome prediction for manipulation, in Proc. IEEE Int. Conf. Robot. Autom., (Shanghai, China, 2011), 3828-3834.
5. T. Matsuno, J. Huang and T. Fukuda, Fault Detection Algorithm for External Thread Fastening by Robotic Manipulator Using Liner Support Vector Machine Classifier, in Proc. IEEE Int. Conf. Robot. Autom., (Karlsruhe, Germany, 2013), 3428-3435.
6. 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.
7. A. Nakamura and T. Kotoku, Systematization of error recovery in skill-based manipulation, Artificial Life and Robotics, Springer, 14(2) (2009) 203-208.
8. A. Nakamura, K. Nagata, K. Harada, N. Yamanobe, T. Tsuji, T. Foissotte and Y. Kawai, Error recovery using task stratification and error classification for manipulation robots in various fields, in Proc IEEE/RSJ Int. Conf. on Intell. Robots Syst., (Tokyo, Japan, 2013), 3535-3542.
9. 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.
10. T. Hasegawa, T. Suehiro and K. Takase, A model-based manipulation system with skill-based execution, IEEE Trans. Robot. Autom., 8(5) (1992) 535-544.
11. A. Nakamura, T. Ogasawara, T. Suehiro and H. Tsukune, Skill-based backprojection for fine motion planning, in Proc. IEEE/RSJ Int. Conf. on Intell. Robots Syst., (Osaka, Japan, 1996), 2, 526-533.
12. A. Nakamura, T. Ogasawara, K. Kitagaki and T. Suehiro, Using robust and simplified geometric models in skill-based manipulation, in Proc. IEEE/RSJ Int. Conf. on Intell. Robots Syst., (Hawaii, USA, 2001), 1, 138-145.

Real-time Planning Robotic Palletizing Tasks using Reusable Roadmaps

Takumi Sakamoto
Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama
Toyonaka, 560-8531, Japan

Kensuke Harada
Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama
Toyonaka, 560-8531, Japan
Artificial Intelligence Research Center, National Institute of Advanced Industrial Science and Technology (AIST),
2-3-36, Aomi,Koto-ku, Tokyo, 135-0064, Japan

Weiwei Wan
Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama
Toyonaka, 560-8531, Japan
Intelligent Systems Research Institute, Advanced Industrial Science and Technology (AIST), 1-1 Umezono
Tsukuba, 305-8560, Japan
E-mail: sakamoto@hlab.sys.es.osaka-u.ac.jp, harada@sys.es.osaka-u.ac.jp, wan@sys.es.osaka-u.ac.jp
www.roboticmanipulation.org

Abstract

This paper focuses on robotic motion planning for performing the palletizing or de-palletizing tasks. In such tasks, a robot usually iterates similar pick-and-place for several times. Considering such feature of the tasks, we propose two motion planning approaches named R-PRM and R-RRT* where both methods utilize the previously constructed roadmaps in the conventional PRM and RRT*, respectively. We experimentally confirm that both methods significantly save the calculation time needed for motion planning compared to the conventional planning methods.

Keywords: Palletizing, De-palletizing, Motion Planning, Path Planning, PRM, RRT*

1. Introduction

In a logistic process, human workers loads cargo boxes onto a pallet and to unload them from a pallet where such tasks are often called the palletizing and de-palletizing tasks, respectively. Since the palletizing tasks are known to be very labor intensive, this research proposes a robotic motion planning method for automating such tasks.

When a robot loads a lot of cargo boxes onto a pallet, a robot performs the pick-and-place for a number of times

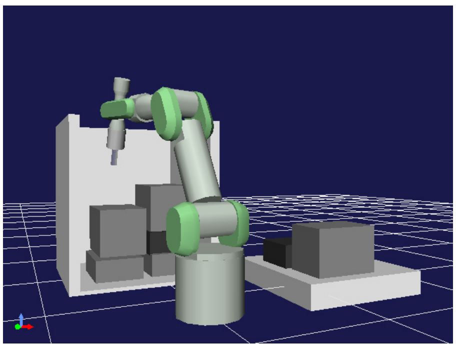


Fig. 1 . Image of palletizing and de-palletizing tasks

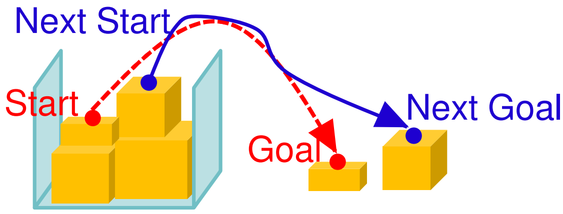


Fig. 2 . In a robotic palletizing task, a robot iterates similar pick-and-place for a number of times.

with slightly changing the initial and target configurations of the robot. Let us consider planning the motion of a robot performing the palletizing task by using a conventional sampling based motion planning method such as the PRM (Probabilistic Roadmap Method) and RRT (Rapidly-exploring Random Tree). Although the configurations of the robot and the environment around it are similar for each pick-and-place, a motion planner usually constructs a number of roadmaps and the constructed roadmaps are not used in the next pick-and-place. This results in unnecessarily increasing the calculation time needed for motion planning. Moreover, the calculation time of the random-sampling based motion planners such as PRM and RRT changes depending on each pick-and-place. It means that the calculation time may sometimes become extremely large if the roadmaps are placed at the wrong places in the initial planning phase.

To cope with this problem, this paper proposes an approach for robotic motion planning methods suitable for the palletizing tasks by re-using the roadmaps constructed in the previous pick-and-place. We propose two planning methods where one is named the R-PRM (Re-usable PRM) and the other is the R-RRT* (Re-usable RRT*). In both methods, we consider reducing the calculation time by re-using the previously constructed roadmaps. In R-PRM, we construct a graph of roadmap in the first pick-and-place. After the second pick-and-place, we just modify the roadmaps located around the initial and the target object pose. In the R-RRT*, we execute the RRT* in the first pick-and-place. After the second pick-and-place, we use the solution path obtained in the first pick-and-place. We confirm that both methods enable us to significantly reduce the calculation time after the second pick-and-place.

The rest of the paper is organized as follows: after explaining the related works in the section 2, we explain the proposed methods in the section 3. In the section 4, we confirm the effectiveness of our approach through experiments by using an industrial robot.

2. Related Works

In robotic motion planning, random-sampling based methods such as the PRM (Probabilistic Roadmap) and RRT (Rapidly-exploring Random Tree) have been widely used [1], [2], [3], [4]. To plan a smooth path, post processors such as [5], [6] and the motion planning methods considering the asymptotic optimality such as the RRT* [4] have been proposed.

While the conventional PRM and RRT assume the static obstacles, some methods on motion planner can be used under dynamically changing environments. We can roughly divide the researches into two groups. In the first group, the roadmap graph is modified according to the change of environments [7], [8], [9], [10], [11], [12]. Ferguson et al. [10] proposed the Anytime-RRT where they update an invalid part of the tree structure according to the change of environment. Anytime-RRT has been also used for the path planning of mobile robots [11], [12]. On the other hand, the second group, the solution path is locally modified according to the change of environment [13], [14]. Quinlan et al. [13] proposed a method for modifying the solution path according to the gradient of energy field. Combining the above two approaches, Yoshida et al. [15] proposed a method of humanoid motion planning that can deal with the change of environments.

On the other hand, this paper proposes a method of robotic motion planning for changing environments. Different from the previous approach, we consider utilizing a feature of the palletizing tasks where the environment only the position of the target object changes for each pick-and-place. Assuming such small change of environment, we consider modifying the roadmap graph.

3. Proposed Motion Planners

In this research, we propose two methods on robotic motion planning re-using the previously constructed roadmaps, i.e., P-PRM (Re-usable Probabilistic Roadmap) and R-RRT*(Re-usable Rapidly-exploring Random Tree Star). First, we explain a method for collision checking introduced in this research in the subsection 3.1. Then, we explain the R-PRM and R-RRT* in the subsections 3.2 and 3.3, respectively.

3.1. *Collision detection*

We tackle the motion planning problem of a robot moving in a dynamically changing environment. As an example of a dynamically changing environment, we consider a situation where a robot performs the pick-and-place for a number of times. In each pick-and-place, a robot picks an object from the stack and put it at a designated place. After a robot performs a pick-and-place, the configuration around the robot changes only due to the moved object. Since the change of configuration is not large, tt is not efficient to construct the roadmap graph from scratch for each pick-and-place. Rather, taking into account the configuration of moved object, we consider slightly modifying the previously constructed roadmaps. This subsection explains how to check collision around the part of the configuration space where the change occurred.

Fig. 3 shows how to check collision and to modify the previously constructed roadmap graph where the newly appeared obstacle in the configuration space is marked in gray. Since it is almost impossible to precisely obtain the shape of newly appeared obstacle in the configuration space, we first assume a hyper sphere in the configuration space as shown by the green circle. We set the radius of the hyper sphere large enough to include the newly appeared obstacle. We consider checking the collision of the roadmaps constructed in the previous pick-and-place included in this hyper sphere. If the collision is detected, we consider removing the corresponding roadmaps and edges connecting these roadmaps from the roadmap graph.

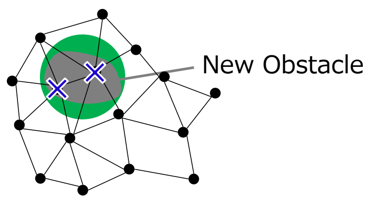


Fig. 3 Collision checking of roadmaps included in the hyper sphere

3.2. *R-PRM*

We first explain the R-PRM (Re-usable PRM) for reusing the previously constructed roadmap graph. The overview of the algorithm is shown in Fig. 4 where white and gray denote the free space and obstacle, respectively. Red and blue lines denote the previously found path and the path to be planned, respectively.

(i) Planning Initial Path
To plan the motion of a robot in the first pick-and-place, we use conventional PRM and memorize the constructed roadmap graph as shown in Fig. 4(a). The box placed at the designated place is regarded as the newly appeared obstacle when planning the next pick-and-place. By using the collision checking method explained in the previous subsection, we consider removing the inefficient edges and nodes included in the roadmap graph as shown in Fig. 4(b).

(ii) Planning Second and Subsequent Pick-and-place
To plan the motion of a robot in the second and subsequent pick-and-place, we consider re-using the previously constructed roadmap graph for the purpose of shortening the calculation time. Connecting the new start and goal configurations to the roadmap graph, the solution path is searched as shown in Fig. 4(c).

(iii) Path Smoothing
After finding the solution path, we apply the shortcut operation [6] for predefined times to smooth the solution path as shown in Fig. 4(d).

3.3. *R-RRT**

We next explain a single-query motion planner that re-uses the previously constructed roadmaps based on the RRT* with satisfying the asymptotic optimality.

(i) Planning Initial Path
To plan the motion of a robot in the first pick-and-place, we use the conventional RRT* as shown in Fig. 5(a). The roadmap tree is constructed from both initial and goal configurations. After finding the solution path, we remove the initial and the target configurations from the solution path. Let $\mathrm{T_{mid}}$ be the obtained path. When planning the next pick-and-place, the object placed in the previous pick-and-place is regarded to be an obstacle. We remove the roadmaps causing collision from $\mathrm{T_{mid}}$ due to the newly appeared obstacle as shown in Fig. 5(b).

(ii) Planning Second and Subsequent Pick-and-place
fter the second and subsequent pick-and-place, we first connect T_{mid} with the new initial configuration, and then connect T_{mid} with the new target configuration. Let T_s and T_g be the roadmap trees rooted at the initial and the target, respectively. First, by expanding both T_s and T_{mid}, we try to connect T_s

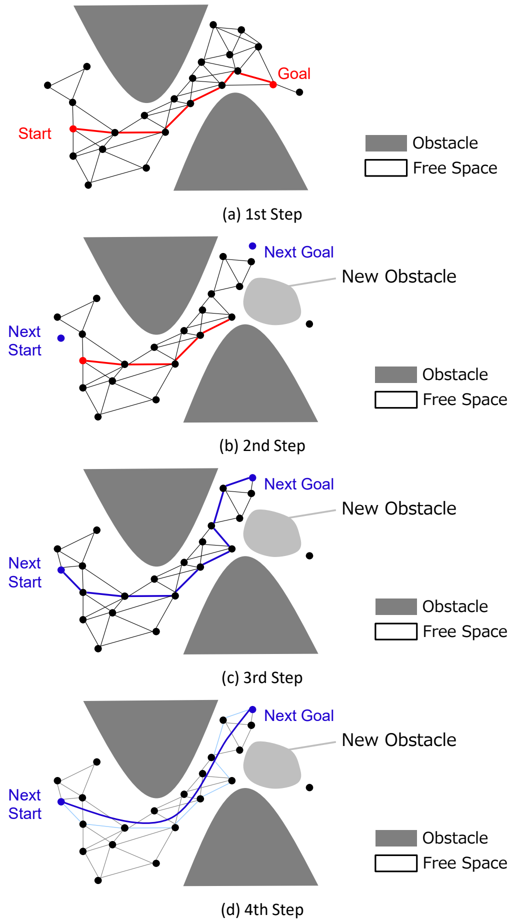


Fig. 4 . Path planning using R-PRM

and T_{mid} by using the RRT* as shown in Fig. 5(c). Then, by expanding both T_g and T_{mid} by using the RRT* as shown in Fig.5(d).

While we can guarantee the asymptotic optimality of the solution path by using the conventional RRT*, our proposed R-RRT* can approximately realize the optimal path if the new initial and target configurations are close enough to the previous initial and target configurations, respectively.

4. Experiments

This section shows experimental results of our proposed methods.

4.1. *Experimental Conditions*

We performed the following three palletizing experiments.

- EX1 Arrange the picked objects from the back to the front as shown in Fig. 6.

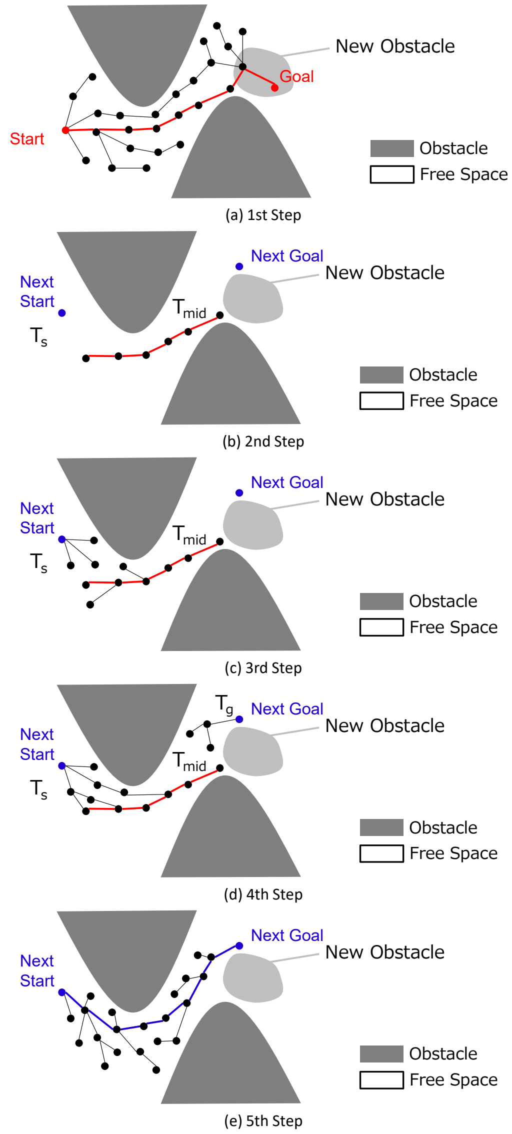


Fig. 5 . Path planning using R-RRT*

- EX2 Arrange the picked objects from the front to the back as shown in Fig. 7
- EX3 Place objects placed in a cage on the table under the table as shown in Fig. 8.

Fig. 6 . Overview of EX1

Fig. 7 . Overview of EX2

Fig. 8 . Overview of EX3

Snapshot of robot motion in Ex2 is shown in Fig. 9.

In each experiment, we tested PRM, R-PRM, RRT*, and R-RRT*. In case of R-PRM and R-RRT*, we planned the first pick-and-place by using RRM and R-RRT* and planned the second and subsequence pick-and-place by using R-PRM and R-RRT*, respectively. In each experiment, we conducted experiments on placing six objects for five times. We especially show the result of motion planning from the grasping configuration (Fig. 9(d)) to the placing configuration (9(h)). We used the uniform sampling in Ex1 and Ex2 while we used the Gaussian sampling in Ex3 since a large number of obstacles are included. When planning the motion by using PRM and R-PRM, we sampled the configuration space until we obtain 50 milestones. We smooth the solution path by applying the shortcut operation [6]. When planning the motion by using RRT* and R-RRT*, we terminated the planning algorithm if 10 feasible paths are planned or the path is updated for 100000 times in Ex1 and Ex2 and for300000 times in Ex3.

We used the industrial robot Nextage [16] equipped with a suction gripper in the right hand. To pick a box, the suction pad is placed at the upper surface of a box. We coded the planning algorithms on the robot simulation environment Choreonoid [17] and sent the motion command from Choreonoid to Nextage. To plan the robot motion, we used the PC with eight 3.10GHz CPUs.

4.2. *Results*

Figs. 10 and 11 show the comparison of calculation time between PRM and R-PRM and between RRT* and R-RRT*, respectively. In each figure, (a), (b) and (c) show the result of Ex1, Ex2 and Ex3, respectively. The vertical and the horizontal axes show the calculation time and the number of pick-and-place trial, respectively. We can see from these figures that the calculation time after second and subsequent pick-and-place has been improved in case of R-PRM and R-RRT* where the calculation time is less than 400[ms] for all the experiments. Variation of the calculation time appeared in PRM and RRT* is reduced in R-PRM and R-RRT*. In Ex1 and Ex2, all the algorithms did not fail in finding a solution. On the other hand, in Ex3, PRM failed in finding a solution for 9 times out of 30 times pick-and-place while R-PRM never failed in finding a solution. This is because R-PRM and R-RRT* re-uses the previous roadmaps and likely to successfully find a solution.

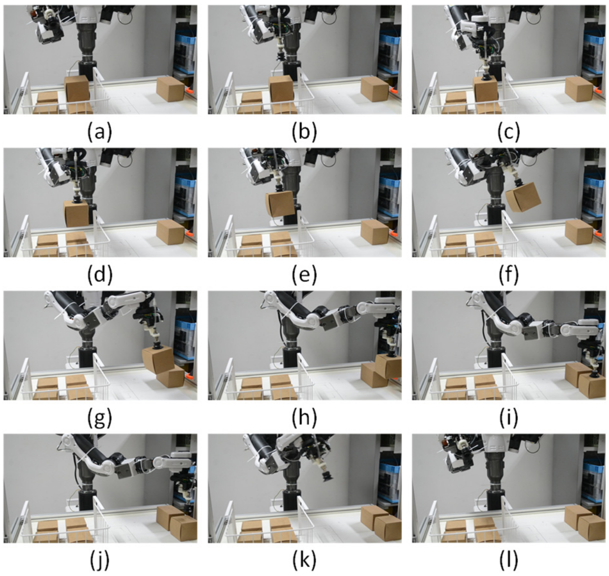


Fig. 9. Snapshot of robot motion in Ex2

5. Conclusions

In this paper, we proposed a robotic motion planning algorithms re-using the previously constructed roadmaps which is well applied for robotic palletizing tasks. We proposed two motion planning algorithms, i.e., R-PRM and R-RRT* based on PRM and RRT*, respectively. Through experiments, we confirmed that we can significantly reduce the calculation time and can reduce

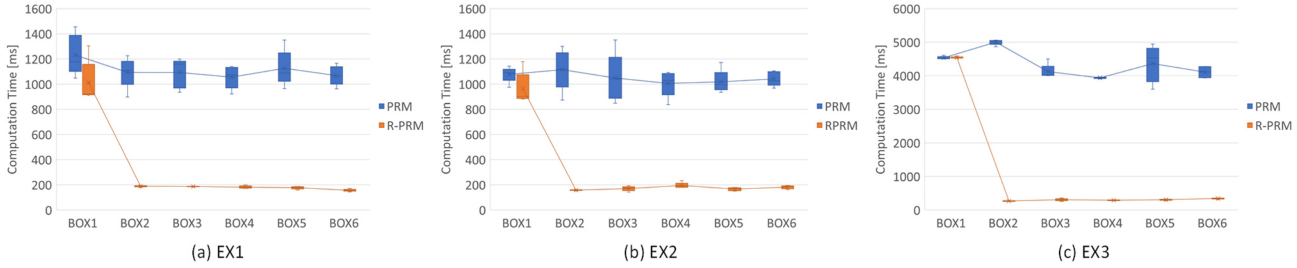


Fig. 10 . Results of PRM and R-PRM

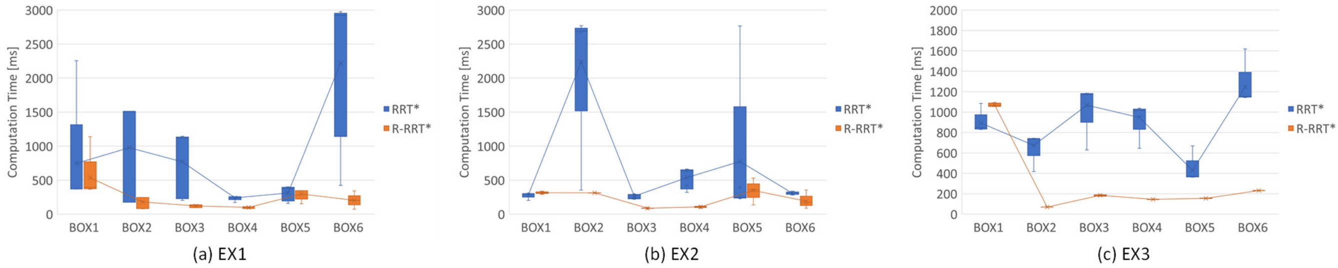


Fig. 11 . Results of RRT* and R-RRT*

the variation of the calculation time if we use our proposed algorithms. Application of our motion planning algorithms to other tasks is considered to be our future research topic.

References

1. L.E. Kavraki, P. Svestka, J.-C. Latombe, M.H. Overmars, Probabilistic roadmaps for path planning in high-dimensional configuration spaces, *IEEE Trans. on Robotics and Automation*, 12(4), 1996.
2. S.M. LaValle and J.J. Kuffner Jr., Randomized Kinodynamic Planning, *The Int. J. Robotics Research*, 20(5), pp. 378400, 2001.
3. J.J. Kuffner and S.M. LaValle, ”RRT-connect: An efficient approach to single-query path planning”, in *Proc. of IEEE Int. Conf. on Robotics and Automation*, 2000.
4. S.Karaman and E.Frazzoli, ”Sampling-based algorithms for optimal motion planning”, *The Int. J. of Robotics Research*, 2011, 30.7: 846-894.
5. R.Geraerts and M.H.Overmars, ”Creating High-Quality Paths for Motion Planning”, *The Int. J. of Robotics Research*, 2007, 30.7: 846-894.
6. H.Kris, N.T.Hing, Victor, ”Fast Smoothing of Manipulator Trajectories Using Optimal Bounded Acceleration Shortcuts, ”*In Robotics and Automation (ICRA)*, 2010 *IEEE Int. Conf. on, pages2493-2498. IEEE*, 2010.
7. P. Leven and S. Hutchinson, ”A framework for real-time path planning in changing environments,” *Int. J. of Robotics Research*, vol. 21, no.12, pp. 999-1030, 2002.
8. L. Jaillet and T. Simeon,“A PRM-based motion planner for dynamically changing environments, ”*Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2004.
9. J. Berg, D. Ferguson, J. Kuffner, “Anytime path planning and replanning in dynamic environments,” *Proc. IEEE Int. Conf. on Robotics and Automation,* pp. 2366-2371, 2006.
10. D. Ferguson and A. Stent, ”Anytime RRTs,” *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 5369-5375, 2006.
11. Y. Yang and O. Brock, ”Elastic roadmaps: Globally task-consistent motion for autonomous mobile manipulation in dynamic environments,” *Proc. Robotics: Science and Systems*, 2007.
12. Y. Kuwata, G. A. Fiore, J. Teo, E. Frazzoli, and J. P. How, ”Motion planning for urban driving using RRT,” in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 1681-1686, 2008.
13. S. Quinlan and O. Khatib, ”Elastic bands: Connecting path planning and control,” in *Proc. 1993 IEEE Int. Conf. on Robotics and Automation*, 1993, pp. 802-807.
14. O. Brock and O. Khatib, ”Elastic strips: A framework for motion generation in human environments,” *Int. J. of Robotics Research*, vol.21, no. 12, pp. 1031-1052, 2002.
15. E. Yoshida, K. Yokoi, and P. Gergondet, ”Online replanning for reactive robot motion: Practical aspects,” *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 5927-5933, 2010.
16. KAWADA Robotics Inc., https://www.kawadarobot.co.jp
17. Choreonoid, http://choreonoid.org/ja

No Free Lunch Principle in Agent Swarm Systems: One Case Study

Yunzhong Song†
School of Electrical Engineering and Automation, Henan Polytechnic University, 2001 Century Avenue Jiaozuo, 454003, 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-mail:songhpu@126.com, daifz@tust.edu.cn, xiaohm@huel.edu.cn, smfei@seu.edu.cn

Abstract

This note comes with information flooding of the control action of the leader in leader follower framework. In this situation, even observer for the leader velocity is built in a distributed style, the agent swarm system in a whole could not be classified into distributed one. One case study example will be borrowed to demonstrate the truth of the no free lunch principle in agent swarm systems.

Keywords: free lunch principle, agent swarm systems, distributed observer, leader-follower framework

1. Introduction

The agent swarm systems are recognized widely and this trend is still flooding[1]. Networks of networks for leader-follower system was explored by Song[2], which emphasizes interdependence among the agents. After that, the root node agents and the leaf node agents, could flock into each other, were touched upon further[3]. For agent swarm systems, on account of the interwoven between information and control, complexities are enlarged unprecedented[4-5]. Leader-follower agent swarms tracking control, was focused upon by several scientists[6-9].

To be set apart from the main results, this note will contribute to the neglected part of the dual roles of information and control. It is well known that the

† Corresponding author
© The 2020 International Conference on Artificial Life and Robotics (ICAROB2020), Jan. 13-16, B-Con Plaza, Beppu, Oita, Japan

distributed control is overly dependent on the neighbor based information exchange among agents, local control based regulation is always the main theme of the agent swarm systems. While besides that, the moving agents, especially the active leader one, must have its own velocity and added control action. For second order leader-follower agent swarms, the velocity of the leader agent was paid much more attention and its observers design is necessarily executed for its unavailable peculiarity for not all the follower agents. To be distinct, the added control action of the leader agent is assumed possible for all the follower agents, and this assumption is overly strong to be possible, for the availability of the added control action of the leader node is always impossible unless there is some broadcasting mechanism about that signal. This note will concentrate on the complexity balance between information and control of the agent swarm systems, the simplicity of one side will deteriorate on the other side. That is no free lunch principle in agent swarm systems.

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. Distributed Observers Design for Leader-Follower Agent Swarm systems

This part will begin with distributed observers design for leader-follower agent swam systems.

2.1. *Problem formulation*

Assume the agent swarm system is composed with one leader agent and n follower agents. Here, the graph G is used to describe the graph of the follower agents, which is undirected graph. We use G denotes the graph with the leader agent. Some agents in the follower ones will connect the leader one with directed arcs. Obviously, the G can be named as connected if at least one of the follower agents is connected with the leader one. To highlight the main theme of this note, the topology of the agent system is assumed to be static. One form of the leader-follower agent topology of this note can be listed as Figure 1. In Fig.1, we have two different connections from the leader agent to the follower agents, one is characterized as solid line with arrow, the other are characterized as discontinuous lines with arrows. Here, the solid line is employed to represent the physical proximity of the interacting agents, while the discontinuous lines are for information flooding from the leader to the followers, broadcasting can be one kind of mechanisms.

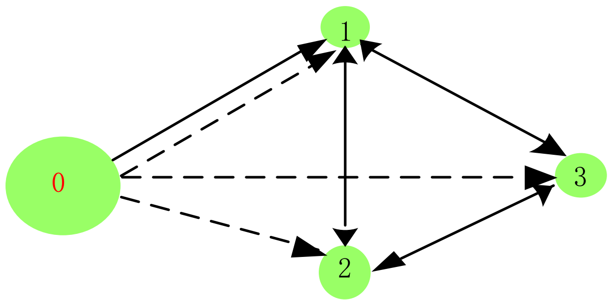


Fig.1 Network consists of 4 nodes

2.2. *The Main Results*

Leader agent is assumed to be active. The dynamic of the leader agent is expressed as:

$$\begin{cases} \dot{x}_0 = v_0, x_0 \in R^m, \\ \dot{v}_0 = u_0, u_0 \in R^m, \\ \quad y = x_0, \end{cases} \tag{1}$$

Where x_0, v_0 and y are position velocity and the only measurable variable, respectively. The dynamic of the follower agents is described as:

$$\begin{cases} \dot{x}_i = v_i, x_i \in R^m, \\ \dot{v}_i = u_i, u_i \in R^m, i = 1, \ldots, n, \end{cases} \tag{2}$$

The distributed observer demonstrated was in the following form:

$$\begin{cases} x_{ij} = x_i - x_j, x_{i0} = x_i - x_0, \\ u_i = u_0 - k(v_i - \hat{v}_i) \\ \qquad -l\Big(\sum_{j \in N_i(t)} a_{ij} x_{ij} + \sum_{i \in N_0(t)} b_i x_{i0}\Big), \\ \dot{\hat{v}}_i = u_0 - (l/k^2)\Big(\sum_{j \in N_i(t)} a_{ij} x_{ij} + \sum_{i \in N_0(t)} b_i x_{i0}\Big), \end{cases} \tag{3}$$

Where $N_i(t), N_0(t)$ denotes the neighbour agents of the follower agent i and the leader agent. a_{ij}, b_i is related with the topology of G. If agent j is in the neighbourhood of agent i, then $a_{ij} = 1$. Similarly, if the follower agent i is connected with the leader agent directly, then $b_i > 0$.

Comment 1: The protocol of (3) for (1) and (2) was put forward by Hong et al in [7], concentrated by Sarras in [8], and replied by Hong et al again in [9].

Comment 2: The seminal work of Hong et al in [7] declare that their protocol to be superior to the other ones in its simplicity in observer structure, while the dynamic of the system is in second order, the observer in distributed form can be in first order. That avoids the intricacy of the full order observer.

Comment 3: The communication between Hong et al and Sarras et al was only constrained in the coefficient of the protocol, without doubt, the selection of the coefficient is vitally important, and to do that, the research was promoted a lot, that can be seen clearly in [8] and [9].

Comment 4: To be set apart from the already existed results, no free lunch principle in leader follower agent swarms would be introduced, and without losing the simplicity of (3).

Comment 5: That availability u_0 for all the follower agents was too strong to be true. In fact, some extra mechanism must be taken to guarantee its availability, broadcasting can be one of them, but broadcasting easily suffer from information leakage and adversarial attacks.

Comment 6: Control simplicity and information complexity are always stand by each other. The gain of the control simplicity comes along with information intricacy. That is the no free lunch principle in leader-follower agent swarm systems. Our improved protocol is like:

$$\begin{cases} x_{ij} = x_i - x_j, x_{i0} = x_i - x_0, \\ u_i = u_0 - k(v_i - \hat{v}_i) \\ \quad -l(\sum\limits_{j\in N_i(t)} a_{ij}x_{ij} + \sum\limits_{i\in N_0(t)} b_i x_{i0}) + \bar{u}, \\ \dot{\hat{v}}_i = u_0 - (l/k^2)(\sum\limits_{j\in N_i(t)} a_{ij}x_{ij} + \sum\limits_{i\in N_0(t)} b_i x_{i0}) + \bar{u}, \end{cases} \tag{4}$$

Where we assume that $u = u_0 + \bar{u}$, and u_0 is virtually possible to drive the controller and the observer of the follower agents, $\bar{u}$ is added in to compensate this virtually dealt scheme, $\bar{u}$ satisfies $\|\bar{u}\| \le M \ge 0$.

Comment 7: The suggested protocol for the follower agents about the controller function of the leader agent, which is in the closed control loop, is feasible. This can be said in the following reasons: At first the on board sensor, actuator for the second integrator in the same series can be guessed on some sense, though with some errors, and second, the error induced disturbance can again be estimated by customized observer.

Lemma 1: *For leader (1) and followers (2), with $\bar{u} = 0$, the controller-observer pair (3) can produce*

$$\lim_{t\to\infty}|x_i(t) - x_0(t)| = 0, \lim_{t\to\infty}|v_i(t) - v_0(t)| = 0; \tag{5}$$

Theorem 1: There exists a constant $c_M > 0$ with $\lim\limits_{M\to 0} c_M = 0$, such that

$$\lim_{t\to\infty}|x_i(t) - x_0(t)| \le c_M, \lim_{t\to\infty}|v_i(t) - v_0(t)| \le c_M; \tag{6}$$

Proof: In case of convenience, set $\xi = (x_1,..,x_n)^T - x_0\mathbf{1}, \eta = (v_1,...,v_n)^T - v_0\mathbf{1},$ and $\zeta = k(v_1,...,v_n)^T - kv_0\mathbf{1}$, where $\mathbf{1} = (1,...,1)^T \in R^m$. Then the closed system can be turned into:

$$\begin{cases} \dot{\xi} = \eta, \\ \dot{\eta} = -l(L+B)\xi - k\eta + \zeta + \bar{u}, \\ \dot{\zeta} = \dfrac{l}{k}(L+B)\xi + \bar{u}, \end{cases} \tag{7}$$

Assume $z = \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix}$, then in a compact form, (7) can be casted into

$$\dot{z} = Fz = \begin{pmatrix} 0 & I & 0 \\ -lH & -kI & I \\ -\dfrac{l}{k}H & 0 & 0 \end{pmatrix} + \delta,$$

Where $H = L + B, \delta = (0, \bar{u}, \bar{u}, 0, ..., 0)^T$. For system (7), a control Lypunov function (CLF) can be constructed as $V(z) = z^T(t)Pz(t),$

where $P = \begin{pmatrix} kI & I & -\dfrac{k}{2}I \\ I & I & -\dfrac{1}{2}I \\ -\dfrac{k}{2}I & -\dfrac{1}{2}I & \dfrac{k}{2}I \end{pmatrix}$, which can be guaranteed positive if parameters l, k are properly chosen. To be detail, they satisfy the following inequalities

$l \ge 2/\lambda_{\min}, k \ge 4 + \lambda_{\max}$, where $\lambda_{\min}, \lambda_{\max}$ are the minimum and maximum eigenvalues of positive matrix H.

Consequently,

$$V(z(t)) \le e^{-\beta t} V(z(0)) + \beta_0 M^2. \quad (8)$$

This implies (6).

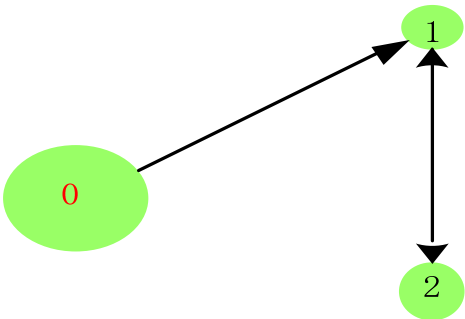


Fig.2 Case study example

Here, we use Fig.2 as the case study example, parameters are selected as:

$$u = u_0 + \bar{u}, u_0 = \cos t, \bar{u} = 0.02\cos t, l = 40, b = 0.2, k = 200.$$

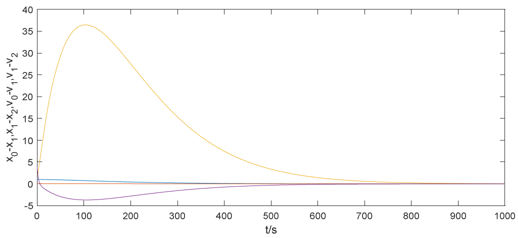


Fig.3 Position and velocity tracking errors of two followers

The corresponding results are demonstrated as Fig.3. From simulation results, we can say the suggested scheme is effective, which is with the virtual control function estimation and perturbation term of amendment.

2.3. *Further Comments*

Inspirations can be always possible if we pay much more attention to the interplay of information and control in agent swarm systems. No free lunch principle can be one example of them, though sometimes it is subtle.

3. Conclusion

No free lunch principle can be a big help for us to promote the research about leader-follower agent swarm systems.

System intricacy, could not be easily mastered if we are kept away from the concrete systems. Case study like high performance vehicle driving[10-11], can be the milestone systems. Time varying delays and other negative points, can also be the adversary factors. Adaptive scheme in Jia[12] provided a good example. Refined control of leader follower system according their roles of the leader agent, was recommended also[7]. So, balance between information and control can be an eternal topic in agent swarm systems.

Acknowledgements

This work is partially supported by NSFC Grant (61340041 and 61374079) and the Project-sponsored by SRF for ROCS, SEM to Yunzhong Song as well as Natural Science Fund of Henan Province (182300410112).

References

1. R. Olfati, R. Murray, Consensus problems in networks of agents with switching topology and time delays, *IEEE Transactions on Automatic Control*, **49**(2004) 1520-1533..
2. Y. Song, 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. Y. Hong, J. Hu, L. Gao, Tracking control for multi-agent consensus with an active leader and variable topology, *Automatica*, **42**(2006) 1177-1182.
7. Y. Hong, G. Chen, L. Bushenell, Distributed observers design for leader-following control of multi-agent netowrks, *Automatica*, **44**(2008) 846-850.
8. I. Sarras, P. Lahanier, J. Marzat, S. Bertrand, Comments on 'distributed observers design for leader-following control of multi-agent netowrks, Automatica, 44: 846-850, 2008', *Automatica*, **105**(2019)453-454.
9. Y. Hong, G. Chen, L. Bushenell, Authors' reply to comments on "distributed observers design for leader-following control of multi-agent netowrks", *Automatica*, **105**(2019) 455-455.
10. 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.
11. 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

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.

Analyzing the Controllability and Observability of Discrete-Time Delayed LTI Systems with Data-Based Methods

Zhuo Wang*

The Research Institute of Frontier Science, Beihang University, 37 Xueyuan Road, Haidian District Beijing, 100191, People's Republic of China

Chenghong Wang

Chinese Association of Automation, 95 Zhongguancun East Road, Haidian District Beijing, 100190, People's Republic of China

E-mail: zhuowang@buaa.edu.cn, chenghwang@163.com

www.buaa.edu.cn

Abstract

We present some data-based methods to analyze the state/output controllability and state observability of discrete-time delayed linear time-invariant (LTI) systems, whose parameter matrices are all assumed unknown. They first augment the system into a high dimensional LTI model, then employ the measured state/output data to directly construct the controllability and observability matrices of this high dimensional model, whose ranks are the criteria of corresponding characteristics of the original system. These data-based methods have low computational load and calculation complexity.

Keywords: Controllability and Observability, Data-Based Analyzing Methods, Discrete-Time Delayed LTI Systems, High Dimensional LTI Model, Measured State/Output Data.

1. Introduction

In the control theory and control engineering area, time-delay systems are a class of important research objects, since time delay is a universal and frequent phenomenon existing in practical life and industries.[1,2] The natural and artificial neural networks,[2–4] the multi-agent systems,[5,6] the electronic and optical systems,[7–10] etc. are some typical examples to have time delays. As an essential property of the real world, the complete prevention of time delays cannot be guaranteed, even by using the state-of-the-art technology. In this realistic circumstance, the extensive research on characteristics of systems with time delays, for instance, the stability and robustness,[11,12] the controllability and observability,[13–15] etc. is both theoretically and practically significant.

But when it is put into practice, people always encounter difficulties in both theoretical and practical areas. One typical difficulty is about the system dimension: when it is in the continuous-time domain, it can be regarded as an infinite dimensional system, which has a transcendental equation with infinite solutions as its characteristic equation[16–18]; when the system is discrete-time, its dimension increases dramatically with the increase of the time-delay terms.[1,5,19] In addition, nowadays modern industries have become so large-scale and complex that correctly building their dynamic models has become more and more difficult and even impractical.[20–22] All these problems have prevented analyzing the characteristics of time-delay systems.

*Corresponding Author

© The 2020 International Conference on Artificial Life and Robotics (ICAROB2020), Jan. 13-16, B-Con Plaza, Beppu, Oita, Japan

Nevertheless, it is fortunate that the modern sensor technologies, the wireless communication techniques, the computer and internet science and technologies, etc. have been rapidly developing and widely applied in many and increasingly more fields. People are having more and more effective, cheap and fast approaches to obtain various kinds of data from industrial production and daily life.[22,23] The information contained in the measured data usually reflect the intrinsic nature of the systems.[24,25] This situation is so clear in the big data era, that naturally leads to the idea of developing data-based methods to analyze the characteristics of time-delay systems only by utilizing the measured data.

To comply with the trends of the times, we have already conducted some investigations on the data-based property analysis of the discrete-time LTI systems with time delays, whose results unfortunately did not have good universality.[26] On the other hand, in most of the published research reports, the data-based characteristics analyzing methods could only deal with systems with particular delay structures, but may not be applied to the systems with a general form of time delays.

To overcome the above problems, in this paper we will develop a series of data-based methods for analyzing the state/output controllability and state observability of a general class of discrete-time delayed LTI systems. They can directly determine the system characteristics just by utilizing a simple augmented state-space model together with the measured historical and current testing data, without identifying the unknown system parameter matrices. We will also discuss the advantages of them from the aspects of identification workload, calculation precision and computational complexity, respectively.

2. Description of The Problem

This study concerns the problem of analyzing the state/output controllability and the state observability of the discrete-time delayed LTI systems, which can be represented by the following general dynamic model:

$$\begin{cases} x(k+1) = \sum_{i=0}^{N_x} A_i x(k-i) + \sum_{j=0}^{N_u} B_j u(k-j), \\ y(k) = \sum_{l=0}^{N_y} C_l x(k-l), \end{cases}$$

(1)

where $u \in \mathbb{R}^m$, $x \in \mathbb{R}^n$, $y \in \mathbb{R}^q$ are the input, the state and the output of system (1), respectively; the time delays $1 \le N_u, N_y \le N_x < \infty$ are integer constants and the time index $k \ge -N_x$; among each group of the parameter matrices $A_1, \ldots, A_{N_x} \in \mathbb{R}^{n\times n}$, $B_1, \ldots, B_{N_u} \in \mathbb{R}^{n\times m}$ and $C_1, \ldots, C_{N_y} \in \mathbb{R}^{q\times n}$, there is at least one nonzero matrix.

In this paper, we suppose that the above A_i, B_j and C_l are all unknown and none of them is random. Next below, we will propose a series of analysis methods for the determination of the above system characteristics without the necessity of identifying the parameter matrices, but just by utilizing the measured data. The advantage of these data-based characteristics determination methods will also be illustrated.

To begin with, system (1) is first expanded into a high dimensional LTI model by state-space augmentation:

$$\begin{cases} X(k+1) = \hat{A}X(k) + \hat{B}U(k) \\ y(k) = \hat{C}X(k) \end{cases} \quad (k \ge 0), \tag{2}$$

where $U(k) = \begin{bmatrix} u(k) \\ u(k-1) \\ \vdots \\ u(k-N_u) \end{bmatrix}$, $X(k) = \begin{bmatrix} x(k) \\ x(k-1) \\ \vdots \\ x(k-N_x) \end{bmatrix}$ are the input and the state of system (2), respectively; and the augmented parameter matrices are defined as

$$\hat{A} = \begin{bmatrix} A_0 & A_1 & A_2 & \cdots & A_{N_x} \\ I_n & 0 & 0 & \cdots & 0 \\ 0 & I_n & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & I_n & 0 \end{bmatrix},$$

$$\hat{B} = \begin{bmatrix} B_0 & B_1 & \cdots & B_{N_u} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \tag{3}$$

$$\hat{C} = \begin{bmatrix} C_0 & C_1 & \cdots & C_{N_y} & 0 & \cdots & 0 \end{bmatrix}.$$

where $\hat{A} \in \mathbb{R}^{n(N_x+1)\times n(N_x+1)}$, $\hat{B} \in \mathbb{R}^{n(N_x+1)\times m(N_u+1)}$, $\hat{C} \in \mathbb{R}^{q\times n(N_x+1)}$ and $I_n \in \mathbb{R}^{n\times n}$ is the identity matrix.

The equivalence between the state controllability, the output controllability as well as the state observability of system (2) and the respective characteristics of system (1) is obvious.

2.1. *Data-based state controllability criterion*

We start with the state controllability in this subsection, where a commonly used criterion is first given below.

Lemma 1. *The necessary and sufficient condition for LTI system $z(k+1) = Hz(k) + Gv(k)$ $(k \geq 0)$ with the input $v \in \mathbb{R}^m$ and the state $z \in \mathbb{R}^n$, to be completely state controllable, is*

$$rank([H^{n-1}G, H^{n-2}G, \ldots, HG, G]) = n.$$

Because the state controllability of system (2) and that of system (1) are equivalent, based on Lemma 1, we can determine this property by checking the rank of

$$W_C = [\hat{A}^{n(N_x+1)-1}\hat{B}, \ldots, \hat{A}\hat{B}, \hat{B}]. \quad (4)$$

In order to obtain the augmented controllability matrix W_C, the conventional analytical methods have to identify the unknown $A_0, A_1, \ldots, A_{N_x}$, $B_0, B_1, \ldots, B_{N_u}$ in system (1) first, and construct matrices $\hat{A}, \hat{B}$ as in (3), then calculate $\hat{A}\hat{B}, \hat{A}^2\hat{B}, \ldots, \hat{A}^{n(N_x+1)-1}\hat{B}$, and build W_C with these identification and calculation results. It is natural that the identification errors are often introduced in the above process, and the calculation workload is really heavy especially when the state dimension n and the time delay N_x in state are large.

In this circumstance, for the sake of avoiding the identification errors and also for reducing the workload in practice, people intend to find a more simple and direct way for the construction of W_C. We can make this goal achieved via the following data-based analyzing method.

This method starts with a set of $n(N_x+1)+1$ tests on system (2). In the first $n(N_x+1)$ tests, we design the input sequences as $U^{\{p\}}(0), \ldots, U^{\{p\}}(n(N_x+1)-1)$ $(p = 1, 2, \ldots, n(N_x+1))$, which can make the following $n(N_x+1)$ vectors

$$V^{\{p\}} = \begin{bmatrix} U^{\{p\}}(0) \\ U^{\{p\}}(1) \\ \vdots \\ U^{\{p\}}(n(N_x+1)-1) \end{bmatrix} \left(1 \leq p \leq n(N_x+1)\right)$$

linearly independent. Therefore, the matrix

$$\bar{V} = \left[V^{\{1\}}, \ldots, V^{\{n(N_x+1)\}}\right] \in \mathbb{R}^{mn(N_u+1)(N_x+1)\times n(N_x+1)} \quad (5)$$

has the rank of $n(N_x+1)$. On the other hand, the last test is done on system (2), where the input sequence is set as $U^{\{n(N_x+1)+1\}}(k) \equiv 0$.

Next, let all the corresponding states of system (2) have the same initial value:

$$X^{\{1\}}(0) = X^{\{2\}}(0) = \cdots = X^{\{n(N_x+1)+1\}}(0) = X_0.$$

Measure and record the values of $X^{\{1\}}(n(N_x+1))$,…, $X^{\{n(N_x+1)\}}(n(N_x+1))$ and $X^{\{n(N_x+1)+1\}}(n(N_x+1))$. Utilizing these special augmented state data, we present the following data-based state controllability criterion.

Theorem 1. *Define*

$$Z_X^{\{p\}} = X^{\{p\}}\left(n\left(N_x+1\right)\right) - X^{\{n(N_x+1)+1\}}\left(n\left(N_x+1\right)\right),$$
$$\left(1 \leq p \leq n\left(N_x+1\right)\right),$$
$$Z_X = \left[Z_X^{\{1\}}, Z_X^{\{2\}}, \ldots, Z_X^{\{n(N_x+1)\}}\right]. \quad (6)$$

Both system (1) and system (2) are completely state controllable, if and only if $\mathrm{rank}(Z_X) = n(N_x+1)$.

Proof. With the specifically designed $n(N_x+1)+1$ testing input sequences, the same initial state X_0, as well as the measured augmented state data, there should be

$$X^{\{n(N_x+1)+1\}}(n(N_x+1)) = \hat{A}^{n(N_x+1)}X_0,$$
$$X^{\{p\}}(n(N_x+1))$$
$$= \sum_{k=0}^{n(N_x+1)-1} \hat{A}^k\hat{B}U^{\{p\}}(n(N_x+1)-1-k) \quad (7)$$
$$+\hat{A}^{n(N_x+1)}X_0,$$
$$Z_X^{\{p\}} = \sum_{k=0}^{n(N_x+1)-1} \hat{A}^k\hat{B}U^{\{p\}}(n(N_x+1)-1-k),$$

for all $1 \leq p \leq n(N_x+1)$. Based on (4), (5), (6) and (7),

$$Z_X$$
$$= \left[\hat{A}^{n(N_x+1)-1}\hat{B}, \ldots, \hat{A}\hat{B}, \hat{B}\right] \times$$
$$\begin{bmatrix} U^{\{1\}}(0) & \cdots & U^{\{n(N_x+1)\}}(0) \\ \vdots & \ddots & \vdots \\ U^{\{1\}}\left(n\left(N_x+1\right)-1\right) & \cdots & U^{\{n(N_x+1)\}}\left(n\left(N_x+1\right)-1\right) \end{bmatrix}$$
$$= W_C\bar{V}.$$

Since $\bar{V}$ is of rank $n(N_x+1)$ due to the special design of $U^{\{p\}}(0)$,…, $U^{\{p\}}(n(N_x+1)-1)$ $\left(1 \leq p \leq n(N_x+1)\right)$, it can be inferred that

$$\mathrm{rank}\left(W_C\right) + \mathrm{rank}\left(\bar{V}\right) - n\left(N_x+1\right)$$
$$\leq \mathrm{rank}\left(W_C\bar{V}\right) \quad (8)$$
$$\leq \min\left\{\mathrm{rank}\left(W_C\right), \mathrm{rank}\left(\bar{V}\right)\right\} = \mathrm{rank}\left(W_C\right)$$

Consequently, $\mathrm{rank}\left(Z_X\right) = \mathrm{rank}\left(W_C\bar{V}\right) = \mathrm{rank}\left(W_C\right)$.

By Lemma 1, the sufficient and necessary condition for system (2) to be completely state controllable is $\operatorname{rank}(Z_X) = n(N_x + 1)$, and the same holds true for system (1). □

2.2. *Data-based output controllability criterion*

For the system characteristics determination, the output controllability is also frequently studied. This subsection will analyze the output controllability in a data-based perspective.

Lemma 2. *The LTI system*

$$\begin{cases} z(k+1) = Hz(k) + Gv(k) \\ w(k) = Fz(k) \end{cases} \quad (k \ge 0),$$

with the input $v \in \mathbb{R}^m$*, the state* $z \in \mathbb{R}^n$ *and the output* $w \in \mathbb{R}^q$*, is completely output controllable, if and only if*

$$\operatorname{rank}([FH^{n-1}G, FH^{n-2}G, \ldots, FHG, FG]) = q.$$

From Lemma 2, the output controllability of system (2), which is equivalent to that of system (1), can be verified by checking the rank of

$$W_{OC} = [\hat{C}\hat{A}^{n(N_x+1)-1}\hat{B}, \ldots, \hat{C}\hat{A}\hat{B}, \hat{C}\hat{B}]. \qquad (9)$$

Similarly as in the last subsection, for the same concern for avoiding the identification errors and reducing the calculation workload, we propose a data-based method to determine the output controllability, which would neither identify $A_0, \ldots, A_{N_x}$, $B_0, \ldots, B_{N_u}, \ldots$, and $C_0, \ldots, C_{N_y}$ nor compute $\hat{C}\hat{B}$, $\hat{C}\hat{A}\hat{B}, \ldots, \hat{C}\hat{A}^{n(N_x+1)-1}\hat{B}$.

We still conduct $n(N_x + 1) + 1$ groups of tests on system (2), as the prior preparation. First, special input sequences $U^{\{p\}}(0)$, $U^{\{p\}}(1)$, …, $U^{\{p\}}(n(N_x + 1) - 1)$ $(1 \le p \le n(N_x + 1))$ are designed for making $\bar{V}$ in (5) have full column rank. Then, another input sequence is also set as $U^{\{n(N_x+1)+1\}}(k) \equiv 0$, and the corresponding initial sates are set to have the same value as in the last subsection, i.e. $X^{\{1\}}(0) = \cdots = X^{\{n(N_x+1)+1\}}(0) = X_0$.

Then, we measure and record the respective outputs $y^{\{p\}}(n(N_x + 1))$ $(1 \le p \le n(N_x + 1) + 1)$, with which we can propose a theorem about the output controllability in the data-based perspective below.

Theorem 2. *Define*

$$Z_y^{\{p\}} = y^{\{p\}}(n(N_x+1)) - y^{\{n(N_x+1)+1\}}(n(N_x+1)),$$
$$(1 \le p \le n(N_x+1)),$$
$$Z_y = \left[Z_y^{\{1\}}, Z_y^{\{2\}}, \ \ldots, \ Z_y^{\{n(N_x+1)\}}\right]. \qquad (10)$$

Both system (1) and system (2) are completely output controllable, if and only if $\operatorname{rank}(Z_y) = q$.

Proof. With the aforementioned special $n(N_x + 1) + 1$ input sequences, the same initial state X_0, as well as the outputs $y^{\{1\}}(n(N_x + 1)), \ldots, y^{\{n(N_x+1)+1\}}(n(N_x + 1))$, we could obtain

$$y^{\{n(N_x+1)+1\}}(n(N_x + 1)) = \hat{C}\hat{A}^{n(N_x+1)}X_0,$$
$$y^{\{p\}}(n(N_x + 1))$$
$$= \textstyle\sum_{k=0}^{n(N_x+1)-1} \hat{C}\hat{A}^k\hat{B}U^{\{p\}}(n(N_x + 1) - 1 - k) \qquad (11)$$
$$+\hat{C}\hat{A}^{n(N_x+1)}X_0,$$
$$Z_y^{\{p\}} = \textstyle\sum_{k=0}^{n(N_x+1)-1} \hat{C}\hat{A}^k\hat{B}U^{\{p\}}(n(N_x + 1) - 1 - k),$$

for all $1 \le p \le n(N_x + 1)$. Hence, it can be inferred from (5), (9), (10) and (11) that

$$Z_y = [\hat{C}\hat{A}^{n(N_x+1)-1}\hat{B}, \ldots, \hat{C}\hat{A}\hat{B}, \hat{C}\hat{B}][V^{\{1\}}, \ldots, V^{\{n(N_x+1)\}}]$$
$$= W_{OC}\bar{V}. \qquad (12)$$

Similarly as in the proof of Theorem 1, $\operatorname{rank}(Z_y) = \operatorname{rank}(W_{OC}\bar{V}) = \operatorname{rank}(W_{OC})$, with $\bar{V}$ having full column rank. Therefore, based on Lemma 2, the necessary and sufficient condition for both system (2) and system (1) to be completely output controllable, is $\operatorname{rank}(Z_y) = q$. □

2.3. *Data-based state observability criterion*

After analyzing the state controllability and the output controllability, in this subsection we will study how to determine the state observability of both system (1) and system (2), only by utilizing the measured data. Similarly, a corresponding lemma is first given below as the start.

Lemma 3. *The LTI system*

$$\begin{cases} z(k+1) = Hz(k) + Gv(k) \\ w(k) = Fz(k) \end{cases} \quad (k \ge 0),$$

with the input $v \in \mathbb{R}^m$*, the state* $z \in \mathbb{R}^n$ *and the output* $w \in \mathbb{R}^q$*, is completely state observable, if and only if*

$$\operatorname{rank}\left(\begin{bmatrix} F \\ FH \\ \vdots \\ FH^{n-1} \end{bmatrix}\right) = n.$$

As observed previously, the state observability of system (2) is equivalent to that of system (1), and from Lemma 3, the state observability of system (2) can be determined by calculating the following matrix's rank.

$$W_{OB} = \begin{bmatrix} \hat{C} \\ \hat{C}\hat{A} \\ \vdots \\ \hat{C}\hat{A}^{n(N_x+1)-1} \end{bmatrix}. \quad (13)$$

Similar to the controllability cases, if people want to avoid introducing identification errors and performing a large amount of calculations, it is better to develop an analyzing method for determining the state observability only with the measured data, just like in the previous subsections.

This novel method still begins with some tests on system (2), where we set $n(N_x+1)$ linearly independent initial states $X^{\{1\}}(0)$, $X^{\{2\}}(0)$,…, $X^{\{n(N_x+1)\}}(0)$ first. Such that the following matrix

$$\Phi_X = \left[X^{\{1\}}(0), X^{\{2\}}(0), \ldots, X^{\{n(N_x+1)\}}(0) \right] \quad (14)$$

has rank of $n(N_x+1)$. Then, set all their inputs as

$$U^{\{1\}}(k) = U^{\{2\}}(k) = \cdots = U^{\{n(N_x+1)\}}(k) \equiv 0. \quad (15)$$

At the meantime, measure the respective outputs $y^{\{1\}}(k)$, $y^{\{2\}}(k)$,…, $y^{\{n(N_x+1)\}}(k)$ at the sampling instants $k = 0, 1, \ldots, n(N_x+1)-1$. With these data, define

$$Y_p = \begin{bmatrix} y^{\{p\}}(0) \\ y^{\{p\}}(1) \\ \vdots \\ y^{\{p\}}(n(N_x+1)-1) \end{bmatrix} \ \big(1 \le p \le n(N_x+1)\big),$$

$$Y = \left[Y_1, Y_2, \ldots, Y_{n(N_x+1)}\right] \in \mathbb{R}^{qn(N_x+1)\times n(N_x+1)}. \quad (16)$$

Next, a data-based criterion on the state observability will be presented below.

Theorem 3. *Suppose that the initial states of system (2) can be set as in (14). Then, system (1) and system (2) are both completely state observable, if and only if*

$$\text{rank}(Y) = n(N_x+1),$$

where Y is defined in (16).

Proof. With the constant input sequences $U^{\{p\}}(k) \equiv 0$ $\big(1 \le p \le n(N_x+1)\big)$, the respective output sequences of system (2) should be

$$\begin{cases} y^{\{p\}}(0) = \hat{C}X^{\{p\}}(0) \\ \quad\vdots \\ y^{\{p\}}(k) = \hat{C}\hat{A}^k X^{\{p\}}(0) \\ \quad\vdots \\ y^{\{p\}}\big(n(N_x+1)-1\big) = \hat{C}\hat{A}^{n(N_x+1)-1} X^{\{p\}}(0) \end{cases} \quad (17)$$

for all $1 \le p \le n(N_x+1)$. From (13), (14), (16) and (17), we can obtain $Y_p = W_{OB}X^{\{p\}}(0)$ and $Y = W_{OB}\Phi_X$. Since $\text{rank}(\Phi_X) = n(N_X+1)$, then

$$\begin{aligned} &\text{rank}(W_{OB}) + \text{rank}(\Phi_X) - n(N_x+1) \\ &\le \text{rank}(W_{OB}\Phi_X) \\ &\le \min\{\text{rank}(W_{OB}), \text{rank}(\Phi_X)\} \\ &= \text{rank}(W_{OB}). \end{aligned} \quad (18)$$

As a result, $\text{rank}(Y) = \text{rank}(W_{OB}\Phi_X) = \text{rank}(W_{OB})$. Therefore, based on Lemma 3, system (2) is completely state observable if and only if $\text{rank}(Y) = n(N_x+1)$, which is also the necessary and sufficient condition for the complete state observability of system (1) due to the equivalence between the two systems. □

Similar to the data-based state/output controllability determination methods in the foregoing subsections, the data-based analyzing method for state observability in Subsection 2.3 has a few merits:

(i) it can determine the state observability of discrete-time delayed LTI systems either with (at least one of $A_1, \ldots, A_{N_x}, C_1, \ldots, C_{N_y}$ in system (1) is nonzero) or without time delay;
(ii) it does not need to identify these parameter matrices, such that the relevant identification errors can be avoided;
(iii) it does not need to calculate $\hat{C}\hat{A}$, …, $\hat{C}\hat{A}^{n(N_x+1)-1}$ for the construction of the observability matrix W_{OB} in (13), which greatly reduces the workload.

3. Discussion on Calculation Precision and Computational Complexity

As aforementioned, when analyzing the general discrete-time delay LTI system (1), the traditional model-based approaches have to identify all the parameter matrices first, and in principle, the identification results could be expressed as

$$\begin{aligned}\tilde{A}_i &= A_i + \Delta A_i \quad (0 \le i \le N_x),\\ \tilde{B}_j &= B_j + \Delta B_j \quad (0 \le j \le N_u),\\ \tilde{C}_l &= C_l + \Delta C_l \quad (0 \le l \le N_y),\end{aligned} \tag{19}$$

where ΔA_i, ΔB_j and ΔC_l denote the corresponding matrix identification errors. With them, we can construct the augmented matrices $\tilde{A}$, $\tilde{B}$ and $\tilde{C}$ like $\hat{A}$, $\hat{B}$ and $\hat{C}$ in (3), respectively.

In this section, we will first take the data-based state controllability determination method given in Theorem 1 as a representative, to illustrate the advantage of higher calculation precision. Since having the same advantage, discussions on the rest data-based determination methods for the output controllability and the state observability, are omitted here for brevity.

Recall the state equation in (1). For the convenience of discussion, we apply a numerical example below, where $m = 1$, $n = 2$ and $N_u = N_x = 1$:

$$x(k+1) = A_0 x(k) + A_1 x(k-1) + B_0 u(k) + B_1 u(k-1),$$

$$A_0 = \begin{bmatrix} 0.1 & -0.4 \\ 0.3 & -0.2 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0.8 & 0.5 \\ -0.5 & 0.6 \end{bmatrix},$$

$$B_0 = \begin{bmatrix} 0 \\ 0.1 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 0.1 \\ 0 \end{bmatrix}.$$

Suppose the followings are the errors introduced during the identification process:

$$\Delta A_0 = \begin{bmatrix} -0.03 & 0.08 \\ -0.07 & 0.04 \end{bmatrix}, \quad \Delta A_1 = \begin{bmatrix} -0.08 & -0.06 \\ 0.06 & -0.07 \end{bmatrix},$$

$$\Delta B_0 = \begin{bmatrix} 0.01 \\ -0.06 \end{bmatrix}, \quad \Delta B_1 = \begin{bmatrix} -0.04 \\ -0.01 \end{bmatrix}.$$

Consequently, $\tilde{A}_0$, $\tilde{A}_1$, $\tilde{B}_0$ and $\tilde{B}_1$ are calculated by (19), and the respective augmented parameter matrices $\tilde{A}$ and $\tilde{B}$ are constructed as in (3).

The conventional analyzing methods that are based on model identification, will utilize $\tilde{A}$ and $\tilde{B}$ to obtain the state controllability matrix $\tilde{W}_C = [\tilde{A}^3\tilde{B}, \tilde{A}^2\tilde{B}, \tilde{A}\tilde{B}, \tilde{B}]$ for the sake of checking its rank. In this example, we have

$$\tilde{W}_C = \begin{bmatrix} -0.0134 & 0.0249 & 0.0253 & 0.0344 & -0.0121 & 0.0074 & 0.01 & 0.06 \\ 0.0066 & 0.018 & 0.0147 & -0.0325 & -0.0041 & 0.0154 & 0.04 & -0.01 \\ 0.0253 & 0.0344 & -0.0121 & 0.0074 & 0.01 & 0.06 & 0 & 0 \\ 0.0147 & -0.0325 & -0.0041 & 0.0154 & 0.04 & -0.01 & 0 & 0 \end{bmatrix}.$$

If we round off the elements of $\tilde{W}_C$ as those in A_0, A_1, B_0 and B_1, then $\operatorname{rank}(\tilde{W}_C) = 2 < 4$, and the system will be determined not completely state controllable, which is not conformed with the real case. On the contrary, if we apply the data-based state controllability determination method given in Theorem 1, where A_0, A_1, B_0 and B_1 are not identified, the identification errors ΔA_0, ΔA_1, ΔB_0, ΔB_1 will then not occur.

Recall (5), set

$$\begin{aligned}\mathrm{V}^{\{1\}} &= [1,\ 0,\ 0,\ 0,\ 0,\ 1,\ 1,\ 0]^T,\\ \mathrm{V}^{\{2\}} &= [0,\ 0,\ 0,\ 1,\ 1,\ 1,\ 1,\ 1]^T,\\ \mathrm{V}^{\{3\}} &= [1,\ 1,\ 1,\ 0,\ 0,\ 0,\ 0,\ 1]^T,\\ \mathrm{V}^{\{4\}} &= [0,\ 1,\ 1,\ 1,\ 1,\ 0,\ 0,\ 0]^T.\end{aligned} \tag{20}$$

So $\operatorname{rank}(\bar{V}) = \operatorname{rank}([V^{\{1\}}, V^{\{2\}}, V^{\{3\}}, V^{\{4\}}]) = 4$, satisfying the linear independency requirement on the inputs. Then, we can obtain

$$Z_X = \begin{bmatrix} -0.0574 & 0.0511 & 0.054 & 0.069 & -0.04 & 0.01 & 0 & 0.1 \\ 0.0138 & 0.0443 & 0.052 & -0.053 & -0.02 & 0.03 & 0.1 & 0 \\ 0.054 & 0.069 & -0.04 & 0.01 & 0 & 0.1 & 0 & 0 \\ 0.052 & -0.053 & -0.02 & 0.03 & 0.1 & 0 & 0 & 0 \end{bmatrix},$$

where Z_X is defined in (6), and in this case we have $\operatorname{rank}(Z_X) = 4 = n(N_x + 1)$. From Theorem 1, it can be inferred that the system is completely state controllable, which obviously is the real case. If we continue similarly rounding off the elements of Z_X as those in A_0, A_1, B_0

and B_1, Z_X will still have full rank, and the analyzing conclusion will remain the same.

This example implies that our data-based analyzing methods have strong robustness against the identification errors, and therefore have higher calculation precision than the conventional model-based approaches. In the particular situation that the highest-order minor of the controllability (observability) matrix is close to zero, the analyzing conclusion drawn by utilizing the developed data-based determination methods, will not be influenced by the accumulated round off errors. These advantages are especially obvious when the dimensions m, n, q and the maximum delays N_u, N_x, N_y of system (1) are large, since there will be much heavier identification workload and much more accumulated calculation errors.

Furthermore, the lower computational complexity as another advantage is the next to be discussed. Here, to simplify the discussion, we regard both a summation and a multiplication of two matrix elements as one operation. For the general discrete-time delayed LTI system (1), when we multiply $\hat{A}^i$ $(1 \le i \le n(N_x+1)-2)$ with $\hat{A}$, there should be $n(N_x+1)$ multiplication operations and $n(N_x+1)-1$ summation operations to compute each element. Because the number of elements in $\hat{A}^{i+1}$ is $n^2(N_x+1)^2$, there are $n^2(N_x+1)^2(2nN_x+2n-1)$ operations in total, to obtain $\hat{A}^{i+1}$.

When $\hat{A}^k$ $(1 \le k \le n(N_x+1)-1)$ multiplies $\hat{B}$, there are $n(N_x+1)$ multiplications and $n(N_x+1)-1$ summations for each calculated element, while there are $mn(N_u+1)(N_x+1)$ elements in the obtained $\hat{A}^k\hat{B}$. As a consequence, people need to perform $mn(N_u+1)(N_x+1)(2nN_x+2n-1)$ operations. And in general, there are totally (nN_x+n-1) such kind of matrix multiplications. To compute the state controllability matrix $W_C = \left[\hat{A}^{n(N_x+1)-1}\hat{B}, \ldots, \hat{A}\hat{B}, \hat{B}\right]$, people perform

$$mn(N_u+1)(N_x+1)(2nN_x+2n-1)(nN_x+n-1)$$
$$+\sum_{i=1}^{n(N_x+1)-2} i\,n^2(N_x+1)^2(2nN_x+2n-1)$$

operations. In summary, the conventional model-based state controllability analyzing methods have the computational complexity of $\mathcal{O}\left(n^5{N_x}^5\right)$.

As a contrast, by observing (5) and (6), our data-based state controllability analyzing method only does $n(N_x+1)+1$ groups of experiment tests together with $n^2(N_x+1)^2$ element subtractions. So that the developed data-based analyzing method only has the computational complexity of $\mathcal{O}\left(n^2{N_x}^2\right)$.

By extension, it can be inferred that our data-based analyzing methods have lower computational complexity than the conventional model-based methods.

4. Conclusion and Future Work

This paper developed some novel data-based methods to analyze and determine the state/output controllability and the state observability of a general class of discrete-time delayed LTI systems.

These data-based analyzing methods conducted some specific input tests, and employed the measured historical and current data to directly construct the system matrix, the state/output controllability matrix as well as the state observability matrix, which are utilized to determine the corresponding characteristics. Generally speaking, they have merits in three major aspects over the conventional model-based approaches: the less identification work, the higher calculation precision and the lower computational complexity.

Our data-based methods present a new perspective on characteristics analysis of time-delay systems, yet they still have shortcomings for improvement. For the current stage, they can only analyze discrete-time LTI systems, but cannot deal with continuous-time systems, nonlinear systems or time-variant systems. Besides, if the system is stochastic, especially when there exist random/unknown noises contained in the system structure, we have no other way but to identify all the parameter matrices and the developed data-based methods are no longer feasible for the analyzing job. In the future researches, we would like to study these problems and hope to find suitable methods to overcome the difficulties.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China (Grant No. 61673041), and BAQIS Research Program #18G34.

References

1. L. Zou, Z. Wang, H. Gao and X. Liu, State estimation for discrete-time dynamical networks with time-varying delays and stochastic disturbances under the round-robin protocol, *IEEE Trans. Neural Netw. Learn. Syst.* **28**(5), 2017, pp. 1139–1151.
2. S. Yin, H. Yang, H. Gao, J. Qiu and O. Kaynak, An adaptive NN-based approach for fault-tolerant control of nonlinear time-varying delay systems with unmodeled dynamics, *IEEE Trans. Neural Netw. Learn. Syst.* **28**(8), 2017, pp. 1902–1913.
3. J. Wang, X. M. Zhang and Q. L. Han, Event-triggered generalized dissipativity filtering for neural networks with

time-varying delays, *IEEE Trans. Neural Netw. Learn. Syst.* **27**(1), 2016, pp. 77–88.

4. Y. W. Wang, W. Yang, J. W. Xiao and Z. G. Zeng, Impulsive multisynchronization of coupled multistable neural networks with time-varying delay, *IEEE Trans. Neural Netw. Learn. Syst.* **28**(7), 2017, pp. 1560–1571.
5. Y. Chen, D. W. C. Ho, J. Lu and Z. Lin, Convergence rate for discrete-time multiagent systems with time-varying delays and general coupling coefficients, *IEEE Trans. Neural Netw. Learn. Syst.* **27**(1), 2016, pp. 178–189.
6. D. Wang, N. Zhang, J. Wang and W. Wang, A PD-like protocol with a time delay to average consensus control for multi-agent systems under an arbitrarily fast switching topology, *IEEE Trans. Cybern.* **47**(4), 2016, pp. 898–907.
7. Z. Huszka and A. Chakravorty, Implementation of delay-time based nonquasi-static bipolar transistor models in circuit simulators, *IEEE Trans. Electron Devices*. **61**(8), 2014, pp. 3004–3006.
8. J. Chen, X. Zhang and J. Peng, Time-delayed chaotic circuit design using all-pass filter, *IEEE Trans. Circuits Syst. I, Reg. Papers*. **61**(10), 2014, pp. 2897–2903.
9. L. Sliwczynski and P. Krehlik, Multipoint joint time and frequency dissemination in delay-stabilized fiber optic links, *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*. **62**(3), 2015, pp. 412–420.
10. H. Guerboukha, A. Markov, H. Qu and M. Skorobogatiy, Time resolved dynamic measurements at THz frequencies using a rotary optical delay line, *IEEE Trans. Terahertz Sci. Technol.* **5**(4), 2015, pp. 564–572.
11. Z. Meng, W. Xia, K. H. Johansson and S. Hirche, Stability of positive switched linear systems: weak excitation and robustness to time-varying delay, *IEEE Trans. Autom. Control*. **62**(1), 2017, pp. 399–405.
12. T. Chen and X. Liu, μ-Stability of nonlinear positive systems with unbounded time-varying delays, *IEEE Trans. Neural Netw. Learn. Syst.* **28**(7), 2017, pp. 1710–1715.
13. Z. Lijun and Z. Kuize, Controllability and observability of boolean control networks with time-variant delays in states, *IEEE Trans. Neural Netw. Learn. Syst.* **24**(9), 2013, pp. 1478–1484.
14. S. Lam and E. J. Davison, Computation of the real controllability radius and minimum-norm perturbations of higher-order, descriptor, and time-delay LTI systems, *IEEE Trans. Automat. Contr.* **59**(8), 2014, pp. 2189–2195.
15. F. Li, J. Sun and Q.-D. Wu, Observability of boolean control networks with state time delays, *IEEE Trans. Neural Netw.* **22**(6), 2011, pp. 948–954.
16. M. T. Ozgun and M. Torlak, Effects of random delay errors in continuous-time semi-digital transversal filters, *IEEE Trans. Circuits Syst. I, Reg. Papers*. **61**(1), 2014, pp. 183–190.
17. F. Cacace, F. Conte and A. Germani, Filtering continuous-time linear systems with time-varying measurement delay, *IEEE Trans. Automat. Contr.* **60**(5), 2015, pp. 1368–1373.
18. L. Zou, Z. Wang, H. Gao and X. Liu, Event-triggered state estimation for complex networks with mixed time delays via sampled data information: the continuous-time case, *IEEE Trans. Cybern.* **45**(12), 2015, pp. 2804–2815.
19. J. Shen and J. Lam, On the decay rate of discrete-time linear delay systems with cone invariance, *IEEE Trans. Automat. Contr.* **62**(7), 2017, pp. 3442–3447.
20. G. Scarciotti and A. Astolfi, Model reduction of neutral linear and nonlinear time-invariant time-delay systems with discrete and distributed delays, *IEEE Trans. Automat. Contr.* **61**(6), 2016, pp. 1438–1451.
21. Y. Zhao, A. Fatehi and B. Huang, A data-driven hybrid ARX and Markov chain modeling approach to process identification with time-varying time delays, *IEEE Trans. Ind. Electron*. **64**(5), 2017, pp. 4226–4236.
22. W. Yan, D. Tang and Y. Lin, A data-driven soft sensor modeling method based on deep learning and its application, *IEEE Trans. on Ind. Electron*. **64**(5), 2017, pp. 4237–4245.
23. H. Sun, Z. Wang, J. Wang, Z. Huang, N. Carrington and J. Liao, Data-driven power outage detection by social sensors, *IEEE Trans. Smart Grid*. **7**(5), 2016, pp. 2516–2524.
24. S. Yin, Y. Jiang, Y. Tian and O. Kaynak, A data-driven fuzzy information granulation approach for freight volume forecasting, *IEEE Trans. Ind. Electron.* **64**(2), 2017, pp. 1447–1456.
25. A. Marinoni and P. Gamba, Unsupervised data driven feature extraction by means of mutual information maximization, *IEEE Trans. Comput. Imag.* **3**(2), 2017, pp. 243–253.
26. Binquan Zhou, Zhuo Wang, Yueyang Zhai, Heng Yuan, Data-Driven Analysis Methods for Controllability and Observability of A Class of Discrete LTI Systems with Delays, *Proceedings of 2018 IEEE the 7th Data Driven Control and Learning Systems Conference*, Enshi, Hubei Province, P. R. China, May 2018, pp. 380–384.

An Optimal Collective Control Strategy Based on Vicsek Model

Yongnan Jia*, Qing Li, Yue Liu, 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
www.ustb.edu

Abstract

Vicsek model is one of the most famous models taking advantage describing the collective behaviors of self-propelled particles by using simple interaction rules. These particles update their directions according to the average value of their neighbors, who located in its communication range. However, as we all know, individual difference exists widely among social animals. Different individuals may have different influence on the same particle. Therefore, we propose an optimal method considering the individual difference when updating the directions of each particle. According to the simulation results, compared with the standard Vicsek model, the optimal model is more efficient under specific conditions.

Keywords: Collective control, vicsek model, optimization.

1. Introduction

Inspired by the collective behaviors of social animals such as flocks of birds, school of fish, and swarm of bees, collective Control of multi-agent system has been widely studied for its potential applications on underwater robotic system[1], sensor network[2], and even aerial swarm[3-5].

Vicsek model is one of the most famous mechanisms that are introduced to reappear the collective behavior of multiple self-propelled particles[6]. The most important contribution of Vicsek model is the introduction of noise for the physical model. And they have proved that the noise is beneficial to the collective behavior of these self-propelled particles. Since its first appearance, various of variants of Vicsek models have been proposed[7-10]. However, seldom literatures have mentioned that how to improve the convergence effect of the collective behavior. In this paper, we propose an optimal strategy, based on which the stability of the multi-agent system can be improved.

We all know that the individual differences exist widely[11]. In Vicsek model, particles are supposed to be equal, which doesn't exactly conform to the reality. However, how these differences exist, we still do not know. Therefore, we propose a kind of optimal model from the evolutionary point of view. We suppose that the rank of each particle is decided by the number of its neighbors. If one particle can communicate with more particles, it has a higher level during the system, which means that it has a higher influence factor to decide the behavior at the next time step. We also compared this optimal strategy with the traditional one, the simulation results tell us that the optimal strategy can improve the stability of the whole system under bigger noise, only for a range of group sizes.

This paper is organized as follows: the Vicsek model and its optimal version are presented in section II. Section

*Corresponding author.

III shows the simulation results and illustrates the advantage of the proposed optimal strategy for collective control problem. Finally, this paper is concluded in section IV.

2. Modeling

In this paper, N particles are considered to move freely without any boundary limitation. The initial position and direction of N particles are generated randomly. For collective tasks, these particles are expected to move not only in order state but also coherently over time.

Without losing generality, suppose that the time interval during the simulation process is $\Delta t = 1$. At $t = 0$, N particles were randomly distributed within an area of a given size L and have the same absolute velocity v_d as well as randomly distributed directions θ. At each time step, the position of the ith particle is updated according to the following equation

$$\mathrm{x}_i(t+1) = \mathrm{x}_i(t) + \mathrm{v}_i(t)\Delta t \tag{1}$$

In each time step, the velocity of a particle $\mathrm{v}_i(t+1)$ is updated by

$$\mathrm{v}_i(t+1) = \mathrm{v}_i^{align}(t) + \mathrm{v}_i^{rep}(t) + \mathrm{v}_i^{adh}(t) \tag{2}$$

Where $\mathrm{v}_i^{align}(t)$ is the alignment term, $\mathrm{v}_i^{rep}(t)$ is the repulsion term, and $\mathrm{v}_i^{adh}(t)$ is the attraction term.

2.1. *Vicsek model*

The alignment term was constructed to have an absolute value v 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

$$< \theta_i(t) \geq arctan\left(\frac{\sum_{j=1}^{N} l_{ij}(t) sin\left(\theta_j(t)\right)}{\sum_{j=1}^{N} l_{ij}(t) cos\left(\theta_j(t)\right)}\right) \tag{4}$$

Neighbor matrix $L_N(t) = \left[l_{ij}(t)\right]_{N\times N}$ describes the neighbor relationships of particles at time t, where

$$l_{ij}(t) = a_{ij}(t), \forall i,j = 1,\cdots,N \tag{5}$$

The definition of adjacency matrix $A_N(t) = \left[a_{ij}(t)\right]_{N\times N}$ is

$$a_{ij}(t) = \begin{cases} 1, & i = 1,\cdots,N, j \in N_i(t) \\ 0, & otherwise \end{cases} \tag{6}$$

where $N_i(t) = \{j \mid \|\mathrm{x}_i(t) - \mathrm{x}_j(t)\| \leq r\}$. Here, r denotes the interaction radius. Using the above expressions the alignment term can be written as

$$\mathrm{v}_i^{align}(t+1) = \mathrm{c}^{\mathrm{align}} v \mathrm{e}_i(t) \tag{7}$$

where $\mathrm{c}^{\mathrm{align}}$ is the coefficient of the alignment term. $\mathrm{e}_i(t)$ is a unit vector with direction angle $\theta_i^{align}(t)$.

The repulsion term exists only when the distance between any two particles is smaller than the repulsive radius r_{rep}. And the repulsion term is defined as

$$\mathrm{v}_i^{rep}(t+1) = \mathrm{c}^{rep} \sum_{j=1}^{N} \left(\frac{r_{rep} - \|\mathrm{x}_{ij}(t)\|}{r_{rep}} \cdot \frac{\mathrm{x}_{ij}(t)}{\|\mathrm{x}_{ij}(t)\|}\right) \tag{8}$$

where $\|\mathrm{x}_{ij}(t)\| \leq r_{rep}$ and c^{rep} is the coefficient of the repulsion term.

The attraction term is only considered for the boundary particles[12] of the whole system when the distance between two particles is between r_{rep} and r_{adh}.

$$\mathrm{v}_i^{adh}(t+1) = \mathrm{c}^{adh} \sum_{j=1}^{N} \left(\frac{r_{rep} - \|\mathrm{x}_{ij}(t)\|}{r_{adh} - r_{rep}} \cdot \frac{\mathrm{x}_{ij}(t)}{\|\mathrm{x}_{ij}(t)\|}\right) \tag{9}$$

where $r_{rep} \leq \|\mathrm{x}_{ij}(t)\| \leq r_{adh}$ and c^{adh} is the coefficient of the attraction term. This term is introduced in order to prevent the flock spreading due to noises.

2.2. *Optimal Vicsek model*

For the optimal Vicsek model, the repulsion term and the attraction term are same as Vicsek model, while we try to optimize the Vicsek model according to amend the alignment term. We redefine the Neighbor matrix $L_N(t) = \left[l_{ij}(t)\right]_{N\times N}$ as

$$l_{ij}(t) = c_{ij}(t) * a_{ij}(t), \forall i,j = 1,\cdots,N \tag{10}$$

Where $c_{ij}(t) = \frac{B[j](t)}{B[i](t)}$. $B[i](t)$ denotes the number of neighbors of particle i at time t.

From an evolutionary point of view, at the very beginning, the individual differences exist, and the rank of each particle is different, that is a hierarchical system. However, following with the time passes, under the proposed control strategy, the system becomes in order and forms a stable cohesive formation. As shown in Fig. 1, the numbers of the neighbors will follow such a rule: the number of the neighbors becomes small from the central agent to the boundary agent. That means the rank of the particle in the system also becomes in order from chaotic law.

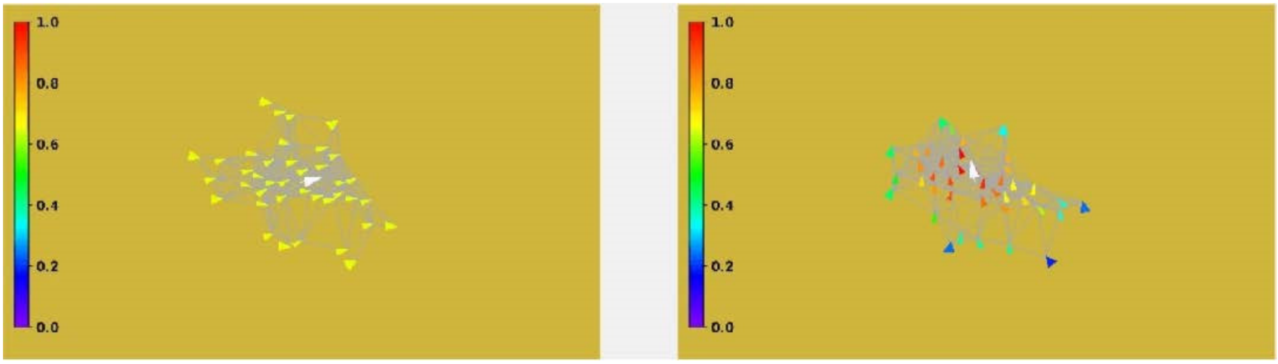


Fig. 1. Graphical representation of Vicsek model and optimal Vicsek model for 40 particles (Left one: Vicsek model; Right one: Optimal Vicsek model.

3. Simulation Results

In order to give the results more generally, simulations are carried out under several group sizes for 10 to 160. Suppose that the density of the system is defined as

$$\rho = \frac{N}{L^2} \tag{11}$$

For different group sizes, they all have the same initial density. Each experiment will be repeated n times according to the following rules:

$$n = \begin{cases} 1000, & if\ N = 10 \\ 1000, & if\ N = 20 \\ 500, & if\ N = 40 \\ 200, & if\ N = 80 \\ 100, & if\ N = 160 \end{cases} \tag{12}$$

In order to measure the stability of the particle group, we used the following average normalized velocity as the order parameter

$$\phi^{ave} = \frac{1}{T}\frac{1}{N}\int_0^T \|\sum_{i=1}^{N} \mathrm{v}_i(t)\| dt \tag{13}$$

where $T = 1000$ is the simulation time for each experiment.

The simulation results are shown in Fig. 2. The black and dashed line denotes egalitarian strategy (Vicsek model), while the blue solid line denotes the hierarchical strategy (optimal Vicsek model). Note that simulation results show that the egalitarian system is almost the same as the simple contribution-based hierarchical system (optimal Vicsek model) when the noise is small. Following with the increase of the group size, the simple contribution-based hierarchical system is more stable than the egalitarian system when the noise is big. However, the advantage of the optimal Vicsek model decreases when the group size is 80. For group size 160, the difference between the two models almost disappears.

4. Conclusion

In this paper, we have investigated the stability of a kind of optimal Vicsek model compared with the classical Vicsek model by using quantitative analysis method. The control protocol consists of three parts, alignment term, repulsion term and attraction term. The only difference between the two models is only the coefficient when calculating the average direction of each particle for the next time step. Average velocity is chosen as the order parameter to evaluate the stability of the multiple particles system. Simulation results illustrate that when the group size is around 100, the optimal Vicsek model will show advantage in stability than the classical Vicsek model under big noise. However, it seems that the coefficients will influence the simulation results. We would like to recover more details on the relationship between coefficients and the stability of the system in the next paper.

Acknowledgements

This work was partly supported by the following grants: National Natural Science Foundation of China (Grant No. 61603362); Fundamental Research Funds for the Central Universities (No. FRF-TP-19-031A2).

References

1. Jia Y. and Wang L., Experimental implementation of distributed flocking algorithm for multiple robotic fish. *Control Engineering Practice*, 2014, **30**, 1–11.
2. Tang Y., Hu Y., Cui J., Liao F., Lao M., Lin F. and Teo R.S.H., Vision-aided multi-uav autonomous flocking in gps-denied environment. *IEEE T. Ind. Electron.*, 2019, **66**(1), 616–626.

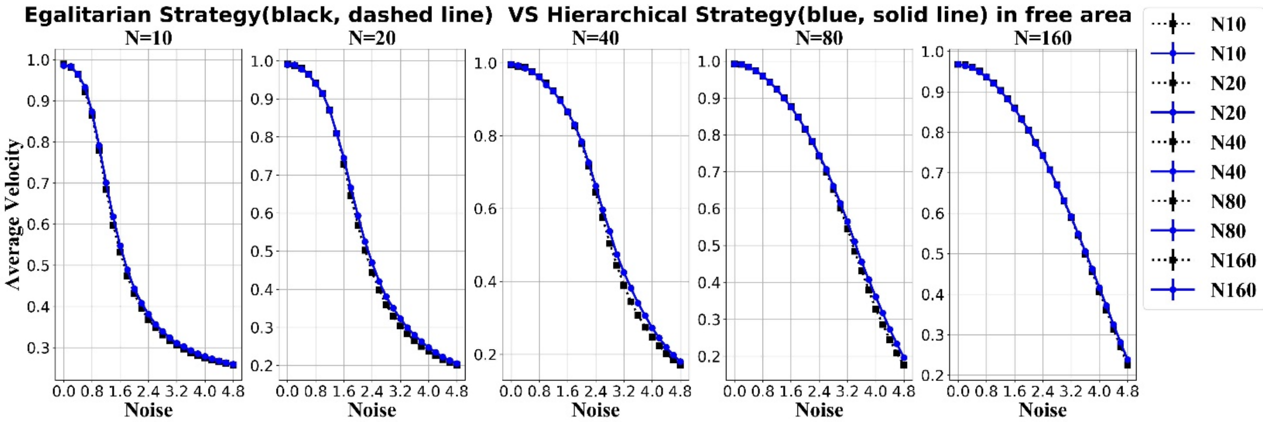


Fig. 2. Quantitative comparison of Vicsek model and optimal Vicsek model for various group sizes.

3. Bayraktar S., Fainekos G.E. and Pappas G.J., Experimental cooperative control of fixed-wing unmanned aerial vehicles. *Proc. 43rd IEEE Conf. Decision Contr.*, 2004, **4**, 4292–4298.
4. Brambilla M., Ferrante E., Birattari M. and Dorigo M., Swarm robotics: A review from the swarm engineering perspective. *Swarm Intelligence*, 2013, **7**(1), 1–41.
5. Jia Y., Swarming coordination of multiple unmanned aerial vehicles in three-dimensional space. In AIAA Modeling and Simulation Technologies Conference, AIAA AVIATION Forum, 2016.
6. Vicsek T., Czirok A., Ben-Jacob E., Cohen I. and Shochet O., Novel type of phase transition in a system of selfdriven particles. *Physical Review Letters*, 1995, **75**(6), 1226.
7. Vicsek T. and Zafeiris A., Collective motion. *Physics Reports*, 2012, **51**7(3), 71–140.
8. Vasarhelyi G., Viragh C., Somorjai G., Nepusz T., Eiben A.E. and Vicsek T., Optimized flocking of autonomous drones in confined environments. *Science Robotics*, **3**(20), 2018.
9. Huepe C. and Aldana M., New tools for characterizing swarming systems: A comparison of minimal models. *Physica A: Statistical Mechanics & Its Applications*, 2008, **387**(12), 2809–2822.
10. Wei W., Wentao D., Suzanne A., Ayusman S. and Mallouk T.E., From one to many: dynamic assembly and collective behavior of self-propelled colloidal motors. *Accounts of Chemical Research*, 2015, **48**(7), 1938–1946.
11. Zafeiris A. and Vicsek T., Why we live in hierarchies? A quantitative treatise. Springer, Berlin, 2018.
12. Pearce D.J.G. and Turner M.S., Density regulation in strictly metric-free swarms. *New Journal of Physics*, **16**, 2014.

Distributed Rotating Encirclement Control of Strict-Feedback Multi-Agent Systems Using Bearing Measurements

Tengfei Zhang, Yingmin Jia*

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: tfzhang@buaa.edu.cn, ymjia@buaa.edu.cn

www.buaa.edu.cn

Abstract

This paper investigates the distributed multi-target rotating encirclement formation problem of strict-feedback multi-agent systems using the targets' bearing angles and the agents' known positions, where all agents are forced to achieve even circular formation around the targets' geometric center. Firstly, an estimator is proposed for each agent to localize the neighbor targets. Secondly, based on the trajectory planning method, a reference trajectory is constructed by three estimators, which are used to obtain the targets' geometric center, the reference rotating radius and angular. Then, the proposed adaptive neural dynamic surface control law forces each agent to move along the reference trajectory, which satisfies the multi-target rotating encirclement formation conditions.

Keywords: Strict-feedback multi-agent systems, rotating encirclement control, target localization, trajectory planning, trajectory tracking

1. Introduction

Recent years the rotating encirclement formation problem of multi-agent systems have attracted considerable attention due to its significant potential applications in both military and civilian areas such as surveillance, search-and-rescue, reconnaissance, etc[1]. Many interesting results have been achieved for the rotating formation or surrounding/encirclement control problem[1-7].

As one of the most important high-order systems, the strict-feedback system can be used to model a variety of physical systems including robotic manipulators, vessel, unmanned aerial vehicle and so on[8]. And recent years have witnessed the emergence of researches with respect to the strict-feedback single/multi-agent system[8-11]. However, there is no research to date on the rotating encirclement control of high-order multi-agent system.

Motivated by above discussion, for the first time, we consider the multi-target rotating encirclement formation problem of strict-feedback multi-agent systems, and only bearing measurements of targets can be obtained. To this end, we divide the problem into three subproblems: target localization, trajectory planning and trajectory tracking. Four estimators are designed to construct a reference trajectory for each agent, and an adaptive neural dynamic surface control law is proposed to make the agent move along the desired trajectory.

2. Preliminaries and Problem Statement

2.1. *Graph Theory*

Let $\mathcal{G}(\mathcal{V},\mathcal{E},\mathcal{A},\mathcal{B})$ be a weighted undirected graph corresponding to n agents and m targets, where $\mathcal{V}=\{v_1,v_2,\cdots,v_n,s_1,\cdots,s_m\}$ denotes the set of vertexes, $\mathcal{E}\subset$

* Corresponding author.

$\mathcal{V} \times \mathcal{V}$ denotes the set of edges, $\mathcal{A} = [a_{ij}] \in \mathrm{R}^{n\times n}$ denotes the weighted adjacency matrix of targets, $\mathcal{B} = [b_{ik}] \in \mathrm{R}^{n\times m}$denotes the weighted adjacency matrix from targets to agents. Let $d(v_i, v_j)$ denote the shortest distance from the vertex v_i to v_j, for instance, $d(v_i, v_j) = 1$ if $(v_i, v_j) \in \mathcal{E}$. The neighbor agents set of the agent v_i is denoted by $\mathcal{N}_i = \{v_j \in \mathcal{V} | (v_i, v_j) \in \mathcal{E}\}$ and the neighbor targets set of the agent v_i is denoted by $\mathcal{N}_i^{\mathcal{T}} = \{s_k \in \mathcal{V} | (v_i, s_k) \in \mathcal{E}\}$. The neighbor agents set of the target s_k is denoted by $\mathcal{N}_k^{\mathcal{I}} = \{v_i \in \mathcal{V} | (v_i, s_k) \in \mathcal{E}\}$. And let $|\cdot|$ denote the number of elements in the set $\cdot$.

2.2. *Problem Statement*

Consider a multi-agent system consisting of n agents (Index set $\mathcal{I} = \{1,2,\cdots,n\}$) and m stationary targets (Index set $\mathcal{T} = \{1,2,\cdots,m\}$) with bearing-only measurements, where the dynamic of agent v_i is in the following q_i-order strict-feedback form.

$$\begin{cases} \dot{x}_{ij} = f_{ij}(\bar{x}_{ij}) + x_{ij+1} \\ \dot{x}_{iq_i} = f_{iq_i}(\bar{x}_{iq_i}) + u_i \\ y_i = x_{i1} \end{cases} \tag{1}$$

where $\bar{x}_{ij} = [x_{i1}^{\mathrm{T}}, \cdots, x_{ij}^{\mathrm{T}}]^{\mathrm{T}}$, and $\bar{x}_{iq_i}, y_i, u_i \in \mathrm{R}^2$ represent the states, output and control input of agent v_i, respectively. $f_{ij}(\bar{x}_{ij})$ is an unknown continuous nonlinear function.

The *objective* of this note is to design the distributed control scheme using bearing-only measurements of targets and the neighbor position information of agents, such that strict-feedback agents are capable of achieving the multi-target rotating encirclement formation, which is properly formulated by Definition 1 using the polar coordinate transformation using the polar coordinate transformation $y_i = \bar{r} + [l_i\cos(\theta_i), l_i\sin(\theta_i)]^{\mathrm{T}}$.

Definition 1[5]**.** The multi-agent system is said to achieve the *multi-target rotating encirclement formation* if

$$\lim_{t\to\infty} \left[l_i - \lambda \max_{k\in\mathcal{T}} \| r_k - \bar{r} \| \right] = 0 \tag{2}$$

$$\lim_{t\to\infty} \left[\theta_i - \theta_j - \tfrac{2\pi(i-j)}{n} \right] = 0 \tag{3}$$

$$\lim_{t\to\infty} \left[\dot{\theta}_i - \omega \right] = 0 \tag{4}$$

Where $i, j \in \mathcal{I}$, r_k and $\bar{r}(t) = 1/m \sum_{k\in\mathcal{T}} r_k$ denote the position of the k-th target and the geometric center of all targets respectively. The design parameter $\lambda > 1$ determines the radius of the desired rotation formation. And ω is the desired angular velocity.

To facilitate the later control design and analysis, we make some reasonable assumptions.

Assumption 1. *All agents are connected in some undirected communication topologies and each target connectes to at least one agent via the directed edge.*

Assumption 2. *The radius of the desired rotation formation is bounded, i.e., there exists a positive constant d^* satisfying* $\max_{k\in\mathcal{T}} \| r_k - \bar{r} \| \le d^*$.

Assumption 3. *The desired angular velocity ω and angular acceleration $\dot{\omega}$ are continuous and bounded, i.e., there exists positive constants ω^* and ω_d^* such that* $\|\omega\| \le \omega^*, \|\dot{\omega}\| \le \omega_d^*$.

3. Control Design

In this section we present in detail the distributed multi-target rotating encirclement control scheme, which includes three parts: target localization, trajectory planning and trajectory tracking.

3.1. *Target Localization*

To estimate the neighbor target's position of agent v_i with bearing-only measurements, the following estimator is proposed according to Ref. 7.

$$\dot{\hat{r}}_{ik} = \alpha_{ik}(I - \varphi_{ik}\varphi_{ik}^{\mathrm{T}})(x_{i1} - \hat{r}_{ik}) \tag{5}$$

Where $k \in \mathcal{N}_i^{\mathcal{T}}$, α_{ik} is a positive design parameter, and φ_{ik} is the unit vector from x_i to r_k, i.e., $\varphi_{ik} = (r_k - x_i)/\|r_k - x_i\|$.

3.2. *Trajectory Planning*

For each agent v_i, to plan a reference trajectory satisfying the Definition 1, we design the following distributed estimators to obtain the estimations $p_i, \hat{l}_i$ and $\hat{\theta}_i$ of the desired geometric center $\bar{r}$, polar radius l_i and polar angle θ_i, respectively.

$$\begin{cases} \dot{p}_{ik} = \beta_i \sum_{j\in\mathcal{N}_i} a_{ij}[p_{jk} - p_{ik}] \\ \qquad +\beta_i b_{ik}[\hat{r}_{ik} - p_{ik}] \\ p_i = \frac{1}{m}\sum_{k=1}^{m} p_{ik} \end{cases} \tag{6}$$

$$\begin{cases} \dot{\rho}_{i1} = \gamma_{i1} \max_{k\in\mathcal{N}_i^{\mathcal{T}}}(\|\hat{r}_{ik} - p_i\|) - \rho_{i1} \\ \dot{\rho}_{i2} = \gamma_{i2} \max_{j\in\mathcal{N}_i \cup\{i\}}(\|\rho_{j1}\|) - \rho_{i2} \\ \quad \vdots \\ \dot{\rho}_{iM} = \gamma_{iM} \max_{j\in\mathcal{N}_i \cup\{i\}}(\|\rho_{jM-1}\|) - \rho_{iM} \\ \hat{l}_i = \lambda\rho_{iM} \end{cases} \tag{7}$$

$$\dot{\hat{\theta}}_i = \delta_i \sum_{j\in\mathcal{N}_i} a_{ik} \left[\hat{\theta}_j - \hat{\theta}_i - \tfrac{2\pi(j-i)}{n} \right] + \omega \tag{8}$$

Where $\beta_i, \gamma_{i1}, \cdots, \gamma_{iM}, \delta_i$ are positive design parameters, and $M = \max_{i,j\in\mathcal{I}}\{d(i,j)\}$, which can be chosen as $M = n - 1$ if it is not prior information.

Then, with the polar coordinate transformation, the reference trajectory of agent v_i is provided as follows.

$$\hat{y}_i = p_i + \left[\hat{l}_i\cos(\hat{\theta}_i), \hat{l}_i\sin(\hat{\theta}_i)\right]^{\mathrm{T}} \quad (9)$$

3.3. *Trajectory Tracking*

Similar to the backstepping-based DSC design procedure, we define dynamic surface errors as follows.

$$\begin{cases} z_{i1} = x_{i1} - \hat{y}_i \\ z_{ij} = x_{ij} - \hat{\eta}_{ij} \end{cases} \quad (10)$$

Where $\hat{\eta}_{ij}(t)$ is the first-order filter estimation of the virtual controller $\eta_{ij}(t)$ with the time constant $\tau_{ij} > 0$ and the filter error is denoted by $\tilde{\eta}_{ij} = \hat{\eta}_{ij} - \eta_{ij}$.

$$\tau_{ij}\dot{\hat{\eta}}_{ij} + \hat{\eta}_{ij} = \eta_{ij}, \hat{\eta}_{ij}(0) = \eta_{ij}(0) \quad (11)$$

Then, we will present the following virtual and actual controllers and adaptive law such that each agent moves along its desired reference trajectory.

$$\begin{cases} \eta_{ij} = -\kappa_{ij}z_{ij} - \widehat{W}_{ij}^{\mathrm{T}}S_{ij}(\zeta_{ij}) \\ u_i = -\kappa_{iq_i}z_{iq_i} - \widehat{W}_{iq_i}^{\mathrm{T}}S_{iq_i}(\zeta_{iq_i}) \end{cases} \quad (12)$$

$$\dot{\widehat{W}}_{ij} = -\Gamma_{ij}^{-1}\left[\sigma_{ij}\widehat{W}_{ij} - S_{ij}(\zeta_{ij})z_{ij}^{\mathrm{T}}\right] \quad (13)$$

Where $\Gamma_{ij} = \Gamma_{ij}^{\mathrm{T}} > 0$ is an adaptive gain matrix, $\widehat{W}_{ij}$ and $S_{ij}(\zeta_{ij})$ represent the estimation of the optimal weight matrix W_{ij} and the basis function vector respectively. And κ_{ij}, σ_{ij} are positive design parameters.

4. Main Results

With the proposed control scheme in Section 3, we can easily obtain the following reasonable results.

Lemma 1. *Consider the estimator (5) under Assumptions 1-2. Then for any $i \in \mathcal{I}, k \in \mathcal{T}$, the estimation position $\hat{r}_{ik}$ will asymptotically converge to the actual position r_k of the k-th target.*

Proof. The proof is similar to Theorem 3.1 in Ref. 7. □

Then, we define the estimation of the targets' geometric center as follows.

$$\hat{\bar{r}} = \frac{1}{m}\sum_{i=1}^{n}\sum_{k\in\mathcal{N}_i^{\mathcal{T}}}\frac{1}{|\mathcal{N}_k^{\mathcal{I}}|}\hat{r}_{ik} \quad (14)$$

Apparently, $\hat{\bar{r}}$ will asymptotically converge to the actual geometric center $\bar{r}$.

Lemma 2. *Consider the estimator (6) under Assumptions 1-2. For any $i \in \mathcal{I}$, the estimation position p_i will asymptotically converge to $\hat{\bar{r}}$.*

Proof. The proof is similar to Lemma 4 in Ref. 5.. □

Then, combining lemma 1 with lemma 2, we can conclude that the estimation position p_i of the i-th agent will asymptotically converge to the actual geometric center $\bar{r}$.

Lemma 3. *Consider the estimator (7) under Assumptions 1-2. For any $i \in \mathcal{I}$, the following equation holds.*

$$\lim_{t\to\infty}\left[\hat{l}_i - \lambda\max_{i\in\mathcal{I},k\in\mathcal{N}_i^{\mathcal{T}}}(\|\hat{r}_{ik} - p_i\|)\right] = 0 \quad (15)$$

In other words, the estimation $\hat{l}_i$ of the polar radius will asymptotically converge to the above value.

Proof. The proof is similar to Lemma 5 in Ref. 5. □

Furthermore, with Lemma 1 and Lemma 2, it is easy to see that $\lim_{t\to\infty}\left[\hat{l}_i - \lambda\max_{k\in\mathcal{T}}(\|r_k - \bar{r}\|)\right] = 0$, which implies that $\hat{l}_i$ satisfies the condition (2).

Lemma 4. *Consider the estimator (8) under Assumptions 1-3. For any $i, j \in \mathcal{I}$, the following equations hold.*

$$\lim_{t\to\infty}\left[\hat{\theta}_i - \hat{\theta}_j - \frac{2\pi(i-j)}{n}\right] = 0 \quad (16)$$

$$\lim_{t\to\infty}\left[\dot{\hat{\theta}}_i - \omega\right] = 0 \quad (17)$$

In other words, the estimation $\hat{\theta}_i$ of the polar angle satisfies conditions (3) and (4).

Proof. The proof is similar to Lemma 6 in Ref. 5. □

Thus, from the polar coordinate transformation of (9), we can easily conclude that the reference trajectory $\hat{y}_i$ satisfies conditions of the multi-target rotating encirclement formation in Definition 1.

Then, we will carry on the stability analysis of the proposed control scheme (12) and (13), which drives the output y_i of the agent v_i to the reference trajectory $\hat{y}_i$. Since the neural network is capable of approaching any continual nonlinear function with free precision (See Lemma 2 in Ref. 8 for details), we make the following reasonable approximation.

$$\begin{aligned} \phi_{i1} &= W_{i1}^{\mathrm{T}}S_{i1}(\zeta_{i1}) + \varepsilon_{i1}(\zeta_{i1}) \\ &= f_{i1}(\bar{x}_{i1}) - \dot{\hat{y}}_i \\ \phi_{ij} &= W_{ij}^{\mathrm{T}}S_{ij}(\zeta_{ij}) + \varepsilon_{ij}(\zeta_{ij}) \\ &= f_{ij}(\bar{x}_{ij}) + \tilde{\eta}_{ij}/\tau_{ij} + z_{ij-1} \end{aligned} \quad (18)$$

Where $\|\varepsilon_{ij}(\zeta_{ij})\| < \varepsilon_{ij}^*$, ε_{ij}^* is an arbitrarily small constant and denote $\widetilde{W}_{ij} = W_{ij} - \widehat{W}_{ij}$.

Choose a common Lyapunov function candidate as $V_i = \frac{1}{2}\sum_{j=1}^{q_i}[z_{ij}^{\mathrm{T}}z_{ij} + \mathrm{tr}(\widetilde{W}_{ij}^{\mathrm{T}}\Gamma_{ij}\widetilde{W}_{ij})] + \frac{1}{2}\sum_{j=2}^{q_i}\tilde{\eta}_{ij}^{\mathrm{T}}\tilde{\eta}_{ij}$. Thus, by calculating the time derivative of V_i, we have

$$\begin{aligned} \dot{V}_i = &\sum_{j=1}^{q_i} z_{ij}^{\mathrm{T}}[-\kappa_{ij}z_{ij} + \widetilde{W}_{ij}^{\mathrm{T}}S_{ij}(\zeta_{ij}) + \varepsilon_{ij}(\zeta_{ij})] \\ &+ \sum_{j=2}^{q_i} z_{ij-1}^{\mathrm{T}}\tilde{\eta}_{ij} + \sum_{j=1}^{q_i}\mathrm{tr}\left(\widetilde{W}_{ij}^{\mathrm{T}}\Gamma_{ij}\dot{\widetilde{W}}_{ij}\right) \\ &+ \sum_{j=2}^{q_i}\tilde{\eta}_{ij}^{\mathrm{T}}\dot{\tilde{\eta}}_{ij} \end{aligned} \quad (19)$$

The dynamic of the filter error $\tilde{\eta}_{ij}$ can be written as

$$\dot{\tilde{\eta}}_{ij} = -\frac{\tilde{\eta}_{ij}}{\tau_{ij}} + \pi_{ij}\left(\bar{z}_{ij}, \bar{\tilde{\eta}}_{ij}, \bar{\widetilde{W}}_{ij-1}, Y_i\right) \quad (20)$$

Where we denote $\bar{z}_{ij} = \left[z_{i1}^{\mathrm{T}}, \cdots, z_{ij}^{\mathrm{T}}\right]^{\mathrm{T}}$, $\bar{\tilde{\eta}}_{ij} = \left[\tilde{\eta}_{i2}^{\mathrm{T}}, \cdots, \tilde{\eta}_{ij}^{\mathrm{T}}\right]^{\mathrm{T}}$, $\bar{\widetilde{W}}_{ij} = \left[\widetilde{W}_{i1}^{\mathrm{T}}, \cdots, \widetilde{W}_{ij}^{\mathrm{T}}\right]^{\mathrm{T}}$ and $Y_i = \left[\hat{y}_i^{\mathrm{T}}, \dot{\hat{y}}_i^{\mathrm{T}}, \ddot{\hat{y}}_i^{\mathrm{T}}\right]^{\mathrm{T}}$.

Therefore, substituting the adaptive law (13) and (20) into (19), we obtain

$$\dot{V}_i = -\sum_{j=1}^{q_i} \kappa_{ij} z_{ij}^{\mathrm{T}} z_{ij} - \sum_{j=1}^{q_i} z_{ij}^{\mathrm{T}} \varepsilon_{ij}(\zeta_{ij}) + \sum_{j=1}^{q_i} \sigma_{ij} \mathrm{tr}(\widetilde{W}_{ij}^{\mathrm{T}} \hat{W}_{ij}) - \sum_{j=2}^{q_i} \frac{\tilde{\eta}_{ij}^{\mathrm{T}} \tilde{\eta}_{ij}}{\tau_{ij}} + \sum_{j=2}^{q_i} z_{ij-1}^{\mathrm{T}} \tilde{\eta}_{ij} + \sum_{j=2}^{q_i} \tilde{\eta}_{ij}^{\mathrm{T}} \pi_{ij} \tag{21}$$

According to Assumption 2 and 3, we know that the desired rotating radius, angular velocity and angular acceleration are bounded. Then $\hat{y}_i$, $\dot{\hat{y}}_i$ and $\ddot{\hat{y}}_i$ are bounded, i.e., there exists a positive constant Y_i^* such that $\Xi_i = \{Y_i | \|\hat{y}_i\| + \|\dot{\hat{y}}_i\| + \|\ddot{\hat{y}}_i\| \leq Y_i^*\}$. In addition, we denote that $\Pi_i = \{(\bar{z}_{iq_i}, \bar{\tilde{\eta}}_{iq_i}, \overline{\widetilde{W}}_{iq_i}) | V_i \leq 2\mu_i\}$, where μ_i is a positive constant. Then, it is not hard to see that Ξ_i and Π_i are compact sets. Thus, there exists a positive constant π_{ij}^* satisfying $\|\pi_{ij}(\cdot)\| \leq \pi_{ij}^*$.

Moreover, with Young's inequality[8], we have

$$\begin{cases} z_{ij}^{\mathrm{T}} \varepsilon_{ij}(\zeta_{ij}) \leq ({\varepsilon_{ij}^*}^2 / 2\varrho_{ij}) \|z_{ij}\|^2 + \varrho_{ij}/2 \\ \mathrm{tr}(\widetilde{W}_{ij}^{\mathrm{T}} \hat{W}_{ij}) \leq -(1/2)\|\widetilde{W}_{ij}\|_F^2 + (1/2)\|W_{ij}\|_F^2 \\ z_{ij-1}^{\mathrm{T}} \tilde{\eta}_{ij} \leq (1/2)\|z_{ij-1}\|^2 + (1/2)\|\tilde{\eta}_{ij}\|^2 \\ \tilde{\eta}_{ij}^{\mathrm{T}} \pi_{ij} \leq ({\pi_{ij}^*}^2 / 2\Delta_{ij}) \|\tilde{\eta}_{ij}\|^2 + \Delta_{ij}/2 \end{cases} \tag{22}$$

Make the following names.

$$\begin{cases} \epsilon_{ij}^1 = \kappa_{ij} - \frac{1}{2} - \frac{{\varepsilon_{ij}^*}^2}{2\varrho_{ij}}, \epsilon_{iq_i}^1 = \kappa_{iq_i} - \frac{{\varepsilon_{iq_i}^*}^2}{2\varrho_{iq_i}} \\ \epsilon_{ij}^2 = \frac{\sigma_{ij}}{2}, \epsilon_{ij}^3 = \frac{1}{\tau_{ij}} - \frac{1}{2} - \frac{{\pi_{ij}^*}^2}{2\Delta_{ij}} \end{cases} \tag{23}$$

Then, $\dot{V}_i$ can be rewritten as

$$\dot{V}_i \leq -c_{i0} V_i + c_{i1} \tag{24}$$

Where

$$\begin{cases} c_{i0} = \min\{2\epsilon_{ij}^1, 2\epsilon_{ij}^2, 2\epsilon_{ij}^3\} \\ c_{i1} = \sum_{j=1}^{q_i} \frac{\varrho_{ij}}{2} + \sum_{j=1}^{q_i} \frac{\sigma_{ij}\|W_{ij}\|_F^2}{2} + \sum_{j=2}^{q_i} \frac{\Delta_{ij}}{2} \end{cases} \tag{25}$$

Theorem 1. *Consider the multi-agent system (1) in the strict-feedback form with stationary multi-targets. Suppose that Assumptions 1-3 hold. For any bounded initial condition* $V_i(0) \leq \mu_i$, *if we choose design parameters satisfy* $c_{i0} > 0$, *then all agents will achieve the multi-target rotating encirclement formation with the proposed control scheme in Control Design.*

Proof. By integrating both ends, it is obvious that the solution of (24) satisfies the following inequality.

$$V_i \leq \left[V_i(0) - \frac{c_{i1}}{c_{i0}}\right] e^{-c_{i0}t} + \frac{c_{i1}}{c_{i0}} \tag{26}$$

Then, with Lemma 1 in Ref. 8, we know that the tracking error z_{i1} is bounded, and the upper bound is associated with c_{i1}/c_{i0}. By reasonably selecting design parameters, z_{i1} can be sufficiently reduced to 0.

Combining with Lemma 1,2,3 and 4, we can easily conclude that our proposed control scheme will drive all agents achieve the multi-target rotating encirclement formation. □

5. Conclusion

The collective multi-target rotating encirclement formation problem of strict-feedback multi-agent systems is investigated by dividing into three subproblems. Our proposed control scheme can solve this problem well.

Acknowledgements

This work was supported by the NSFC (61327807, 61521091, 61520106010, 61134005) and the National Basic Research Program of China (973 Program: 2012CB821200, 2012CB821201).

References

1. F. Chen, W. Ren and Y. Cao, Surrounding control in cooperative agent networks, *Syst. Control Lett.*, **59**(11), 2010, pp. 704-712.
2. P. Lin and Y. Jia, Distributed rotating formation control of multi-agent systems, *Syst. Control Lett.*, **59**(10), 2010, pp. 587-595.
3. M. Deghat, I. Shames, B.D.O. Anderson and C. Yu, Localization and circumnavigation of a slowly moving target using bearing measurements, *IEEE T. Automat. Contr.*, **59**(8), 2014, pp. 2182-2188.
4. L. Mo, X. Yuan and Q. Li, Finite-time rotating target-encirclement motion of multi-agent systems with a leader, *Chinese J. Phys.*, **56**(5), 2018, pp. 2265-2274.
5. T. Zhang, J. Ling and L. Mo, Distributed Finite-Time Rotating Encirclement Control of Multiagent Systems With Nonconvex Input Constraints, *IEEE Access*, **7**, 2019, pp. 102477-102486.
6. R. Li, Y. Shi and Y. Song, Localization and circumnavigation of multiple agents along an unknown target based on bearing-only measurement: A three dimensional solution, *Automatica*, **94**, 2018, pp. 18-25.
7. J. Shao and Y. Tian, Multi-target localisation and circumnavigation by a multi-agent system with bearing measurements in 2D space, *Int. J. Syst. Sci.*, **49**(1), 2018, pp. 15-26
8. T. Zhang and Y. Jia, Adaptive neural network control of uncertain strict-feedback systems with full-state constraints by integral-barrier Lyapunov functions, *Proc. 37th Chin. Control Conf.*, Wuhan, China, 2018, pp. 846-851.
9. S.J. Yoo, Distributed consensus tracking of a class of asynchronously switched nonlinear multi-agent systems, *Automatica*, **87**, 2018, pp. 421-427.
10. T. Zhang and Y. Jia, Adaptive neural network control of uncertain strict-feedback systems with full-state constrains by integral-barrier Lyapunov functions, *Proc. 2018 Chin. Intel. Syst. Conf.*, Wenzhou, China, 2018, pp. 71-84.
11. W. He, S. Zhang and S.S. Ge, Adaptive control of a flexible crane system with the boundary output constraint, *IEEE T. Ind. Electron.*, **61**(8), 2014, pp. 4126-4133.

Effects of Variable Arm Length on UAV Control Systems

M. Rizon, CK. Ang, MI. Solihin
Faculty of Engineering, Technology and Built Environment, UCSI University, 56000 Cheras, Kuala Lumpur, Malaysia

Zuradzman M. R., H. Desa, Shahriman A. B., Wan Khairunizam
School of Mechatronic, University Malaysia Perlis (UniMAP), 02600 Arau, Perlis, Malaysia

I. Zunaidi
Faculty of Technology, University of Sunderland, St Peter's Campus, Sunderland, SR6 0DD, United Kingdom

E-mail: MohdRizon@ucsiuniversity.edu.my, angck@ucsiuniversity.edu.my, mahmudis@ucsiuniversity.edu.my, zuradzman@unimap.edu.my, hazry@unimap.edu.my, shahriman@unimap.edu.my, khairunizam@unimap.edu.my, zunaidi.ibrahim@sunderland.ac.uk

Abstract

Quadrotor is a type of unmanned aerial vehicle that has been widely used in many applications, such as, policing, surveillance, aerial photography and agriculture. Conventionally, the control of quadrotor flight direction is accomplished by varying speeds of motors or manipulating torques. In this paper, a novel mechanism is proposed. The mechanism uses stepper motors to control the arm length for changing flight directions, while maintaining motors' speed at constant. A mathematical model has been created. The analysis results have shown that varying arm length can effectively control the moment of bending of quadrotors. Increasing the length of arms can result in the increase of the moment of bending without changing speed of motors, thus saving energies. Experimental results have shown that the new mechanism is able to carry more payloads which the motor speed can be utilized fully at 100% while the flight direction is been controlled by changing of the arm length compared to conventional flight control mechanisms.

Keywords: UAV, Quadrotor, Moment of Bending, Arm length

1. Introduction

UAV is a short name for “Unmanned Aerial Vehicle” defined as aircrafts without the onboard presence of pilots [1]. UAV is commonly used in military and police forces in situations where the risk in sending a human piloted air craft is unacceptable. In many developed countries, UAV is also used to perform tasks, for example intelligence, surveillance, and reconnaissance missions. Small UAVs can also be used for entertainment industry, such as, aerial filming, aerial photography and others. Most of the UAVs are made in quadrotor due to its easy to design, small rotors, and excellent manoeuvrability [2]. The study requires analyzing the relationship between UAV’s motor support bar length and the UAV control systems. Quadrotor is a type of UAV that is lifted and propel by four rotors. The propellers are connected with two pairs of support bars. The flight direction of quadrotor can be controlled by increasing or decreasing the speed of motors [3].

Larger motor required quite high current, which are difficult to control using current or voltage regulating circuitry. Also, power usage of the UAV systems is a major problem to withstand the desired endurance of flight time. A study to manage the power usage is important and a study from fluid mechanic’s point of view is much needed. Besides that, flight movement of quadrotor limits the power of motor. Quadrotor motors cannot operate in 100% power because of another 30% of power is reserved to control for flight direction. Furthermore it is difficult to control the stability and precision of flight movement of quadrotor that powered by fuel or petrol [4][5].

Therefore, the theory is implied to control the quadrotor’s direction of flight by increasing or decreasing the length of the support bar. The length of support bar helps in balancing and stabilizing the UAV. The study focus on relation between support bar length and moment of bending of various position of motor. Thus, this study shall be the basis to run the dynamic analysis at the UAV’s motor support bar length control systems and also to enhance the UAV’s mathematical modelling by using SOLIDWORKS® software, CAD and CAE systems [6].

2. Quadrotor Dynamic Model

Quadrotor is an UAV that consist of four motors located at the end of the cross configuration. Each of the motor consists of propeller that generate thrust force for lifting the quadrotor [5]. The front and rear motor rotate counter-clockwise, while the left and the right motor rotate clockwise [7]. The rotations allow to nearly canceling the gyroscopic effects and aerodynamic torques in trimmed flight [8]. The quadrotor is able to move in yaw, roll, pitch and hover direction with the change of thrust outputs of each propeller. In Figure 1(a), hover movement is obtained when the net thrust of all the rotors is equal to zero. In order to achieve zero net thrust for all rotors, the direction of rotation of rotor 1 and rotor 3 must be in the opposite direction of rotor 2 and rotor 4. Besides that, all the rotors are operated in equal speed. When the speed of motor (1, 2, 3, 4) are increased equally as shown in Figure 1, the quadrotor flies in vertical motion along z-axis according to the thrust output produce due to the propeller rotation. In Figure 1(b), pitch movement is obtained when the torque of motor 3 is increase/decrease and the torque of motor 1 is decrease/increase and keeping constant torque of motor 2 and 4. The quadrotor will pitch forward or backward. In Figure 1(c) roll movement is obtained when the torque of motor 4 is increase/decrease and the torque of motor 2 is decrease/increase and keeping constant torque of motor 1 and 3. The quadrotor will roll right or roll left. In Figure 1(d), yaw movement is obtained when the torque of motor 2 and 4 is increase while decrease the torque of the motor 1 and 3 to move left and vice versa for the right direction. The quadrotor will rotate clockwise or counter-clockwise.

Quadrotor is non-linear system with 6 degree of freedom and contains only 4 motors input [8]. Figure 2 represents the free body diagram and axes of the quadrotor. In Figure 2, l is the distance of each motor towards the center of the pivot. ϕ, θ and ψ are representing the Euler angles about the body axes x, y, z. $F1$, $F2$, $F3$ and $F4$ are the thrust force produce by the propeller.

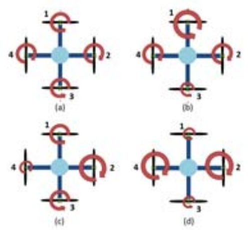


Figure 1. Movement of the quadrotor due to propeller rotation

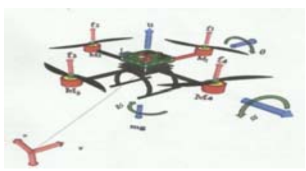


Figure 2. Free body diagram of the quadrotor [4]

The position and velocity vectors in Earth frame stated as

$$x = [x\ y\ z]^T \tag{1}$$

$$\dot{x} = [\dot{x}\ \dot{y}\ \dot{z}]^T \tag{2}$$

The Pitch, Roll and Yaw angle in body frame stated as

$$\theta = [\theta\ \phi\ \psi]^T \tag{3}$$

$$\dot{\theta} = [\dot{\theta}\ \dot{\phi}\ \dot{\psi}]^T \tag{4}$$

Then, R indicates the rotational matrix from body to earth frame as

$$R = \begin{pmatrix} c\psi c\theta & -s\psi c\phi + c\psi s\theta s\phi & c\phi c\psi s\theta + s\theta s\psi \\ s\psi c\phi & c\psi c\theta + s\theta s\phi s\psi & c\phi s\theta s\psi - c\psi s\phi \\ -s\theta & c\theta s\phi & c\theta c\phi \end{pmatrix} \tag{5}$$

In equation (5), c indicates cos( ) and s indicates sin( ).

The thrust force and control torque act on the body and produced by the propeller rotation. The vector of thrust moves from Body frame to Earth frame. Applying the Newton-Euler method for rigid body as:

$$m\ddot{x} = FR + [0\ 0 - mg]^T - F_a \tag{6}$$

$$r\dot{w} = w \times rw + \tau \tag{7}$$

$\ddot{x}$ represents the linear acceleration vector, F is the thrust that produced by the motors, F_a indicates the frictional force, τ is the control torque produced by motors, m is the mass of the body. Thus, from the equation (6) and (7), quadrotor dynamic equation is derived as:

$$\begin{gathered} \boldsymbol{m\ddot{x} = u1(cos\phi sin\theta cos\psi + sin\phi sin\psi) - k1\dot{x}} \\ \boldsymbol{m\ddot{y} = u1(cos\phi sin\theta cos\psi - sin\phi sin\psi) - k2\dot{y}} \\ \boldsymbol{m\ddot{z} = u1(cos\theta cos\phi) - mg - k3\dot{z}} \\ I_y\ddot{\phi} = (I_z - I_x)\dot{\phi}\dot{\psi} + u3 - k5\dot{\phi} \\ I_z\ddot{\psi} = (I_x - I_y)\dot{\theta}\dot{\phi} + u4 - k6\dot{\psi} \\ I_x\ddot{\theta} = (I_x - I_z)\dot{\theta}\dot{\psi} + u2 - k4\dot{\theta} \end{gathered} \tag{8}$$

Where I is the moment of inertia. The thrust induced variation the speed of the motors is the input vector and stated as:

$$U = [u1\ u2\ u3\ u4]^T \tag{9}$$

Thus, the input vectors are defined as:

$$\begin{gathered} u1 = b(w1^2 + w2^2 + w3^2 + w4^2) \\ u2 = bl(w3^2 - w1^2) \\ u3 = bl(w4^2 - w2^2) \\ u4 = d(w2^2 + w4^2 - w1^2 - w3^2) \end{gathered} \tag{10}$$

Where, w is the motor speed, b and d are thrust coefficient and drag coefficient. The lift force to hoover the quadrotor is defined as $u1$. $u2$, $u3$, and $u4$ are the input torques that locate the quadrotor to the pitch, roll and yaw attitude. In equation (10), the increasing and decreasing of motor speed produce the input torque that oriented toward pitch, roll or yaw attitude.

3. Varying Arm Length

Figure 3(a) indicates the pitch movement that is produced by the decreasing the arm length l1 and increasing the arm length l3 and keeping both arm length l2 and l4 length unchanged. Figure 3(b) shows the pitch movement produced when the arm length l1 is increased and the arm length l3 is decreased and kept constant length in both arm length l2 and l4. Figure 3(c) illustrates the roll movement that is obtained when the arm length l2 is decreased and the arm length l4 is increased while keeping the length of both l1 and l3 fixed. Figure 3(d) shows the roll movement that is obtained when the arm length l2 is increased and the arm length 14 is increased while keeping the length of both l1 and l3 fixed. In Figure 3(e), the yaw movement is accomplished by increasing the arm length l2, l4 and decreasing the arm length l1, l3. Thus, the quadrotor tilted in counter-clockwise direction. Figure 3(f) indicates the yaw movement in clockwise direction produced by decreasing the arm length l1, l3 and increasing the arm length l2, l4.

Equation (10) that stated earlier explains the input torque produced by the increasing and decreasing the motor speed. According to torque law:

$$\tau = f_i l \tag{11}$$
$$f_i = b w_i^2 \tag{12}$$

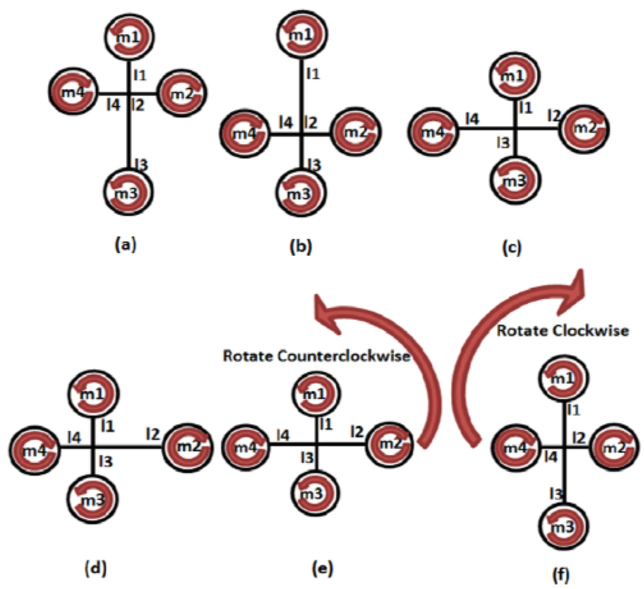


Figure 3. Pitch, roll, yaw moment respect to arm length

The input torque is derived as:

$$\begin{aligned} u1 &= b(w1^2 + w2^2 + w3^2 + w4^2) \\ u2 &= b w_i^2 (l3 - l1) \\ u3 &= b w_i^2 (l4 - l2) \\ u4 &= d w_i^2 (l2 + l4 - l3 - l1) \end{aligned} \tag{13}$$

Therefore, the pitch, yaw, roll movement of the quadrotor can be controlled by increasing or decreasing arm length to generate the certain input torque. The motor speed is kept at constant where $w_i = w1 = w2 = w3 = w4$.

4. Mathematical Analysis of Quadrotor Arm Length

The analysis is conducted by referring KDE data performance. KDEdirect.com provides data performance of specific motor version for its suitable propeller size and voltage. Motor version that suited well with 18.5” diameter and 6.3 mm pitch propeller is KDE5215XF-435(435KV). The data consists of amperage, power input, thrust output, rpm and efficiency for different throttle range between 25% to 100%. The data is obtained from the experiments conducted by the KDE Company.

Table 1. KDE data performance

MOTOR VERSION	VOLTAGE [V]	PROPELLER SIZE	THROTTLE RANGE	AMPERAGE [A] (LOWER IS BETTER)	POWER INPUT [W] (HIGHER IS BETTER)	POWER INPUT [hp]	THRUST OUTPUT [g] (HIGHER IS BETTER)	THRUST OUTPUT [N]	THRUST OUTPUT [lb]	RPM [rev/min] (HIGHER IS BETTER)	EFFICIENCY [g/W] (HIGHER IS BETTER)	EFFICIENCY [lb/HP]
KDE5215XF-435(435Kv) KDEXF-UAS75HVC S.R.ENABLED	14.8V(4S) 16.8V MAX	18.5"×6.3 KDE-CF185-DP DUAL BLADE	25.00%	2.1	31	0.04	430	4.22	0.95	1980	13.87	22.8
			37.50%	4.6	68	0.09	860	8.43	1.9	2820	12.65	20.79
			50.00%	8.3	122	0.16	1390	13.63	3.06	3600	11.39	18.73
			62.50%	13.6	201	0.27	2010	19.71	4.43	4320	10	16.44
			75.00%	20.6	304	0.41	2710	26.58	5.97	4980	8.91	14.66
			87.50%	29.7	439	0.59	3510	34.42	7.74	5640	8	13.14
			100.00%	39.3	581	0.78	4420	43.35	9.74	6240	7.61	12.51

5. Determining Moment of Bending for Every Motor Speed (RPM)

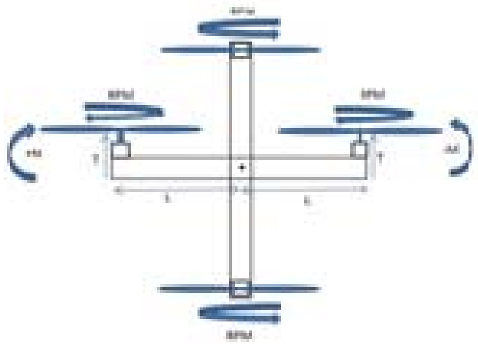


Figure 4. Free body diagram respect to moment thrust and length

Figure 4 shows the free body diagram of the quadrotor. In order to produce hover movement, the equivalent net torque of the quadrotor is equal to zero and all motor spinning in equal speeds. Moment of bending acting on quadrotor arm can be calculated using the formula as shown in equation (14).

(14)

$$M_i = T_i \times L$$

M is the moment of force, T is the force or thrust output produce by the propeller rotation and L is the quadrotor arm length, and i=1, 2, 3, 4, 5, 6, 7. In the first step of the analysis, there are 2 variables, a fixed variable and a manipulated variable. The fixed variable is the length of the quadrotor arm, L. The manipulated variable is the thrust output produced from variation motor speeds. The length of the quadrotor arm is 400 mm suitable for 18.5" propeller. Quadrotor hover in vertical direction, when the all four-motor spinning in equal speed, produces the same amount of thrust. Higher speed produces more thrust so that the quadrotor flies higher and vice versa for lower speed. Therefore, the moment of bending can be calculated for every speed of motor provided from KDE data performance.

6. Determining The Change in Quadrotor Arm Length

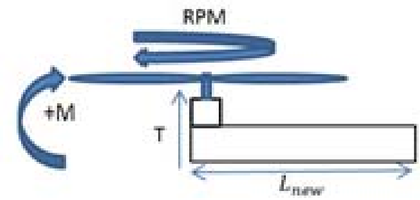


Figure 5: Free body diagram for 1 arm

Finding arm length needed to produce a required moment from previous data using thrust output produce from 3 motor speeds of 3600 RPM, 4320 RPM and 4980 RPM. The constant variable now is change to thrust output produced from a fixed motor speed and the manipulated variable is the moment of bending, M_i. Their relationship is shown in the equation (15).

$$L_{inew} = M_i / T \tag{15}$$

Where i= 1,2,3,4,5,6,7 and M is the moment of force, T is the thrust output.

7. Assembly Design of The Quadrotor Arm

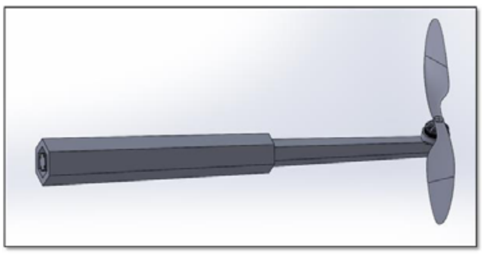


Figure 6. Assembly design of the quadrotor arm

The design requires a stepper motor, a fixed arm, a moving arm, 4 x 18.5" propellers and a brushless DC motor.

8. Results

Figure 7 shows the relation of moment of bending against the thrust output. In the diagram, we can see that the moment of bending is increased directly that is proportional to the thrust output. Thrust output produced by the rotation of the propeller is also related to the speed of motor. The increased in motor speed results in higher thrust generated. The minimum value of moment of bending is 1.69 N.m acting on 0.4 m quadrotor arm length, with the thrust output of 4.22 N. The moment of bending increases until it reaches to the maximum value of 17.34 N.m with the thrust output of 43.35 N.

Table 2. Data moment of bending due to various motor speed

MOTOR VERSION	PROPELLER SIZE	THROTTLE RANGE	RPM [rev/min] (HIGHER IS BETTER)	THRUST OUTPUT (HIGHER IS BETTER) [g]	[N]	[lb]	ARM LENGTH (m)	MOMENT OF BENDING (N.m)
KDE5215XF-435(435Kv) KDEXF-UAS75HVC S.R.ENABLED	18.5"× 6.3 KDE-CF185-DP DUAL-BLADE	25.00%	1980	430	4.22	0.95	0.4	1.69
		37.50%	2820	860	8.43	1.9	0.4	3.37
		50.00%	3600	1390	13.63	3.06	0.4	5.54
		62.50%	4320	2010	19.71	4.43	0.4	7.88
		75.00%	4980	2710	26.58	5.97	0.4	10.63
		87.50%	5640	3510	34.42	7.74	0.4	13.77
		100.00%	6240	4420	43.35	9.74	0.4	17.34

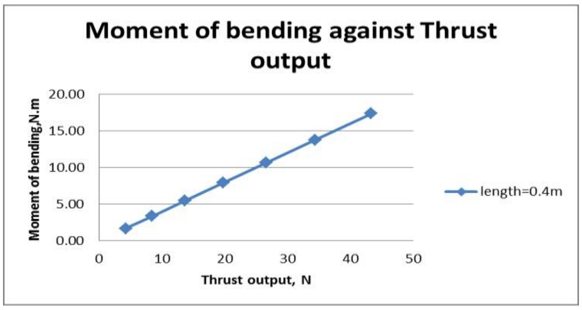


Figure 7. Graph moment of bending against thrust output

9. Result of Arm Length for Required Moment of Bending

Figure 8 shows the arm length against moment of bending for 3 different motor speeds 3600 RPM, 4320 RPM and 4980 RPM. In the graph, we can see that the arm length is proportional to the moment of bending. Additionally, the slope showing in the graphs differs from each other. The slope of the motor speed of 3600 RPM is highest compared to 2 other motor speeds, 4320 RPM and 4980 RPM. Motor speeds of 3600 RPM, 4320 RPM and 4980 RPM generate thrust output of 13.63 N, 19.71 N and 26.58 N respectively.

Table 3: Data of arm length for required moment of bending (3600 RPM)

MOTOR VERSION	VOLTAGE [V]	PROPELLER SIZE	THRUST[N]	RPM [rev/min]	MOMENT [N.m]	Lnew [m]
			13.63	3600	1.69	0.12
			13.63	3600	3.37	0.25
KDE5215XF-435(435Kv)		18.5"× 6.3	13.63	3600	5.45	0.40
KDEXF-UAS75HVC	14.8V(4S)	KDE-CF185-DP	13.63	3600	7.88	0.58
S.R.ENABLE	16.8V(MAX)	DUAL-BLADE	13.63	3600	10.63	0.78
			13.63	3600	13.77	1.01
			13.63	3600	17.34	1.27

Table 4: Data of arm length for required moment of bending (4320 RPM)

MOTOR VERSION	VOLTAGE [V]	PROPELLER SIZE	THRUST[N]	RPM [rev/min]	MOMENT [N.m]	Lnew [m]
			19.71	4320	1.69	0.09
			19.71	4320	3.37	0.17
KDE5215XF-435(435Kv)		18.5"× 6.3	19.71	4320	5.45	0.28
KDEXF-UAS75HVC	14.8V(4S)	KDE-CF185-DP	19.71	4320	7.88	0.40
S.R.ENABLE	16.8V(MAX)	DUAL-BLADE	19.71	4320	10.63	0.54
			19.71	4320	13.77	0.70
			19.71	4320	17.34	0.88

Table 5: Data of arm length for required moment of bending (4980 RPM)

MOTOR VERSION	VOLTAGE [V]	PROPELLER SIZE	THRUST[N]	RPM [rev/min]	MOMENT [N.m]	Lnew [m]
			26.58	4980	1.69	0.06
			26.58	4980	3.37	0.13
KDE5215XF-435(435Kv)		18.5"× 6.3	26.58	4980	5.45	0.21
KDEXF-UAS75HVC	14.8V(4S)	KDE-CF185-DP	26.58	4980	7.88	0.30
S.R.ENABLE	16.8V(MAX)	DUAL-BLADE	26.58	4980	10.63	0.40
			26.58	4980	13.77	0.52
			26.58	4980	17.34	0.65

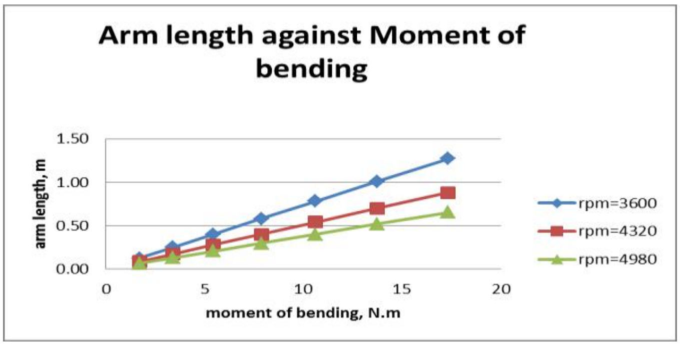


Figure 8. Graph arm length against moment of bending

In order to generate moment of bending of 1.69 N.m to 17.34 N.m with motor speed of 3600 RPM, we need to increase the arm length from 0.12 m to 1.7 m. For 4320 RPM motor speed, the arm length needs to be increased from 0.09 m to 0.88 m. For motor speed of 4980 RPM, the arm lengths needs to be increased from 0.06 m to 0.65 m in order to generate moment of bending of 1.69 N.m to 17.34 N.m. It shows that lower thrust output or lower motor speed requires longer arms to produce sufficient moment.

10. Result of Solidworks Motion Simulation

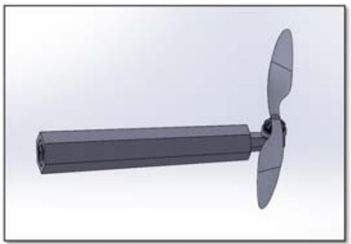


Figure 9: Arm moves inside the fixed arm

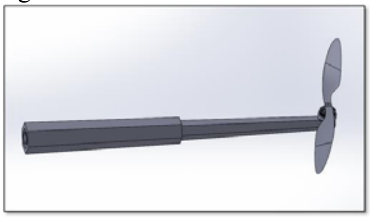


Figure 10: Arm move outside the fixed arm

Figure 9 shows the moving arm moving inside the fixed arm. This occurs when the stepper motor rotates counter-clockwise thus the moving arm moves in negative x-direction. Figure 10 shows that the moving arm is move outside the fixed arm. This motion occurred when the stepper motor rotates clockwise and achieved positive linear motion (forward motion).

11. Conclusions

Based on the mathematical analysis conducted, the varying in arm lengths affects the moment of bending of the quadrotor. Increasing the arm length changes the moment of bending of the quadrotor to increase or decrease the arm length will decrease the moment of bending. The speed of motor is kept at constant so that the thrust generated by the propeller rotation is at constant. The analysis also proves that the change in mechanism of the existing UAV's quadrotor that is the quadrotor flight direction can be obtain by increasing or decreasing motors speed. The existing methods have been extended to varying the speed of the motor, the quadrotor flight direction can be controlled by varying the length of the quadrotor's arms.

References

[1]S. Norouzi Ghazbi, Y. Aghli, M. Alimohammadi, and A. A. Akbari, "Quadrotors unmanned aerial vehicles: A review," *Int. J. Smart Sens. Intell. Syst.*, 2016.

[2] S. A. Raza, "Design and control of a quadrotor unmanned aerial vehicle," *ProQuest Diss. Theses*, 2010.

[3] P. Castillo, R. Lozano, and A. Dzul, “Stabilization of a mini rotorcraft with four rotors: Experimental implementation of linear and nonlinear control laws,” *IEEE Control Syst. Mag.*, 2005.
[4] N. Yasameen Kamil, D. Hazry, K. Wan, and Z. M. Razlan, “Trajectory tracking based on arm’s length variation,” *J. Theor. Appl. Inf. Technol.*, 2015.
[5] N. Yasameen Kamil, D. Hazry, K. Wan, and Z. M. Razlan, “A novel VAL: Quadrotor control technique for trajectory tracking based on varying the Arm’s Length,” *ARPN J. Eng. Appl. Sci.*, 2016.
[6] L. M. Gómez-López, V. Miguel, A. Martínez, J. Coello, and A. Calatayud, “Simulation and modeling of single point incremental forming processes within a solidworks environment,” in *Procedia Engineering*, 2013.
[7] T. Krokowicz, M. Gasca, H. Voos, and D. Ucinski, “Indoor navigation for quadrotor UAVs using schematic environment maps,” in *19th International Workshop on Robotics in Alpe-Adria-Danube Region, RAAD 2010 - Proceedings*, 2010.
[8] A. A. Mian and D. Wang, “Modeling and backstepping-based nonlinear control strategy for a 6 DOF quadrotor helicopter,” *Chinese J. Aeronaut.*, 2008.
[9] CH Shen, FYC Albert, CK Ang, DJ Teck, KP Chan, "Theoretical development and study of takeoff constraint thrust equation for a drone," *2017 IEEE 15th Student Conference on Research and Development (SCOReD)*, 18-22, 2017.
[10]Y Yie, MI Solihin, AC Kit, "Development of swarm robots for disaster mitigation using robotic simulator software," *9th International Conference on Robotic, Vision, Signal Processing and PowerApplications*, 2017, 377-383.
[11] SH Tang, CK Ang, MKABM Ariffin, SB Mashohor, "Predicting the motion of a robot manipulator with unknown trajectories based on an artificial neural network," *International Journal of Advanced Robotic Systems,* 11 (10), 2014, 176.
[12] TS Hong, AC Kit, DN Nia, MKAM Ariffin, W Khaksar, "Planning for Redundant Manipulator Based on Back-Propagation Neural Network," Advanced Science Letters 19 (11), 2013, 3307-3310.
[13] CK Ang, SH Tang, S Mashohor, MKAM Arrifin, "Solving continuous trajectory and forward kinematics simultaneously based on ANN," International Journal of Computers Communications &Control 9(3),2014,253-260.

EEG based drowsiness detection using relative band power and short time fourier transform

Pranesh Krishnan, Sazali Yaacob, Annapoorni Pranesh Krishnan
*Intelligent Automotive Systems Research Cluster, Electrical Electronic and Automation Section, Universiti Kuala Lumpur Malaysian Spanish Institute, Kulim, Kedah, 09000, Malaysia**

Mohamed Rizon
Faculty of Engineering, Technology and Build Environment, UCSI University, No 1, Jalan Menara Gading, Taman Connaught, 56000 Kuala Lumpur, Malaysia
E-mail: sazali.yaacob@unikl.edu.my, pranesh@unikl.edu.my, annapoorni.mani@s.unikl.edu.my, MohdRizon@ucsiuniversity.edu.my

Abstract

Sleeping on the wheels due to drowsiness is one of the major causes of death tolls all over the world. The objective of this research article is to classify drowsiness with alertness based on the EEG signals using spectral and band power features. A publicly available ULg DROZY database used in this research. The five EEG channels from the raw multimodal signal are extracted and by using a higher order Butterworth low pass filter the high frequency components above 50 Hz are removed. Another bandpass filter bank is designed to separate the raw signals into eight sub bands namely delta, theta, low alpha, high alpha, low beta, mid beta, high beta and gamma. The preprocessed signals are segmented into equal number of frames with a frame duration of 2 seconds using a rectangular time windowing approach with an overlap of 50%. The relative band power based on the short time fourier transform (STFT) was computed for each frame and the features are extracted. The extracted feature sets are further normalized and labelled as drowsy and alert and then combined to form the final dataset. K-fold cross validation method is used. The dataset is trained using KNN and SVM classifiers and the results are compared. The KNN classifier produces 96.1% (dataset1) and 95.5%(dataset2) classification accuracy.

Keywords: drowsiness, polysomnography, band power, short time fourier transform.

1. Introduction

Amongst the various drowsiness detection techniques, the physiology-based methods involve electroencephalogram (EEG), electromyogram (EMG), electrocardiogram (ECG) and electrooculogram (EOG). The EEG method yields high accuracy and is referred to as gold standard. Thus, the physiology-based method produces high efficiency and reliability in detecting drowsiness at an early stage. The EEG signal, which is the most preferred physiological signal, has the closest association with drowsiness. The drawbacks of the EEG signals are that, as they have high temporal resolution and can be easily interfered by EMG, eye blink and electromagnetic noise [2].

Drowsy driving is a vital issue and hence need to be identified at the earliest. Drowsiness has subsequent effects. a) Driver's reduced attention to surroundings, b) Considerable delay in reaction time, and c) Affect the driver's capability to make decisions. The literature study attempts to review all the previous approaches in detecting drowsiness. Drivers involved in sleep-related crashes rate their quality of sleep was either bad or fair [3]. Missing an hour of sleep can lead to car crash risk. The only antidote for drowsiness is sleep [4]. The remainder of this paper is organized as follows: Section 2 presents the multimodal DROZY database. Section 3 describes the feature extraction and modelling of the drowsiness detection. The results are presented and discussed in detail in Sections 4 and 5 respectively. Section 6 presents the conclusions presented in this paper.

2. DROZY database

The ULg multimodality drowsiness (DROZY) database is considered in this research as it contains multimodal approach. The complete details of the multimodal data collection, data description and the protocol are represented in Figure 1(Massoz, Langohr, Francois, & Verly, 2016)

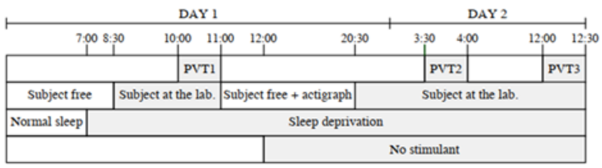


Fig. 1. Protocol design during the data collection (Massoz et al., 2016).

The database contains physiology related signals i.e., EEG, EOG, ECG, and EMG. By using Embla Titanium system, the signals were recorded from EEG channels Fz, Pz, Cz, C3, and C4 referenced at A1 in the international 10-20 system. The vertical and horizontal EOG signals, EMG and ECG signals were also recorded at a sampling rate of 512 Hz. Along with these signals video signals were also recorded. The drowsiness and alert signals were recorded from 14 male and female subjects. The test was conducted in a controlled environment in three trials. Before the first trial, the subjects were asked to have a good sleep pattern for the past week. In the first trial, the subjects were asked to perform an action watching the screen. After the first trial, the subjects were asked to stay awake for 36 to 38 hours to keep them sleep deprived. In the second and third trials, the subjects performed the same previous experiment. Post the final test, the subjects were asked to take a good sleep before they drive back home.

3. Methodology

3.1. *EEG Channels*

The experiment was conducted in three trials in two days as mentioned in the protocol. The authors conducted a widely used tool to measure the performance impairments due to drowsiness using a 10-minute psychomotor vigilance test (PVT). This test gives the reaction time to visual or auditory stimuli that occurs at random inter-stimulus interval. The database contains PSG signals from 11 electrophysiological signals (5 EEG, 2 EOG, 1 ECG and 1 EMG) are considered. The five EEG channels are recorded from C3, C4, Cz, Pz, and Fz locations present in the central lobe of the brain. The placement of the EEG channels is depicted in Figure 4. The sensors placed around C3, C4 and Cz locations deal with the sensory and motor functions. C3 and C4 gave the best seizure detection performance. This research, uses monopolar montage with C3 as reference. The vertical and horizontal EOG signals are recorded from above and at the side of the right eye to capture the eye blinks. An ECG channel is recorded from the electrode placed on the chest and an EMG signal is recorded from the electrode placed on the neck of the participant. The placement of electrodes is pictorially represented in Figure 2

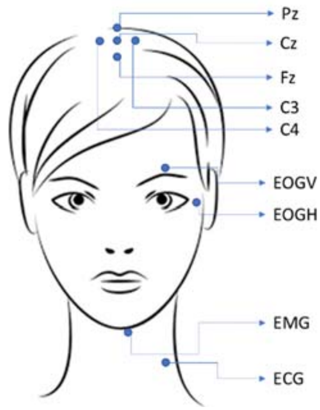


Fig. 2. Multimodal electrode placement during the trials

3.2. *EEG signal processing*

EEG signals are the non-invasive physiological means of measuring the brain activity. It has the closest relationship with drowsiness. The EEG has a neural domain signal information has high temporal resolution but can be easily interfered by EMG, eye blink and electromagnetic noise (Ha & Yoo, 2016). The EEG, along with the EMG, ECG, and EOG signals remain the European data format (EDF). A MATLAB function is used to extract only the EEG signals which is of the interest in this research. The experiment was conducted for 10 minutes and the EEG signal contains information up to 600 seconds. The signals are recorded at 512 Hz sampling frequency. The signals include both EOG and EMG which contribute towards the drowsiness detection. Hence the eye blink artefacts and the EMG artefacts which are removed in conventional bio signal processing is avoided in this research. The raw signal is processed directly without any artefact's removal. However, a Butterworth low pass filter with a cutoff frequency of 50

Hz is designed and the high frequency components are eliminated.

3.3. *EEG sub banding*

From the literature, the specific frequency ranges from the EEG signals such as alpha, beta, theta, gamma and delta are extracted using a suitable band pass filter. A Butterworth bandpass filter bank with 8 frequency bands is designed and used to extract the sub bands of the EEG signals. The details of the eight sub bands are listed in Table 1.

Table 1 EEG subband filtering.

Wave type	Freq (Hz)	Nature	Optimal
Delta	0.1 - 4	Very slow	Deep sleep
Theta	4 - 8	Slow	Depression
Low Alpha	8 - 9	Moderate	Relaxation
High Alpha	9 – 12	Moderate	Relaxation
Low beta	12 – 16	High	Conscious
Mid beta	16 – 20	High	Blinking
High beta	20 - 30	High	Agitation
Gamma	30 -50	Very high	Cognition

3.4. *Frame analysis and overlap*

The raw signal is recorded for 10 minutes. It is difficult to apply any feature extraction methods to the whole signal. Hence the signal is divided using a time window. Signal epoch: Specific time windows are extracted from the continuous EEG signal. These time windows are called epochs. On a trial and error method, the time windows are selected based on the performance metrics. In this approach a 2 second window is chosen to segment the raw EEG signal. Sliding window is applied on the signal

4. Feature extraction

Dimensionality reduction plays an important role in the classifier performance. Dimensionality reduction is achieved either by feature extraction or feature selection. For both feature extraction and feature selection approach, feature evaluation criterion, dimensionality of the feature space and optimization procedure are required. Feature extraction is the transformation of the original data to a data set with the selected number of variables which contains the most discriminatory information. Feature extraction on the other hand, considers the whole original data and maps the useful information into a lower dimensional space

4.1. *Windowing and STFT*

During each trial of the experiment is recorded for an average of 10 minutes which is approximately 600 seconds. All the signals are read and the maximum length of the signal is computed and taken as reference, for those signals where the length is short, zeros are padded at the end of the signal to make all signals even. The raw EEG signal is segmented into 2 seconds pulses. After segmenting the signals into frames of 2 seconds with the windowing technique, the original raw signal grows to 300 per each subject per each trial, thereby enhancing the data.
For each frame, the average relative power (power in the specific band over total power) in the low alpha, high alpha and delta bands are calculated from the EEG power spectrum, computed using a short time Fourier transform (STFT) using Eq. (1) [18]

$$S_x^{(w)}(t, f) = \int_{-\infty}^{\infty} \left[x(t) w^*(t - t'') \right] . e^{-j2\pi ft} dt \qquad (1)$$

Short time Fourier transform (STFT) is a frequently used feature extraction technique in which separation of stationary signals is performed into small fragments. The process is computed to extract the features from both alert and drowsy databases.

5. Data processing

Finally, the output class is mapped according to the state of alertness or drowsiness. The database contained one trial for alert and two trials for drowsy situations. The extracted feature sets were further normalized and labelled as 1 for the features from the drowsy trial, and ‘0’ for the features extracted from the alert trial. The mixture of Trial 1 (alert) and Trial 2 (drowsy1) formed the final dataset 1 while the Trial 1 (alert) and Trial 3 (drowsy2) formed the final dataset 2. The dataset is split into training and testing using the K-fold cross validation method. The final dataset was rescaled following the bipolar normalization method between 0 to 1.

6. Classification

K-Nearest neighbor algorithm (k-NN) is implemented using ‘*fitcknn*’ function in MATLAB. The K factor, training error rate and validation error rate and adjusted accordingly. The number of neighbor’s value was set to 3 while ‘*minkowski*’ method is used as a distance metric to classify the two classes ‘alert’ and ‘drowsy’.

Support Vector Machine (SVM) classifies between classes by finding the hyperplanes which is a very crucial task in SVM. The '*fitcsvm'* function in MATLAB is used to model the SVM. '*Radial basis function*' kernel is used to classify between drowsy and alert classes.

7. Results

Classifiers are evaluated based on the ease to interpret output, calculation time and prediction power. In order to develop a binary classification, the trail 1 data are labelled as alert while the trail 2 and trial 3 data are labelled as drowsy. Once the feature set is extracted, the output class is labelled accordingly and the final dataset is processed.

Table 2 Average Classification Accuracy.

	Dataset 1	Dataset 2
KNN	96.1 %	95.5%
SVM	93.7 %	94.9%

8. Discussion

The anomaly in the delta and alpha sub bands are extracted using the entropy and band power features. The hypothesis, during this study is to evaluate the alpha and delta activities. The alpha activity is inversely proportional to the delta activity during drowsiness. Microsleep is a temporary occurrence of sleep or drowsiness which may last from a fraction of a second to two minutes. Microsleeps occur while the person's eyes are open. It is a straight result of sleep deprivation. Frequent blinking, sudden body jerk, blank stare, head dropping are the results of microsleep. Also called as daytime drowsiness. During microsleep, there is an increase in theta and alpha waves for more than 3 seconds. (Poudel, Innes, Bones, Watts, & Jones, 2014).

9. Conclusion

This study presented a method for detection and classification of drowsiness based on the EEG signal using short time fourier transformation (STFT) technique. The publicly available DROZY database from ULg was considered for the reason that it contains multimodal signals. Five EEG channels (C3, C4, Cz, Pz, Fz) were extracted from the database. The high frequency components are not of the interest for this research and they are removed using a higher order low pass filter with a cut off frequency at 50 Hz. The low pass filtered signal were further divided into eight sub bands namely delta, theta, low alpha, high alpha, low beta, mid beta, high beta and gamma using a Butterworth band pass filter The signal was recorded for 600 seconds (10 minutes) which is tedious to analyze the whole signal, and hence the raw signal was divided into frames using a rectangular window with an overlap of 50%. The STFT was computed for all the sub bands on a frame by frame analysis. The database contained one trial for alert and two trials for drowsy situations. The extracted feature sets were further normalized and labelled as 1 for the features from the drowsy trial, and '0' for the features extracted from the alert trial. The mixture of Trial 1 (alert) and Trial 2 (drowsy1) formed the final dataset 1 while the Trial 1 (alert) and Trial 3 (drowsy2) formed the final dataset 2. The dataset is split into training and testing using the K-fold cross validation method. KNN and SVM models are trained using the datasets. The performance of the model to classify drowsiness is reported. KNN achieved better classification accuracy than SVM model (96.1% in dataset 1 and 95.5% in dataset2).

Acknowledgements

The authors would like to thank the Ministry of Higher Education for the financial support provided for this project (FRGS/1/2019/TK08/UNIKL/02/3). The authors would also thank the members of Intelligent Automotive Systems research cluster and the post-graduate students of University Kuala Lumpur Malaysian Spanish Institute for their constructive and critical suggestions during the preparation of this manuscript.

References

1. U. Ha and H. J. Yoo, "A multimodal drowsiness monitoring ear-module system with closed-loop real-time alarm," Proc. - 2016 IEEE Biomed. Circuits Syst. Conf. BioCAS, 2016, pp. 536–539, 2016.
2. J. W. W. Jane C. Stutts, B. V Vaughn, J. C. Stutts, J. W. Wilkins, and B. V Vaughn, "Why do people have drowsy driving crashes: input from those who just did," no. November, pp. 1–81, 1999.
3. E. Leefeldt, "Losing an hour of sleep can heighten car crash risk: AAA warns," CBC News, 2019. [Online]. Available: https://www.cbsnews.com/news/daylight-saving-time-losing-an-hour-of-sleep-can-heighten-your-car-crash-risk-aaa-says/. [Accessed: 28-Jun-2019].

4. W. Li, Q. C. He, X. M. Fan, and Z. M. Fei, “Evaluation of driver fatigue on two channels of EEG data,” Neurosci. Lett., vol. 506, no. 2, pp. 235–239, 2012.
5. H. Pranoto, A. M. Leman, D. Sukmajati, B. Hanum, and I. Baba, “Drivers Drowsiness Detection with Speed Limiter Integrated Fatigue Analyzer (SLIFA) on Fuel Tank Truck,” Int. J. Integr. Eng., vol. 10, no. 2, pp. 66–70, 2018.
6. D. Ribeiro, A. Cardoso, and C. Teixeira, “Online demonstration of an EEG-based drowsiness detector,” Proc. 2017 4th Exp. Int. Conf. Online Exp. exp.at, 2017, pp. 93–94, 2017.
7. H. Albalawi and X. Li, “Single-Channel Real-Time Drowsiness Detection Based on Electroencephalography,” in Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2018, vol. 2018-July, pp. 98–101.
8. H. Albalawi and X. Li, “Single-Channel Real-Time Drowsiness Detection Based on Electroencephalography,” Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS, vol. 2018-July, pp. 98–101, 2018.
9. D. Leger et al., “In-Flight Automatic Detection of Vigilance States Using a Single EEG Channel,” IEEE Trans. Biomed. Eng., vol. 61, no. 12, pp. 2840–2847, 2014.
10. M. Pathak and A. K. Jayanthy, “Designing of a single channel EEG acquisition system for detection of drowsiness,” in 2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), 2017, pp. 1364–1368.
11. M. Pathak and A. K. Jayanthy, “Development of a Real-Time Single Channel Brain-Computer Interface System for Detection of Drowsiness,” Biomed. Eng. Appl. Basis Commun., vol. 29, no. 03, p. 1750019, 2017.
12. T. L. T. T. Da Silveira, A. J. Kozakevicius, and C. R. Rodrigues, “Automated drowsiness detection through wavelet packet analysis of a single EEG channel,” Expert Syst. Appl., vol. 55, pp. 559–565, 2016.
13. A. Picot, S. Charbonnier, and A. Caplier, “On-line detection of drowsiness using brain and visual information,” IEEE Trans. Syst. Man, Cybern. Part ASystems Humans, vol. 42, no. 3, pp. 764–775, 2012.
14. M. Ogino and Y. Mitsukura, “Portable drowsiness detection through the use of a prefrontal single-channel electroencephalogram,” Sensors (Switzerland), vol. 18, no. 12, pp. 1–19, 2018.
15. A. Mehreen, S. M. Anwar, M. Haseeb, M. Majid, and M. O. Ullah, “A Hybrid Scheme for Drowsiness Detection using Wearable Sensors,” IEEE Sens. J., vol. 19, no. 13, pp. 1–1, 2019.
16. S. K. Alluhaibi, M. S. N. Al-Din, and A. Moyaid, “Driver Behavior Detection Techniques: A survey,” 2018.
17. Z. Mu, J. Hu, and J. Min, “Driver Fatigue Detection System Using Electroencephalography Signals Based on Combined Entropy Features,” Appl. Sci., vol. 7, no. 2, p. 150, 2017.
18. I. Belakhdar, W. Kaaniche, R. Djemal, and B. Ouni, “Single-channel-based automatic drowsiness detection architecture with a reduced number of EEG features,” Microprocess. Microsyst., vol. 58, pp. 13–23, 2018.
19. A. G. Correa and E. L. Leber, “An automatic detector of drowsiness based on spectral analysis and wavelet decomposition of EEG records,” 2010 Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBC’10, pp. 1405–1408, 2010.
20. A. Garcés Correa, L. Orosco, and E. Laciar, “Automatic detection of drowsiness in EEG records based on multimodal analysis,” Med. Eng. Phys., vol. 36, no. 2, pp. 244–249, 2014.
21. N. K. Kasabov, Time-Space, Spiking Neural Networks and Brain-Inspired Artificial Intelligence. 2019.
22. U. Ha and H. J. Yoo, “An EEG-NIRS ear-module SoC for wearable drowsiness monitoring system,” 2016 IEEE Asian Solid-State Circuits Conf. A-SSCC 2016 - Proc., pp. 193–196, 2017.
23. H. Göksu, “BCI oriented EEG analysis using log energy entropy of wavelet packets,” Biomed. Signal Process. Control, vol. 44, pp. 101–109, 2018.
24. S. AydIn, H. M. Saraoğlu, and S. Kara, “Log energy entropy-Based EEG classification with multilayer neural networks in seizure,” Ann. Biomed. Eng., vol. 37, no. 12, pp. 2626–2630, 2009.
25. R. Foong, K. K. Ang, C. Quek, C. Guan, and A. A. Phyo Wai, “An analysis on driver drowsiness based on reaction time and EEG band power,” Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS, vol. 2015-Novem, pp. 7982–7985, 2015.
26. L. A., R. J. A. Horne, “Falling asleep while driving: are drivers aware of prior sleepiness” vol. 44, pp. 120–121, 1998.
27. J. Horne and L. Reyner, “Vehicle accidents related to sleep: A review,” Occupational and Environmental Medicine. 1999.
28. S. Nordbakke and F. Sagberg, “Sleepy at the wheel: Knowledge, symptoms and behaviour among car drivers,” Transp. Res. Part F Traffic Psychol. Behav., vol. 10, no. 1, pp. 1–10, 2007.
29. A. Williamson, R. Friswell, J. Olivier, and R. Grzebieta, “Are drivers aware of sleepiness and increasing crash risk while driving?” Accid. Anal. Prev., vol. 70, pp. 225–234, 2014.
30. S. S. Hashemi Nazari, A. Moradi, and K. Rahmani, “A systematic review of the effect of various interventions on reducing fatigue and sleepiness while driving,” Chinese J. Traumatol. - English Ed., vol. 20, no. 5, pp. 249–258, 2017.
31. G. R. Poudel, C. R. H. Innes, P. J. Bones, R. Watts, and R. D. Jones, “Losing the struggle to stay awake: Divergent thalamic and cortical activity during microsleeps,” Hum. Brain Mapp., vol. 35, no. 1, pp. 257–269, 2014.

Mathematical Model Implementation of SPWM fed Three-phase Induction Motor Drive Using MATLAB Simulink

Amir Rasyadan, Sazali bin Yaacob, Pranesh Krishnan
Intelligent Automotive Systems Research Cluster, Electrical Electronic and Automation Section,
University Kuala Lumpur Malaysian Spanish Institute, Kulim, 09000, Kedah, Malaysia[*]

Mohamed Rizon, Ang Chun Kit, Mahmud Iwan Solihin
Faculty of Engineering, Technology and Build Environment, UCSI University, No 1, Jalan Menara Gading,
Taman Connaught, 56000 Kuala Lumpur, Malaysia
E-mail: sazali.yaacob@unikl.edu.my, pranesh@unikl.edu.my, amir.kelana22@s.unikl.edu.my,
MohdRizon@ucsiuniversity.edu.my, angck@ucsiuniversity.edu.my, mahmudis@ucsiuniversity.edu.my

Abstract

Three-phase induction motors are used in a vast area of applications mainly due to their simplicity, ruggedness and high reliability. With recent advancement in semiconductor technologies, the use of fixed speed induction motor drive is becoming obsolete, majority of the applications now requires inverter-based drives for variable speed operation. In the study of induction motor drive operation, mathematical models are often used to simulate the steady state and transient behavior of induction motor. However, to develop such model is not a straightforward task. Knowing only the equations by themselves are not always enough without some knowledge on solving mathematical equations with the use of computer simulation software. This work presents an approach to implement the mathematical model of a Sinusoidal Pulse Width Modulation (SPWM) fed three-phase induction motor drive in MATLAB Simulink. The sub models include induction motor DQ-model and a voltage source inverter (VSI) fed by SPWM signal generator. The presented model implementation is able to simulate the dynamic behavior of an induction motor operation, this would be useful for further studies on the development of induction motor drive system.

Keywords: Three-phase induction motor, DQ-model, Voltage source inverter, SPWM.

1. Introduction

Recent advancement in semiconductor technologies have further extended the capability of induction motors from the traditional fixed speed to the modern variable speed operation with the use of PWM fed inverter-based drives. Some of the applications include industrial pumps, crane lifters, drilling machines, traction applications such as in electric vehicles, and many others. Developing elaborate models have now become crucial in the designing process of induction motor drive system as the application complexity increases. The model should be able to predict and simulate operational characteristics and performances which would be useful to validate the design process and minimize designing error[1]. However, building such model is not a straight forward task, knowledge of the machine basic electric, magnetic, and mechanical behaviors and also some experience in using computer simulation software are required.

There are several literatures that have presented their respective implementation method of the induction motor mathematical model. Some of the works suggest the use of Simulink s-function[1,2], others use Simulink's mathematic operation blocks [3,4], and there are also works that have proposed a solution to avoid MATLAB/Simulink[5] where the motor are modeled completely using C++ programming language. However, their works are short in terms of modeling the SPWM signal and the inverter equations. Hence, an overview of the equations needed to mathematically model an induction motor drive and the approach to implement them in computer simulation software specifically using MATLAB Simulink are presented here. Three main components of the induction motor drive system are included in this work which are the induction motor DQ-model, the three-phase voltage source inverter [6], and the SPWM signal generator for VSI switching [7,8].

Using MATLAB, the model can be built entirely by writing programming code or one can use Simulink which provides a better system representation in block diagram form as an alternative. The main advantage of Simulink is that it uses functional blocks instead of programming code.

2. Model Implementation

The model describe in this work include the Induction motor Direct-Quadrature (DQ) model, the Sinusoidal Pulse Width Modulation model, and the voltage source inverter model. Each of the sub model is presented as follows.

2.1. *Induction Motor Direct-Quadrature (DQ) Model*

The first step of modeling an induction motor involves reference frame transformation of the three-phase supply (V_{as}, V_{bs}, V_{cs}) voltage into two-phase voltage [10,11]. To simplify the implementation, it is convenient to first transform the three-phase voltage using Eq. (1) into a two-phase voltage in stationary reference frame (V_{ds}, V_{qs}), then from the stationary reference frame to the arbitrary reference frame using Eq. (2).

$$\begin{bmatrix} v_{qs}^s \\ v_{ds}^s \end{bmatrix} = \frac{2}{3}\begin{bmatrix} 1 & -1/2 & -1/2 \\ 0 & -\sqrt{3}/2 & \sqrt{3}/2 \end{bmatrix}\begin{bmatrix} v_{as} \\ v_{bs} \\ v_{cs} \end{bmatrix} \tag{1}$$

$$\begin{bmatrix} v_{qs} \\ v_{ds} \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} v_{qs}^s \\ v_{ds}^s \end{bmatrix} \tag{2}$$

Once v_{ds} and v_{qs} are obtained, the DQ model can be calculated. The DQ-model implementations are divided into 4 parts as follows.

2.1.1. *Flux Linkages*

The equations for simulating the induction motor DQ-model can be established by first solving the flux linkage equations defined as in Eq. (3)-(6):

$$\frac{dF_{qs}}{dt} = \omega_b\left[\upsilon_{qs} - \frac{\omega}{\omega_b}F_{ds} + \frac{R_s}{x_{ls}}\left(\frac{x_{ml}^*}{x_{lr}}F_{qr} + \left(\frac{x_{ml}^*}{x_{ls}} - 1\right)F_{qs}\right)\right] \tag{3}$$

$$\frac{dF_{qs}}{dt} = \omega_b\left[\upsilon_{ds} - \frac{\omega}{\omega_b}F_{qs} + \frac{R_s}{x_{ls}}\left(\frac{x_{ml}^*}{x_{lr}}F_{dr} + \left(\frac{x_{ml}^*}{x_{ls}} - 1\right)F_{ds}\right)\right] \tag{4}$$

$$\frac{dF_{qr}}{dt} = \omega_b\left[0 - \frac{(\omega_e - \omega_r)}{\omega_b}F_{dr} + \frac{R_r}{x_{lr}}\left(\frac{x_{ml}^*}{x_{ls}}F_{qs} + \left(\frac{x_{ml}^*}{x_{lr}} - 1\right)F_{qr}\right)\right] \tag{5}$$

$$\frac{dF_{dr}}{dt} = \omega_b\left[0 - \frac{(\omega_e - \omega_r)}{\omega_b}F_{qr} + \frac{R_r}{x_{lr}}\left(\frac{x_{ml}^*}{x_{ls}}F_{ds} + \left(\frac{x_{ml}^*}{x_{lr}} - 1\right)F_{dr}\right)\right] \tag{6}$$

Where:

ω_b	: Motor angular base frequency
ω_e	: Electrical supply angular frequency
ω_r	: Rotor angular frequency (rad/s)
v_{qs}, v_{ds}	: Stator voltages (V) of q and d-axis
v_{qr}, v_{dr}	: Rotor voltages (V) of q and d-axis

2.1.2. *Magnetizing Flux*

The second part calculates the magnetizing flux with the equation shown by Eq. (7) and (8):

$$F_{mq} = x_{ml}\left[\frac{F_{qs}}{x_{ls}} + \frac{F_{qr}}{x_{lr}}\right] \tag{7}$$

$$F_{md} = x_{ml}\left[\frac{F_{ds}}{x_{ls}} + \frac{F_{dr}}{x_{lr}}\right] \tag{8}$$

$$x_{ml} = 1\Big/\left[\frac{1}{x_m} + \frac{1}{x_{ls}} + \frac{1}{x_{lr}}\right] \tag{9}$$

The stator and rotor current (A) in the arbitrary reference frame can then be solved as follows.

$$i_{qs} = \frac{1}{x_{ls}}\left(F_{qs} - F_{mq}\right) \quad (10)$$

$$i_{ds} = \frac{1}{x_{ls}}\left(F_{ds} - F_{md}\right) \quad (11)$$

Where:

i_{qs} , i_{ds} : stator current (A) of quadrature & direct axis

x_{ls} : leakage reactance (Ω) of stator

F_{qs} , F_{ds} : stator flux linkage (Wb-t) of quadrature & direct axis

2.1.2. *Electromagnetic torque and rotor speed*

Finally, the electromagnetic torque T_e (Nm) and rotor speed ω_r (rad/s) can be calculated by using Eq. (12) and (13).

$$T_e = \frac{3}{2}\left(\frac{p}{2}\right)\frac{1}{\omega_b}\left(F_{ds}i_{qs} - F_{qs}i_{ds}\right) \quad (12)$$

$$\frac{d\omega_r}{dt} = \frac{p}{2J}\left(T_e - T_l\right) \quad (13)$$

Where p is the number of poles, J (Kg-m^2) is the moment of inertia and T_l (Nm) is the input load torque.

2.2. ***Sinusoidal Pulse Width Modulation***

Three-phase voltage source inverter (VSI) typically made up of six power switches that demand proper switching sequence to transform the DC voltage supply into a variable frequency three-phase PWM voltage output. In this work, the Sinusoidal Pulse Width Modulation (SPWM) is used. To model the three-phase SPWM switching signals, three sinusoidal reference signals which are set 120-degree phase apart from each other; V_{ref_n} where n represents the 3 phase a, b, c, the signals are then compared with a high frequency triangular signal V_T. By changing the reference signal frequency f_{ref}, the inverter frequency can be change which then provides a simple control over the speed of the induction motor [12]. However, to produce the varying frequency reference signal, the standard sine function as shown in Eq. (14) is not applicable.

$$v_{refn}(t) = \sin\left(2\pi * f_{ref} * t\right) \quad (14)$$

Since the sin function takes argument of angles, simply changing f_{ref} by time will not produce a smooth frequency transition, jitter will be seen in the output as the results of jumping phase from one point of the waveform. To solve this issue, the use of phase accumulator is proposed with the Eq. (15) described below:

$$v_{refn}(t) = \sin\left(2\pi * \int f_{ref}(t)dt\right) \quad (15)$$

The phase accumulator can be built using the Backward Euler Method with Eq (16) and modeled using Simulink's User-Defined Function block.

$$v_{refn}[n] = \sin\left(2\pi * \left(\left(V_{refn}[k-1]\right) + \left(\left(t[k]\right) - t[k-1]\right) * f_{ref}[n]\right)\right) \quad (16)$$

The f_{ref} can now be varied as intended, even in

discontinuous ways, the phase will still change smoothly by time, producing a continuous variable frequency sinusoid.

2.3. ***Voltage Source Inverter***

One way of modeling a three-phase inverter is described by Bose in his book [6] based on the relation of phase voltages and pole voltages as in Eq. (17)-(19).

$$v_{an} = \frac{2}{3}v_{ao} - \frac{1}{3}v_{bo} - \frac{1}{3}v_{co} \quad (17)$$

$$v_{bn} = \frac{2}{3}v_{bo} - \frac{1}{3}v_{ao} - \frac{1}{3}v_{co} \quad (18)$$

$$v_{cn} = \frac{2}{3}v_{co} - \frac{1}{3}v_{ao} - \frac{1}{3}v_{bo} \quad (19)$$

Where:

v_{an}, v_{bn}, v_{cn} : Phase voltages

v_{ao}, v_{bo}, v_{co} : Pole voltages

To validate the mathematical model implementation, the results of the model describe in section 2 is compared with the Simulink's preset asynchronous machine drive which is already built and made available by MATLAB in its Simulink library. During simulation, both the mathematical and the Simulink's preset model use the same 5Hp motor with parameters as shown in Table 1.

Table 1: Induction Motor parameter for mathematical modeling.

Parameter	Abbreviation	Value	Unit
Rotor Resistance	Rr	1.083	Ohm
Stator Resistance	Rs	1.115	Ohm
Rotor Inductance	Lr	5.974E-03	Henry (H)
Stator Inductance	Ls	5.974E-03	Henry (H)
Magnetizing Inductance	Lm	0.2037	Henry (H)
No. of Poles	P	4	
Moment of inertia	J	0.02	kg.m^2

The results of both models output are presented in the next section.

3. Results

The induction rotor speed, phase current, and electromagnetic torque response are as presented in Fig. 13 showing side-by-side comparison of the mathematical model and the Simulink's preset model outputs with the applied load torque profile.

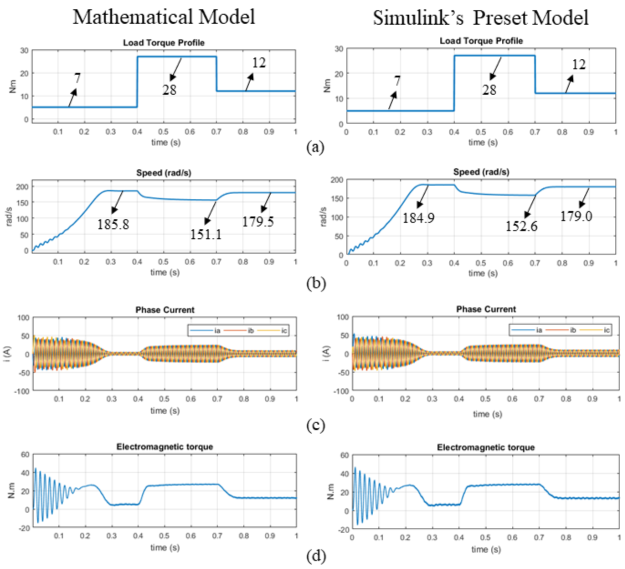


Fig. 13: Side by side Mathematical and Simulink's preset model comparison of (a) Load torque applied (b) Rotor speed (c) Phase currents (d) Electromagnetic torque.

From the results shown, the mathematical model and the Simulink's preset model yield similar response behavior. During starting, transient conditions are observed before the motors got into a steady state condition at time 0.3s. When the load torque is increased from 7Nm to 28Nm, both the model responds with decrease in rotor speed, increase in phase currents amplitude, and increase electromagnetic torque, transient region can also be observed whenever load torque change occurred before getting back into steady state condition.

4. Conclusion

This work has presented an approach to implement an SPWM fed three-phase induction motor drive. The model was built in Simulink entirely based on the equations describe in section 2. The implementation process is also presented in a step-by-step approach so that the work can be easily followed. The results produce by this model have shown similar behavior to that of Simulink's preset induction motor drive model, which validates that the mathematical model implementation is working as expected. This model implementation would be useful for future development of induction motor drive systems.

Note

The complete Simulink model implementation described in this paper can be referred at MATLAB Central File Exchange:
https://www.mathworks.com/matlabcentral/fileexchange/73260-dq-model-of-3-phase-induction-motor-drive

References

1. Asif MJ, Shahbaz T, Hassan SU, Rizvi STH. Mathematical modelling of 3-phase induction motor to study the Torque vs. Speed characteristics using MATLAB Simulink. In: *2016 19th International Multi-Topic Conference (INMIC)*. IEEE; 2016:1-7. doi:10.1109/INMIC.2016.7840144
2. Zangina U, Zainab U, Ilyasu MI. Implementation of the Three-Phase Induction Machine Model Using Simulink / S-Function. *IOSR J Environ Sci Toxicol Food Technol*. 2013;5(3):67-71. http://www.iosrjournals.org/iosr-jestft/papers/vol5-issue3/N0536771.pdf?id=3481.
3. Ozpineci B, Tolbert LM. Simulink Implementation of Induction Machine Model – A Modular Approach. In: Electric Machines and Drives Conference, 2003. IEMDC'03. IEEE International; 2003:728-734.

http://ieeexplore.ieee.org/document/1210317/.

4. Leedy AW. Simulink/matlab dynamic induction motor model for use as a teaching and research tool. *Int J Soft Comput Eng*. 2013;3(4):102-107.
5. Kim J, Cho E, Lee Y. C++ Based Dynamic Model of AC Induction Motor in Discrete Time Domain. In: *2016 IEEE Transportation Electrification Conference and Expo, Asia-Pacific (ITEC).* ; 2016:272-276.
6. Bose BK. *Modern Power Electronics and AC Drives*. Upper Saddle River, NJ 07458: Prentice Hall PTR; 2002.
7. da Silva ER, dos Santos E, Jacobina C. Pulsewidth Modulation Strategies. *IEEE Ind Electron Mag*. 2011;5(2):37-45. doi:10.1109/MIE.2011.941120
8. Rasyadan A, Ibrahim Z, Bin Tuan Dir TMA, Yaacob S. Modeling of Time Domain Analysis for Single and Double Open-Circuit Inverter Switch Faults in Three-Phase Induction Motor Drives. In: *2018 International Conference on Computational Approach in Smart Systems Design and Applications (ICASSDA).* IEEE; 2018:1-7. doi:10.1109/ICASSDA.2018.8477626
9. Ozpineci B, Tolbert LM. Simulink implementation of induction machine model - a modular approach. In: *IEEE International Electric Machines and Drives Conference, 2003. IEMDC'03.* Vol 2. IEEE; 2003:728-734. doi:10.1109/IEMDC.2003.1210317
10. Paul C. Krause, Oleg Wasynczuk SDS. *Analysis of Electric Machinery and Drive Systems*. Second Edi. (El-Hawary ME, ed.). John Wiley& Sons; 2002.
11. Chee-Mun O. *Dynamic Simulation of Electric-Machinery Using MATLAB Simulink.* Prentice Hall; 1998.
12. Husain I. *Electric and Hybrid Vehicles Design Fundamentals*. Boca Raton London New York Washington, D.C: CRC press; 2003.

Implementation of the X-means Clustering Algorithm for Wireless Sensor Networks

Abdelrahman Radwan, Nazhatul Hafizah Kamarudin, Mahmud Iwan Solihin, Hungyang Leong, Mohamed Rizon
Engineering Department, UCSI University,
Jalan Puncak Menara Gading, Taman Connaught,
Kuala Lumpur, 56000, Malaysia

Abstract

K-means clustering algorithms in wireless sensor networks (WSNs) are potential solutions that prolong the network lifetime. However, there are limitations hampering these algorithms, such as size of clusters and the number of clusters. This paper proposes implementing X-means algorithm as a new clustering technique that overcomes K-means limitations; clusters constructed using tentative CHs and tentative area of centroids in an initial phase, After that, if a cluster meets splitting criteria, new centroids selected and new clusters constructed.

Keywords: K-means, X-means, Clustering, Wireless, Sensors, Networks.

1. Introduction

Machine to machine communication (M2M) and Internet of Things (IoT) are the next wireless communication technologies, these technologies require low power communication networks[1]; Wireless Sensor network (WSN) is a potential candidate that provides an infrastructure for these technologies. WSN is a group of wireless nodes that equipped with sensors and actuators; the network designed as an application specific and perform dedicated tasks. Wireless nodes operate on batteries as a source of energy and transmit unit drains nodes batteries[2]; therefore, routing data considered a costly and complex task. In literature clustering algorithms have shown promising network life enhancements; classical algorithms such as LEACH[3], HEED[4], and EECS[5] divide their networks into random clusters, and each cluster has a cluster head (CH), nodes collect data from surrounding then rely them to CHs; CHs have to forward the received data to a central unit called Sink using a long range communication called backhauling. The role of CH is frequently rotated to distribute the load of backhaul communication among nodes and re-clustering initiated every round of operation. Another clustering Algorithms are K-means algorithms such as Refs. 6-10; these algorithms estimate the number of clusters and their initial cluster centroids, then average Euclidean distance to centroid determined and nodes assigned to minimum distance centroid, the process repeated until cluster centroid fixed and cluster members are static.

K-means algorithms overcome classical algorithms problems such as unbalanced intra-cluster consumption caused by random centroid location and non-optimal number of CHs; but still these algorithm can form clusters that requires energy resource beyond the CH capabilities such as high number of cluster members that exceed CHs channel time slots; therefore, this article propose to form initial clusters using K-means clustering then X-mean algorithm used to evaluate formed clusters based on splitting criteria and to finalize list of clusters.

The rest of the article is as following, in section two the system model that includes Energy model and distance crossover model, then in section three we propose X-mean algorithm and describe its two operation phases with mathematical equations in details, after that in section four the simulation results and discussion on how the proposed algorithm prolong WSN lifetime is presented, finally in section we conclude our work with future recommendations to further enhance X-means algorithm.

2. System Model

The system model obtained from Ref.3, and it includes two parts; a consumption model for transmit, receive and data processing as in Eq. (1)(2)(3), and a distance model that estimates the crossover distance as in Eq. (4).

$$E_{tx}(b,d) = E_{elec}(b) + E_{tx-amp}(b,d) \tag{1}$$

$$E_{Tx} = \begin{cases} b * E_{elec} + b * \varepsilon_{fs} * d^2, \ d < d_o \\ b * E_{elec} + b * \varepsilon_{mp} * d^4, \ d \geq d_o \end{cases} \quad (2)$$

$$E_{Rx} = b * E_{elec} \quad (3)$$

$$d_o = \sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp}}} \quad (4)$$

For part one, consumption model Eq. (1) represents the transmission consumption and it consists of E_{elec} that is the energy required to aggregate and process a single bit of data and $E_{tx\text{-}amp}$ that is the amplification energy required to transmit data over (d) distance. Eq. (1) expanded to Eq.(2) that describes two types of the amplification energies, where ε_{fs} describes the amplification energy required to transmit data using the freespace model, if the distance between source and destination is less than the crossover distance (d_0); otherwise, ε_{mp} describes the amplification energy that is required to transmit data using the multipath fading model. the crossover distance between the freespace model and the multipath model has a threshold (d_o), it determined by Eq. (4) and described in Ref. 7. The total transmission energy (E_{Tx}) is proportional to the distance between a source node and a destination node, where it corresponds to d^2 when d is less than d_0 and corresponds to d^4 when d is greater than or equal d_0. Receiving consumption (E_{Rx}) represented by Eq. (3) and includes the energy required to process a single bit.

3. Proposed Method

K-means algorithm depend on a predefined value of K to form clusters, K cluster centroids selected randomly, but even though K-means algorithm are deterministic because initial K centers have to be provided; a bad choice of initial centroids leads to poor clustering structures and performance. Therefore, we propose to implement X-means as a solution for K-means problems. The concept of the X-means clustering can be summarized into two main phases; at the first phase random k-means applied to select initial centroids, then nodes assigned to these initial centroids and new centroids position determined by Eq. (5), the process repeated recursively until a final copy of centroids is determined, the final copy of centroids called parent centroids. Then in phase two, every parent gives birth to two children $C_{n1,2}$, the initial location of the children determined as in Eq. (6); it is worth noticing that the children initial positions are the parent (x,y) position, but adjusted by the Euclidean distance (d_n); here the Euclidean distance is the distance between the parent centroid and the furthest node in the cluster as in Eq. (7).

$$P_n = \alpha_{n-1}(x,y) + \frac{1}{S_n}\sum_{i=1}^{S}\beta_i(x,y) \ \begin{cases} n \in \ 1,2,3 \, .. \, k \\ i \in \ 1,2,3 \, .. \, S_n \end{cases} \quad (5)$$

$$C_{n1,2} = P_n \pm d_n \quad (6)$$

$$d_n = max\left(\left\{\sqrt{(x_n - x_i)^2 + (y_n - y_i)^2}\right\}\right), \ i \in S_n, n \in P_n \quad (7)$$

P_n: Parent (n) centroid.
$\alpha_{n-1}(x,y)$: x and y position of previous (n-1) centroid.
S_n: Set of nodes assigned to P_n.
$\beta_i(x,y)$: the x and y position of node (i).
d_n: the maximum distance between parent centroid and the furthest member node (i).

The new positions are the new children $C_{n1,2}$ and nodes assigned to these children according to received signal strength or distance. Then X-means adjust children centroid positions. The mechanism of X-means fairly designed to minimize backhaul distance by adjusting cluster centroids according to their distance to sink as in Eq.(8). Nodes grouped in clusters based on their average distance from children centroids and their distance to sink as in Eq. (9).

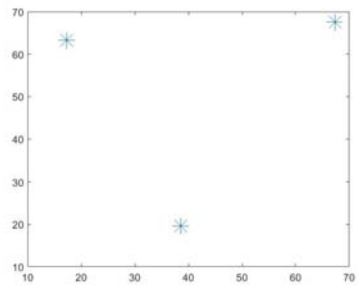
$$CC_{n1,2} = \emptyset(x,y) + \frac{1}{S_n}\sum_{i=1}^{S}\beta_i(x,y) \ \begin{cases} n \in \ 1,2,3 \, .. \, k \\ i \in \ 1,2,3 \, .. \, S_n \end{cases} \quad (8)$$

$$AvgD = \ min((d_c + d_{sink})/2) \quad (9)$$

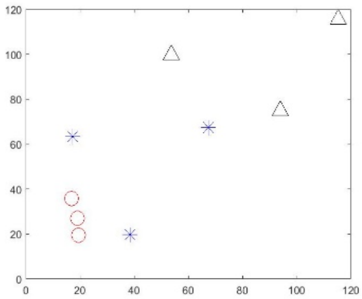
$CC_{n1,2}$: position of parent (n) Children Centroid.
$\emptyset(x,y)$: position of Sink.
AvgD: average distance to centroid and sink.
d_c: A node Euclidean distance to child centroid.
d_{sink}: A node Euclidean distance to sink.

X-means algorithm has to execute recursively until there are no changes at centroids position, the algorithm has three outcomes, first children form their own cluster and parents collapsed, second outcome some of the children and parent diminished, and third outcome children collapsed and parent repositioned to best locations. Fig. 1. shows the new generated children centroids, the first generation is parents centroid generated by k-means as in Fig. 1-A, and the number of generated parents are three, then the implementation of X-means generated children as in Fig. 1-B (marked with triangle and circles), the total of children and parents becomes nine centroids, finally the recursive run of X-means diminishes some of the parents, children, and repositions the remaining children as in Fig. 1-C, the final centroids become static clusters of the network.

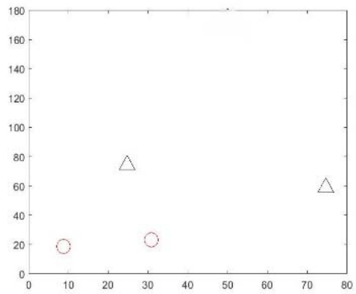
Nodes that have minimum distances to centroids selected to become CHs, then if a CH energy dropped below the energy threshold, it steps down and becomes a cluster member and the next closest node to the cluster centroid



A) Parent centroids generated after running K-means



B) Prents give birth to new children centroids



C) Remaining centroids after running x-means

Fig. 1. X-means two phase of centroid generation

become a CH. Once a node energy dropped below the threshold the node determines a new threshold, the new threshold is a step down from the original threshold by (σ) as in Eq. (10).

$$E_{th} = E_{th} - (\sigma * E_{th}) \quad (10)$$

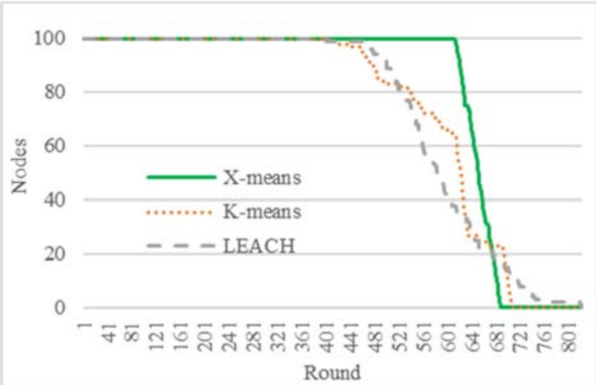
The propose algorithms simulated in MATLAB and the simulation parameters listed in table (1). There are two simulation scenarios considered; first a random deployment of 100 nodes in an area 100 x100, second a random deployment of 200 nodes in an area of 200 x 200. The result is benchmarked against prominent algorithm such as LEACH [3], and a centralized K-means algorithm[7]. X-means implemented to prolong the network lifetime and the deployed sensor nodes always have data to transmit; therefore, the network lifetime measured in round and a single round represents a single packet received from a CH at the sink. At simulated scenarios the sink positioned outside the monitoring field at (50,175) and (100,275) for first and second scenarios respectively. X-means implemented on the sink and its aware of nodes position, thus; clusters formed by sink, but rotating of CH role is a node decision.

Table 1: X-means simulation parameters

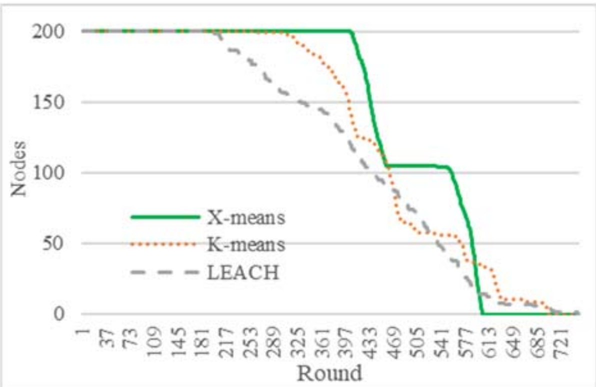
Parameter	Description	Value
d_0	Crossover distance	87m
ε_{fs}	Freespace model amplification energy	10pJ/bit/m²
ε_{mp}	Multipath amplification energy	0.0013pJ/bit/m4
E_{elec}	Single bit processing energy	Tx or Rx = 50nJ/bit Aggregation = 5nJ/bit
E_{init}	Initial Energy	0.5J
K	Initial centroids	3
E_{th}	Energy threshold	0.2J
σ	E_{th} step down	5%
Packet Size	Data size	4.2Kb/packet
Control Packet	Packet Headers	0.2Kb/packet

4. Results and Discussion

The main concept of X-means is to select the best centroid positions, these positions must minimize nodes consumption per round by reducing intra-cluster communication consumption and reducing CHs backhaul consumption. In X-means number of centroids are not determined, but only initial centroids are set. The number of initial centroids has to be set to half of the maximum allowed centroids, for example, in scenario one, the maximum allowed CHs is 5% of the total networks nodes[3], so the number of CHs is 3 after rounding up. After executing X-means the number of parents generated were 3, total children were 6, and final number of centroids were 4, the results of this enhancements shown in Fig. 2-A, the enhancement of centroids number and positions reflect as enhancement of First Node Death (FND), where using X-means algorithm extend FND by 53% over LEACH and 51% K-means. Similarly for the second scenario, the maximum number of CHs is 5%; therefore the maximum number of CHs is 10, and initial centroids/parents required by X-means is 5 centroids and generated children are 10, but the final number of centroids remained 5 and only position of centroids adjusted. The result of simulating the second scenario also shows that x-means prolonged the FND and enhanced the performance over LEACH and K-means. It is worth noticing in X-means network that in both scenarios, after FND occurred a sharp drop in the number of alive nodes occurred too; which indicates that the X-means mechanism in selecting centroid positions and grouping nodes based on their distance from the sink has nearly balanced nodes consumption. In the second scenario, the sharp drop of number of alive nodes stopped at 105 nodes, this because the nodes at the far end of the network depleted more energy during backhaul communication with the sink than those at closer distances to the sink.



A) 100 node networks



B) 200 node networks.

Fig. 2. X-means performance simulation and benchmarking in multiple scenarios.

In our simulation, we also monitored the total network consumption in three periods of network lifetime; the first period represents the duration from the network startup time until 1% of nodes death, then the second period is the energy spent from 1% of nodes death until 50% of nodes death, after that the third period from 50% of nodes death until 100%. X-means algorithm has shown less consumption than LEACH and K-means at the first and the second periods because of centroid optimization, but X-means and K-means consumed higher energy at the end of network lifetime. this is because of static clustering where most of the nodes at close distance from centroids have died and the remaining nodes has high intra-cluster distance, which unlike LEACH that depends on random clustering.

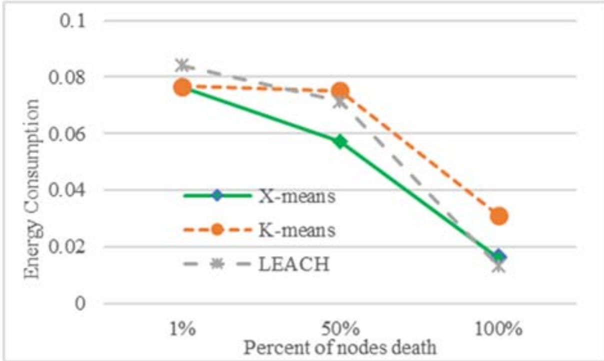


Fig. 3. The average network energy consumption during network lifetime.

5. Conclusion

In this article we proposed X-means as a solution for K-means shortcomings. The algorithm depends on determined value of parents to generate two children for each parent, then children evaluated to generate a final copy of centroids that is less than the maximum number of network centroids and these centroids have to enhance network performance. The simulation scenarios were limited to 100 node deployments and 200 node deployments, thus; the performance is not checked for large scale deployment such as 1000 nodes.

In future work, the design of X-means need to be adjusted to include more than 2 children per parent, in order to increase the chance of finding minimum cost centroids.

Acknowledgements

The author would like to thank UCSI university Trust Fund and Centre of Excellence for Research, Value Innovation and Entrepreneurship (CERVIE) for support this work.

References

1. Li X, Li D, Wan J, Vasilakos AV, Lai CF, Wang S. *A review of industrial wireless networks in the context of industry 4.0. Wireless networks.* 2017 Jan 1;23(1):23-41.
2. Yick J, Mukherjee B, Ghosal D. *Wireless sensor network survey.* Computer networks. 2008 Aug 22;52(12):2292-330.
3. Heinzelman WR, Chandrakasan A, Balakrishnan H. *Energy-efficient communication protocol for wireless*

microsensor networks. In Proceedings of the 33rd annual Hawaii international conference on system sciences 2000 Jan 7 (pp. 10-pp). IEEE.

4. Younis O, Fahmy S. *HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks.* IEEE Transactions on mobile computing. 2004 Oct 1(4):366-79.
5. Ye M, Li C, Chen G, Wu J. *EECS: an energy efficient clustering scheme in wireless sensor networks.* InPCCC 2005. 24th IEEE International Performance, Computing, and Communications Conference, 2005. 2005 Apr 7 (pp. 535-540). IEEE.
6. Periyasamy S, Khara S, Thangavelu S. *Balanced cluster head selection based on modified k-means in a distributed wireless sensor network*. International Journal of Distributed Sensor Networks. 2016 Mar 1;2016:2.
7. Park GY, Kim H, Jeong HW, Youn HY. *A novel cluster head selection method based on K-means algorithm for energy efficient wireless sensor network.* In2013 27th International conference on advanced information networking and applications workshops 2013 Mar 25 (pp. 910-915). IEEE.
8. Sasikumar P, Khara S. *K-means clustering in wireless sensor networks.* In2012 Fourth international conference on computational intelligence and communication networks 2012 Nov 3 (pp. 140-144). IEEE.
9. Razzaq, M., Ningombam, D.D. and Shin, S., 2018, January. *Energy efficient K-means clustering-based routing protocol for WSN using optimal packet size.* In 2018 International Conference on Information Networking (ICOIN) (pp. 632-635). IEEE.
10. Khan A, Tamim I, Ahmed E, Awal MA. *Multiple parameter based clustering (mpc): Prospective analysis for effective clustering in wireless sensor network (wsn) using k-means algorithm.* Wireless Sensor Network. 2012 Jan 1;4(1):18-24.

Robust H_∞ controller design for flexible link manipulator based on constrained meta-heuristics optimization algorithms

Mahmud Iwan Solihin, Lim Wei Hong, Chun Kit Ang, Mohamed Rizon, Abdelrahman Radwan
Faculty of Engineering, UCSI University. Taman Connaught, Cheras, 56000
Kuala Lumpur, Malaysia

Abstract

Robust control of flexible manipulator is interesting and challenging due to its flexible nature and under actuation characteristic. Problems arise due to precise positioning requirements and system flexibility which leads to vibration. A robust tracking control for flexible link manipulators should be able to produce a fast-tracking and accurate response of desired angular position while eliminating the arm's tip vibrations quickly. However, parameters tuning of a robust controller is a mathematically rigorous by conventional approach. In this paper, H-infinity (H_∞) robust controller tuning using meta-heuristics optimization is proposed. The performance of the designed controller will be evaluated and compared for different meta-heuristics optimization algorithms namely Cuckoo Search (CS), Teaching Learning-based Optimization (TLBO) and Black Hole Optimization Algorithm (BHOA). These algorithms are used in the constrained optimization to minimize the H_∞ norm subjected to a specified location of the poles of the feedback (closed-loop) system. The simulation results indicate that these algorithms, especially BHOA, can be a good choice of optimizer since it is free from algorithm-specific preset parameters.

Keywords: robust H-infinity control, flexible link manipulator, cuckoo search, teaching learning-based optimization, black hole optimization algorithm.

1. Introduction

Research on control systems design of flexible manipulators has received a lot of attention since the last some decades due to their several advantages over rigid ones [1,2]. Link flexibility is a consequence of the lightweight constructional feature in manipulator arms that are designed to operate at high speed with low inertia. Thus, flexible link manipulators require less material, are lighter in weight, consume less power, require smaller actuators, are more maneuverable and transportable, have less overall cost and higher payload to robot weight ratio [3,4].

However, the control of flexible manipulators to achieve and maintain accurate positioning is challenging due to its flexible nature and underactuated characteristic. The dynamics is even more complex with parameter variation of the system. Problems arise due to precise positioning requirements and system flexibility which leads to vibration.

On the other hand, optimization is seen as a suitable approach to deal with complex and conflicting requirements in engineering design including control systems design discussed in this paper. Optimization algorithms with stochastic (randomness) components were often referred to as heuristic in the past, though the recent literatures tend to refer them as meta-heuristics. Meta-heuristic algorithm is regarded as further development of the heuristic algorithms as the word 'meta' means beyond or higher level [5,6].

Some famous meta-heuristic algorithms are: genetic algorithm (GA), simulated annealing (SA), taboo search (TS), ant colony optimization (ACO), differential evolution (DE), particle swarm optimization (PSO), bacterial for-aging optimization (BFO), harmony search (HS), bee algorithms (BAs), firefly algorithm (FA), cuckoo search (CS), bat-inspired algorithm (BIA) and recently teaching learning-based optimization (TLBO) and Black Hole Optimization Algorithm (BHOA). Their modified versions (variants) are also proposed and applied to solve global optimization problems [7,8,9,10]. These kind of studies can be considered as applications of artificial intelligence and optimization for robotics, control engineering and other field of applications as found in many other works [11,12,13,14].

In this paper, robust H_∞ controller tuning using meta-heuristics optimization is proposed. The performance of the designed controller will be evaluated and comparison for recently proposed meta-heuristics optimization algorithms using Cuckoo Search (CS)[15], Teaching Learning-based Optimization (TLBO)[16,17] and Black Hole Optimization Algorithm (BHOA)[18,19] will be discussed.

2. Problem Formulation

Consider a plant model of linear time-invariant continuous-time system in state space form:

$$\dot{x}(t) = Ax(t) + Bu(t)$$
$$y(t) = Cx(t) + Du(t) \tag{1}$$

where $x \in R^L$, $u \in R^M$ and $y \in R^P$ are state vector, control input and output vector respectively. It is assumed that the system given in (1) is completely state controllable and all state variables are available for feedback. One can use state feedback controller with forward integral gain (k_i) as shown in Fig. 1. The controller gains, $K = [k_1, k_2, \dots, k_n, k_i]$, can be computed using various methods such as conventional optimal control method to minimize the performance index.

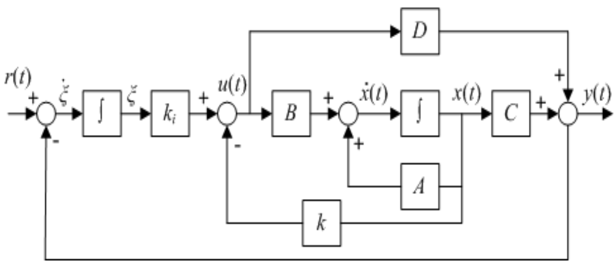


Fig.1. State feedback controller with integral gain (k_i).

In this study, a constrained optimization using metaheuristic algorithms is adopted to search for a set of robust controller gains so that the plant uncertainty is automatically handled with the use of robust H_∞ norm that will be discussed in the next section. In addition, a region of closed loop poles is incorporated as optimization constraint to allow the designers to define the desired time-domain control performance.

Briefly, the H_∞ norm is used as cost function to minimize in this robust feedback control design via optimization. This is performed to simplify and automate the design process using meta-heuristic optimization algorithms since mathematical formulation for the robust control design using conventional H_∞ approach is rigorous. Minimizing the H_∞ norm of the system will guarantee the robust performance of the optimized controller.

The H_∞ norm of a dynamic system is defined as the peak gain of the frequency response when the system is stable and returns to infinity when system is unstable. For MIMO (multi input multi output) systems, H_∞ norm is the maximum singular value over the frequency space. Suppose that the closed loop transfer function of the system is $T(s)$, then the H_∞ norm of this system is defined as:

$$\|T\|_\infty := \underset{\omega}{max_\omega} \quad |T(j\omega)| \tag{2}$$

For the feedback system as shown in Fig. 1, the closed loop transfer function can be expressed as the following:

$$\|T\|_\infty = \left\| C\left(sI - \hat{A}\right)^{-1}\hat{B} \right\|_\infty \tag{3}$$

where $\hat{A}$ and $\hat{B}$ are given by the following equations:

$$\hat{A} = \begin{bmatrix} \tilde{A} & -Bk_i \\ -C_1 & 0 \end{bmatrix}, \qquad \hat{B} = \begin{bmatrix} B \\ 0 \end{bmatrix}, \tag{4}$$

with $\tilde{A} = A - \mathrm{B}[k_1, \dots, k_n]$ and $C_1 = [1 \dots 0]$.

Furthermore, to handle the time domain performance of the system, the closed loop poles (eigenvalues, λ_n) of the system is placed in a specific wedge region (ψ) as shown in Fig. 2. This is treated as the constraint in the optimization. Damping ration of the system is set as a common choice ($\zeta = 0.7$). The transient margin (ρ) is specified according to the desired speed of the response. This is problem-dependent parameter and the value of ρ=2 will be chosen in the case of this study.

Overall, the robust feedback controller design via optimization in this study can be summarized in the following procedure:

Minimize: H_∞ norm of the system: $\|T\|_\infty$

Subject to constraint:

$\lambda_n(X) \in \psi$ for n=1,2,…

and **boundary constraint**:

$X \in [l_b, u_b]$

where $X = K = (k_1, k_2, \ldots, k_n, k_i)$ is the vector solutions such that $X \in S \subseteq R^{n+1}$. S is the search space, and $F \subseteq S$ is the feasible region or the region of S for which the constraint is satisfied. The lower and upper bound (l_b and u_b) of the solution candidates (controller gains) are set according to the designer's common sense such as ± 100 or ± 50.

The more detailed explanation of the constraint handling technique is not presented in this paper for brevity. One can refer to the pervious works [20,21,22].

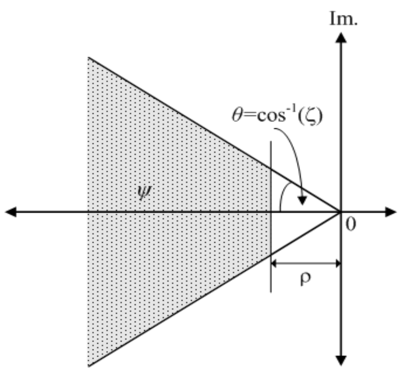


Fig. 2. A wedge region for closed loop poles placement.

3. Overview of The Optimization Algorithms

In this study, three different meta-heuristic algorithms will be used. There have been many meta-heuristic algorithms proposed as mentioned. However, only few recently proposed algorithms are used for comparison in this paper namely: Cuckoo Search Optimization (CSO)[15], Teaching Learning-based Optimization (TLBO)[16,17] and Black Hole Optimization Algorithm (BHOA)[18,19]. The key of meta-heuristic algorithms is that mostly nature-inspired and the solution found is dependent on the set of random variables generation in the solution searching mechanism. Despite many meta-heuristic algorithms have been published with claims of novelty and practical efficacy, many of the publications have been of poor quality, flaws include vagueness, lack of conceptual elaboration, poor experiments, and ignorance of previous literatures [23].

However, the interesting aspect of these recently proposed algorithms is that for example CSO only need algorithm-specific control parameter (discovery rate, P_a). Even TLBO and BHOA does not require any algorithm-specific control parameters. They only need the common preset control parameters such as the population size and the number of generations/iterations. The technical details of these algorithms are not discussed in this paper for brevity. Briefly, these algorithms can be implemented with the main parameters as follows. The parameters shown in Table 1 below are used for this project.

Table 1. The optimization algorithms parameters

CSO	TLBO	BHOA	Note
Np=50	Np=50	Np=50	Np: population sise
Iter=200	Iter=200	Iter=200	Number of iteration/generation.
P_a=0.25	–	–	P_a=discovery rate for CSO. TLBO and BHOA are free from control parameter.
[l_b,u_b]=[-100, 100]; Lower and upper bound for variable solution			
All initial random numbers are set to 'default'			

4. Results for Robust Control Design of Flexible Link Manipulator

The optimization run has been done in MATLAB 2019 environment. The results and evaluation is presented in this section. Fig.3 shows the schematic diagram of the flexible link manipulator. The input of the system is voltage input (V) to the actuator (DC motor). The outputs are motor angular position (θ) and arm's tip deflection angle (α). The linearized state-space model of the system used in this study can be found in ref (1) [1] and is not presented here for brevity.

Table 2 shows the robust feedback controller gain optimized by CSO, TLBO and BHOA with the preset optimization parameters as given in Table 1.

Table 2. The optimization algorithms parameters.

Algorithm	Controller gain, $\boldsymbol{K = [k_1, k_2, k_3, k_4, k_i]}$
CSO	K=[71.32, -55.21, 14.79, -8.66, -98.33]
TLBO	K=[30.83, 0.85, 10.19, -15.62, -28.31]
BHOA	K=[37.68, 10.39, 9.81, -19.20, -46.92]

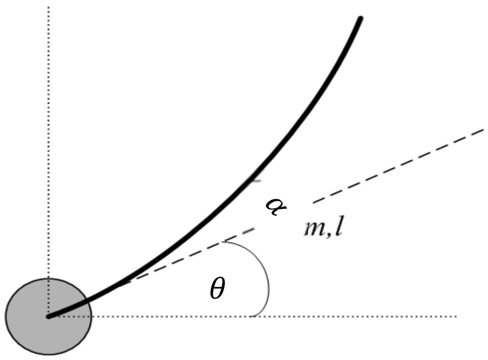


Fig. 3. Flexible link manipulator schematic

The response of the system is shown in Fig. 4 where the tip angular deflection (α) and the arm position angle (θ) are evaluated by giving a square wave input signal with amplitude of ±90°. As can be seen that all three controllers are able to produce good performance in terms of stabilization and time-domain criterion. Table 3 shows the time domain performance in more detailed. The obtained result of BHOA is preferable due to as fast as CSO in the settling time (for θ) but producing lower vibration amplitude (for α) than CSO. TLBO is able to produce lower amplitude due to its slower response as can be seen in Table 3.

Table 3. The time-domain response performance.

Controller	Settling time (for θ)	Max amplitude (for |α|)
CSO	2.1 sec	2.8°
TLBO	2.9 sec	1.1°
BHOA	2.1 sec	1.9°

5. Conclusions

A robust control design for flexible link manipulator using meta-heuristic optimization algorithms have been proposed and their effectiveness have been evaluated. The more detailed comparative analysis and real-time experimentation should be conducted in the future. The initial indication shows that BHOA (black hole optimization algorithm) can be a good choice of optimizer since it is free from algorithm-specific control parameters.

6. References

1. Solihin, M. I., Wahyudi & Akmeliawati, R. PSO-based optimization of state feedback tracking controller for a flexible link manipulator. in *SoCPaR 2009 - Soft Computing and Pattern Recognition* (2009). doi:10.1109/SoCPaR.2009.26
2. Khafaji, A. A. Al & Darus, I. Z. M. Controller Optimization Using Cuckoo Search Algorithm of a Flexible Single-Link Manipulator. *Proc. 2014 First Int. Conf. Syst. Informatics, Model. Simul.* 39–44 (2014). doi:10.1109/sims.2014.16
3. Tokhi, M. O., Mohamed, Z., Martins, J. M., Botto, M. A. & Sá da Costa, J. Approaches for dynamic modelling of flexible manipulator systems. *IEE Proc. - Control Theory Appl.* **150**, 401–411 (2003).
4. Solihin, M. I., Wahyudi, Legowo, A. & Akmeliawati, R. Comparison of LQR and PSO-based state feedback controller for tracking control of a flexible link manipulator. in *ICIME 2010 - 2010 2nd IEEE International Conference on Inf. Management and Engineering* **3**, (2010).
5. Beheshti, Z. A Review of Population-based Meta-Heuristic Algorithm. *SOCO 2013* (2013).
6. Madhuri & Deep, K. A state-of-the-art review of population-based parallel meta-heuristics. in *2009 World Congress on Nature & Biologically Inspired Computing (NaBIC)* 1604–1607 (IEEE, 2009). doi:10.1109/NABIC.2009.5393657
7. Meng, K., Hong, W., Ashidi, N., Isa, M. & Sun, S. A constrained multi-swarm particle swarm optimization without velocity for constrained optimization problems. *Expert Syst. Appl.* **140**, 112882 (2020).
8. El-Sherbiny, M. M. Particle swarm inspired optimization algorithm without velocity equation. *Egypt. Informatics J.* **12**, 1–8 (2011).
9. Zou, F., Chen, D. & Wang, J. An Improved Teaching-Learning-Based Optimization with the Social Character of PSO for Global Optimization. *Comput. Intell. Neurosci.* **2016**, 1–10 (2016).
10. Natarajan, E., Kaviarasan, V., Lim, W. H., Tiang, S. S. & Tan, T. H. Enhanced Multi-Objective Teaching-Learning-Based Optimization for Machining of Delrin. *IEEE Access* **6**, 51528–51546 (2018).
11. Tang, S. H., Ang, C. K., Ariffin, M. K. A. B. M. & Mashohor, S. B. Predicting the Motion of a Robot Manipulator with Unknown Trajectories Based on an Artificial Neural Network. *Int. J. Adv. Robot. Syst.* **11**, 176 (2014).

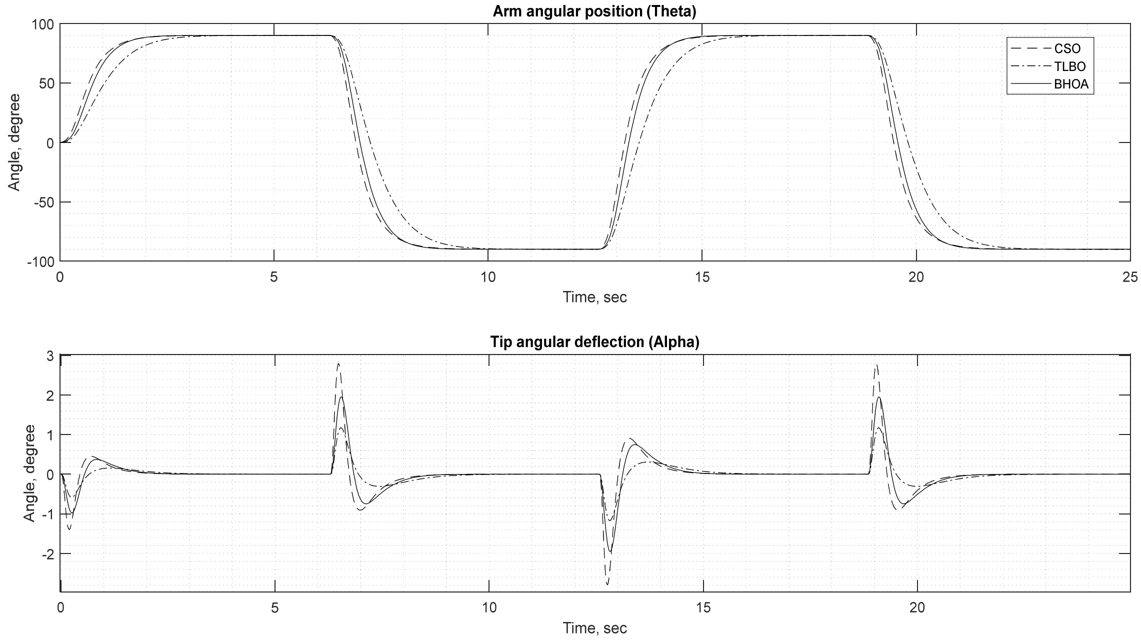


Fig. The system response to the square wave input signal.

12. Hong, T. S., Kit, A. C., Nia, D. N., Ariffin, M. K. A. M. & Khaksar, W. Planning for Redundant Manipulator Based on Back-Propagation Neural Network. *Adv. Sci. Lett.* **19**, 3307–3310 (2013).
13. Abdullah Al-Sanabani, D. G. *et al.* Development of non-destructive mango assessment using Handheld Spectroscopy and Machine Learning Regression. *J. Phys. Conf. Ser.* **1367**, 012030 (2019).
14. Frederick, H., Winda, A. & Iwan Solihin, M. Automatic Petrol and Diesel Engine Sound Identification Based on Machine Learning Approaches. *E3S Web Conf.* **130**, 01011 (2019).
15. Yang, X.-S. & Deb, S. Cuckoo search: recent advances and applications. *Neural Comput. Appl.* **24**, 169–174 (2014).
16. Venkata Rao, R. & Waghmare, G. G. A comparative study of a teaching–learning-based optimization algorithm on multi-objective unconstrained and constrained functions. *J. King Saud Univ. - Comput. Inf. Sci.* **26**, 332–346 (2014).
17. Rao, R. V., Savsani, V. J. & Vakharia, D. P. Teaching–learning-based optimization: A novel method for constrained mechanical design optimization problems. *Comput. Des.* **43**, 303–315 (2011).
18. Hatamlou, A. Black hole: A new heuristic optimization approach for data clustering. *Inf. Sci. (Ny).* **222**, 175–184 (2013).
19. Kumar, S., Datta, D. & Singh, S. K. Black Hole Algorithm and Its Applications. in 147–170 (2015). doi:10.1007/978-3-319-11017-2_7
20. Solihin, M. I. & Akmeliawati, R. Robust Feedback Controller Design Using Cuckoo Search Optimization to Maximize Stability Radius. in *Robot Intelligence Technology and Applications* 62–75 (Springer, 2019). doi:10.1007/978-981-13-7780-8_6
21. Solihin, M. I., Akmeliawati, R., Tijani, I. B. & Legowo, A. ROBUST STATE FEEDBACK CONTROL DESIGN VIA PSO-BASED CONSTRAINED OPTIMIZATION. *Control Intell. Syst.* **39**, (2011).
22. Solihin, M. I., Akmeliawati, R., Muhida, R. & Legowo, A. Guaranteed robust state feedback controller via constrained optimization using Differential Evolution. in *Proceedings - CSPA 2010: 2010 6th International Colloquium on Signal Processing and Its Applications* (2010). doi:10.1109/CSPA.2010.5545279
23. Sörensen, K. Metaheuristics-the metaphor exposed. *Int. Trans. Oper. Res.* **22**, 3–18 (2015).

Classification of Facial Nerve Paralysis Based on Regional Evaluation

Wan Syahirah W Samsudin[1*]
Faculty of Electrical & Electronics Technology Engineering,
Universiti Malaysia Pahang, 26600 Pekan, Pahang, Malaysia.

Rosdiyana Samad
Faculty of Electrical & Electronics Technology Engineering,
Universiti Malaysia Pahang, 26600 Pekan, Pahang, Malaysia.

Kenneth Sundaraj
Faculty of Electronics & Computer Engineering,
Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia.

Mohamed Rizon
[3]Faculty of Engineering, Technology and Built Environment,
UCSI University, No 1, Jalan Menara Gading, Taman Connaught, 56000 Kuala Lumpur, Malaysia.

Mohd Zaki Ahmad
Department of Otorhinolaryngology,
Hospital Tuanku Ampuan Afzan, Jalan Tanah Putih, 25100 Kuantan, Pahang, Malaysia.
E-mail: wansyahirahwsamsudin@yahoo.com, rosdiyana@ump.edu.my, kenneth@utem.edu.my,
MohdRizon@ucsiuniversity.edu.my, drmohdzaki@yahoo.com.my

Abstract

This paper present an approach of regional evaluation using Kanade - Lucas – Tomasi (KLT) method and extract feature points to classify the normal and patients subjects and also to determine the severity level of paralysis for each region of face. There are Individual Score Table, Total Score Table and Paralysis Score Table were presented. A 100% of accuracy has been obtained in identifying the paralysis with k =3 using k-NN classifier for the Individual Score.

Keywords: Facial Nerve Paralysis, Facial Nerve Evaluation, Bells' Palsy, Kanade-Lucas-Tomasi, Optical Flow

1. Introduction

Contributions are to Facial nerve paralysis is the impairment of the motor function of facial muscles due to injury to the facial nerve, which is the seventh (VII) cranial nerve. The nerve emerges from the pons of the brainstem and controls the muscles of facial expression. Besides, it also functions in transporting the taste sensations from the tongue.[1] Hence, the patient with this paralysis will face the numbness on the affected side of face even though no actual sensory loss occurs.[2] They will suffer from the serious functional, cosmetic and physiological problems and demonstrate the impairment in the ability of communication both verbally and non-verbally.[2] Even though it is not the fatal disease, however the facial paralysis is one of the most devastating

*Wan Syahirah W Samsudin, Faculty of Electrical & Electronics Technology Engineering, Universiti Malaysia Pahang, 26600 Pekan, Pahang.

peripheral nerve injuries as the face is the crucial component of the beauty, sexual attractiveness and interest.[3] As proved in the literature, the prevalent cause of this facial paralysis and been reported approximately in 50% of the cases is Bells' Palsy, an idiopathic disease which named after Sir Charles Bell (1774-1842).[4-6]

Based on the International Database of the US Census Bureau[7], an extrapolation of the incidence rate of Bells' Palsy has been made and shows that the highest incidence rate is China which experience about 191,007 cases followed by India and USA, which has 156,628 and 43, 184 cases, respectively.[8] Thus, the clinicians and therapist who deals with this facial paralysis daily are acutely aware of the importance to restore the function of facial nerve to the highest possible results in improving of health, self-esteem and the quality of life of patient. In order to achieve that, a thorough evaluation for assessing the degree of paralysis is needed to measure the facial disability from onset to the various stages of paralysis. The classification of facial nerve paralysis is not only crucial in the decision making of the treatment to the patient, but also helpful in detecting the changes or recovery of patient over time or after the treatment. A dozen of works have been proposed in the past few decades, but most of the existing systems are subjective[9-11], where the evaluation is based on the naked eye of clinicians on the movement of patient and these produce the variations in results. A detailed summarization of these grading systems can be found in our previous work.[12] The most widely used subjective grading scale is House-Brackmann score[13]. To overcome the limitation of subjective system, many objective systems have been developed to provide the clinicians a quantitative measurement of facial nerve function, however, surprisingly, to date, no standard facial nerve evaluation system has been universally accepted and used worldwide.[14]

In addition, regional evaluation is very important to know which region is having more problem and the treatment and rehabilitation can be designed based on the evaluation. The evaluation also can be eliminate the case of synkinesis happens.

Based on above discussion, the purpose of this paper is to present a classification method of facial nerve paralysis based on regional parts of face. This method is not only succeeded in giving the scores for each important regions of face, but also successfully classified the severity of paralysis into six levels of paralysis according to well-known House-Brackmann score. The objective of this system is to assist the clinicians in the assessment of facial nerve functions through the quantitative measurement of patients' performance during facial movement.

2. Materials and Methods

2.1. *Ethical Statement*

The procedures in the experiment was approved by the Medical Research and Ethics Committee (MREC), Ministry of Health, Malaysia (Ref No.: KKM/NIHSEC/800-2/2/2/P13-96). The study was conducted in accordance to Declaration of Helsinki due to human participation as subjects and the consent were obtained prior the experiment.

2.2. *Experimental Setup*

A handy camera (SONY DCR-SX45E) with a 640 x 480 resolution, was placed in front of the subject at a distance of 1.0 m, to capture the frontal view of the subjects' face with a rate of 30 frames per second. An additional lighting system with an adequate illumination was developed to ensure the quality of the video recordings were adequate to be processed. The data was collected at the Physiotherapy and Rehabilitation Department of Hospital Tuanku Ampuan Afzan (HTAA), Kuantan, Malaysia.

2.3. *Data Acquisition*

The publicly video database that satisfies the requirement of this study was not found. Hence, a database was developed with 62 data of normal subjects and 40 data of patients. The normal subjects have symmetrical facial function, where the right side of the face moves symmetrically with the left side. The patients in this study presented with the same etiology which is Bells' Palsy, which is the prevalent cause of facial paralysis and the age of subjects are ranged from 23-75 years old.

The data acquisition was performed through video recordings. Subjects were required to familiarize themselves with the facial exercises needed and follow the designed protocol prior to recording sessions. Subjects were asked to sit upright in comfortable position on chair, keep the head motionless and look straight at the camera while performing all the exercises. Verbal reminder to relax the face also was given between each exercise.

2.4. *Video Processing*

The framework of the proposed regional evaluation method is presented in Figure 1. The details of the development will be explained throughout this section.

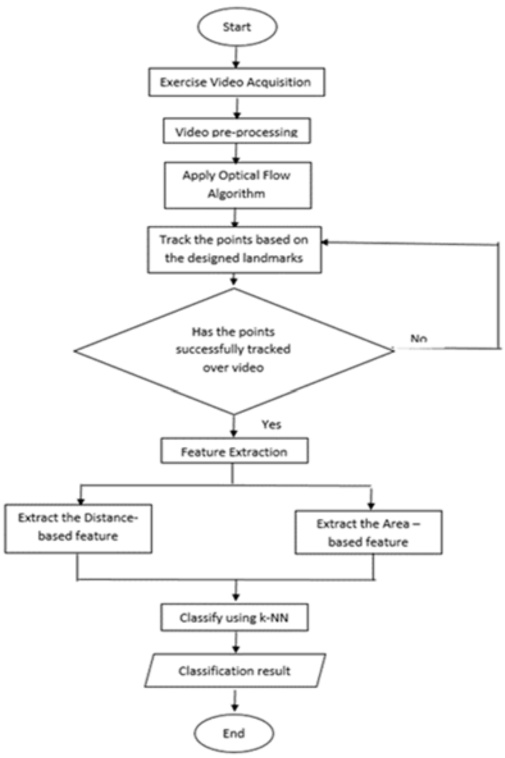


Fig. 1. The development of regional evaluation method

Kanade-Lucas-Tomasi (KLT) method[15] was chosen to track the subjects' face since no markers needed to be attached to the face and also because of its simplicity. This simple method made limited assumptions of the underlying image[16], but still remains as one of the best video-feature tracking algorithm.[17-19]

For the study, the House-Brackmann score was used as the reference to assess the facial nerve function. This score has analyzed three main parts of face that is forehead, eyes and mouth. For having a full assessment of face based on these three regions, suitable facial exercises were developed based on recommendation from the clinicians and for each exercise, only selected landmarks was chosen to characterize the movement of facial muscles. Our previous work[20] has summarized these details.

The initial position of each landmark in the video clip was marked by manually clicking by using the cursor. Once the landmarks were assigned, the KLT algorithm will automatically tracks each point within the Cartesian plane. Then, the pixel coordinates of all the landmarks are acquired by with a custom-written and designed Matlab® application and are recorded for further analysis.

2.5. *Quantitative Evaluation of Facial Paralysis*

To evaluate the facial nerve function quantitatively, initial hypotheses have been made for the study. As the study only investigated on Bell's palsy which only affected the patient on one side of the face, the results will be different with the normal subject who is having normal facial nerve function. For the forehead region, the patient will have difficulty lifting the eyebrows even in mild cases of paralysis. Furthermore, the eyes region also affected where the patient are not able to close the eyelids tightly. In addition, maximum or symmetrical smile cannot be executed in mouth region. Hence, based on these hypotheses, two parameters, Euclidean distance and Eclipse area are appropriate to be examined in this study to find the differences between normal and patient using these formula respectively:

$$D = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{1}$$

$$A = \pi D_1 D_2 \tag{2}$$

By referring to the Eq. (1) and Eq. (2), the x and y are the coordinate of the pixels for each landmark specified. Both distance and area are computed using these values.

Although the position of initial landmarks have been specified for each exercise, however, the face of each subject vary in size and surface area. Therefore, the coordinates of landmark also will be different. Because of that, the difference values between the initial and maximum movement for each parameter was calculated using the Eq. (3) and Eq. (4) before finding the ratio for the differences.

$$D_{diff} = D_{max} - D_0 \tag{3}$$

$$A_{diff} = A_{max} - A_0 \tag{4}$$

The values of the percentage ratio are used to identify the paralysis side of the face and then distinguish between patient and normal subjects. In order to devise a quantitative tool for the assessment of facial nerve function, an individual score was assigned for both sides of the face based on these percentage values. This is done based on the individual score chart presented in Table 1. This score chart was developed based on the results obtained after repeated trials using normal and patient subject and an inference was made with a numerous discussion with a medical professional.

Table 1. Individual Score Chart

Percentage of Ratio, %Δ	**Score Assigned**
< 1	0
1 – 5	1
6 – 10	2
11 – 15	3
16 – 20	4
21 – 25	5
26 and above	6

Subsequently, after all the three facial exercises have been conducted, a total score can be obtained by adding all the individual scores for each region of face. This total score demonstrates the severity level of the paralysis based on House-Brackmann (HB) grade that have been used widely in clinical application. However, in this work, the study will discuss only on individual score because this score lead to the evaluation of each region of face.

For the classification, *k*- nearest neighbor (*k*-NN) classifier was used to classify the subjects into normal and patients with different level of paralysis. It will classify the objects based on the closest training examples in the feature space. For the input to the classifier, the percentage value of the difference in distance and area will be taken.

3. Results and Discussions

For the classification, the parameter *k* of *k*-NN classifier is varies between the value of 1 to 5. Figure 2 and Figure 3 have shown the classification performance of individual score for each region of distance and area respectively.

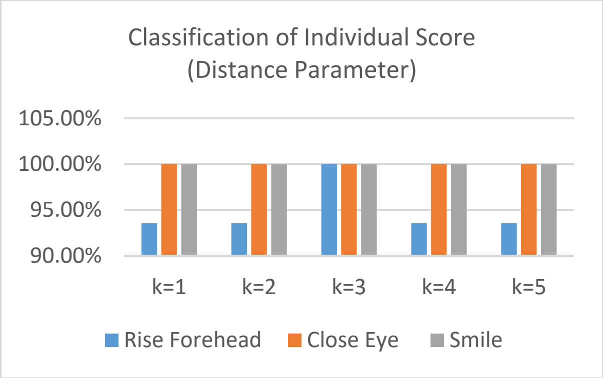


Fig. 2. Performance in Distance Parameter

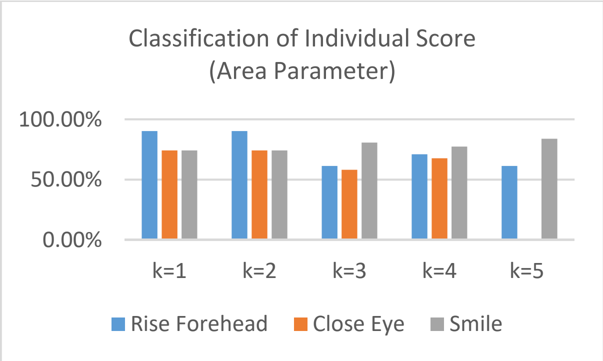


Fig. 3. Performance of Area Parameter

From the figures, it can be seen that the distance parameter is better in performance to classify the region of face between normal and patients subjects. Besides, by referring to Figure 2, it can be seen that by using $k = 3$ produce the best result with the 100% accuracy for all the facial regions. However, the other values of *k* also shows promising results in classification the subjects based on the individual score. Even the distance is better compared to area parameter, however, area performance also is good for the value k=1 and k=2.

Furthermore, for the distance parameter, the eyes and mouth region shows better performance compared to the forehead region. However, in area parameter, the forehead performance is over than the other two regions.

Even though the results show high accuracy in both parameters, however, a lot of works are still have to be done to make the system fully accepted globally. Increasing the number of patients may enhance the accuracy of the system.

Acknowledgements

The authors would like to thank the Medical Research Committee (MREC) of Malaysia for the approval of the data collection (Ref. No.: NMRR-12-1195-14375). Thanks also goes to the Ministry of Higher Education of Malaysia (MoHE) for the financial support.

References

1. M. May and B.M. Schaitkin, Reporting recovery of facial function, *The facial nerve,* 2nd edn. (Thieme Medical Publisher, New York, 2000).
2. P. Singhi and V. Jain, Bell's palsy in children, *Seminars in Pediatric Neurology* **10** (2003), 289-297.
3. J.G. Neely, K.X. Wang, C.A. Shapland, A. Sehizadeh, A. Wang, Computerized objective measurement of facial motion: normal variation and test-retest reliability, *Otol Neurotol* **9** (2010), 1488-92.
4. J.S. Brach and J.M. VanSwearingen, Physical Therapy for Facial Paralysis: A Tailored Treatment Approach, *Physical Therapy* **79** (1999), 397-404.
5. JR. JE. Benecke, Facial paralysis, *Otolaryngol Clin North Am* **2** (2002), 357-65.
6. C.G. Jackson, P.G. Doersten, The facial nerve. Current trends in diagnosis: treatment, and rehabilitation, *Med Clin North Am* **1** (1999), 179-95.
7. US Census Bureau. International Database [Online], 2012, available at : https://www.census.gov/population/international/data/idb/informationGateway.php
8. W.S.W Samsudin, K. Sundaraj, Clinical and non-clinical initial assessment of facial nerve paralysis: A qualitative review, *Biocybernetics and Biomedical Engineering* **34** (2014), 71-78.
9. J.W.M. Botman and L.B.W. Jongkees, Physical Therapy for Facial Paralysis: A Tailored Treatment Approach, *Physical Therapy* 79 (1999), 397-404.
10. M. Fradet, P. Robert, and P. Perez, Clustering Point Trajectories with Various Life-Spans, in: *Conference for Visual Media Production, 2009. CVMP '09*, 2009, pp. 7-14.
11. Y. Satoh, J. Kanzaki, and S. Yoshihara, A comparison and conversion table of 'the House–Brackmann facial nerve grading system' and 'the Yanagihara grading system', *Auris Nasus Larynx* 27 (2000)., 207-212.
12. W.S.W. Samsudin and K. Sundaraj, Evaluation and Grading Systems of Facial Paralysis for Facial Rehabilitation, *Journal of Physical Therapy Science* 25 (2013), 515-519.
13. H. JW and B. DE, Facial nerve grading system, *Otolaryngology Head Neck Surgery* 93 (1985), 146-147.
14. L. Dunwald, Using two- and three-dimensional kinematic analysis to compare functional outcomes in patients who have undergone facial reanimation surgery. [M.S. thesis] Edmonton, Alberta: *Speech-Language Pathology and Audiology*, Alberta University, 2010.
15. B.D. Lucas and T. Kanade, An iterative image registration technique with an application to stereo vision, in: *Proceedings of the 7th international joint conference on Artificial intelligence - Volume 2*, Morgan Kaufmann Publishers Inc., Vancouver, BC, Canada, 1981, pp. 674-679.
16. C. Tomasi and T. Kanade, Detection and Tracking of Point Features, in, Carnegie Mellon University, Pennsylvania, 1991, pp. 91–132.
17. H. Min, S. Ali, and M. Shah, Detecting global motion patterns in complex videos, in: *19th International Conference on Pattern Recognition*, 2008, pp. 1-5.
18. S. Ali, A Lagrangian Particle Dynamics Approach for Crowd Flow Segmentation and Stability Analysis, in: *Computer Vision and Pattern Recognition (CVPR 2007)*, 2007, pp. 1-6.
19. M. Fradet, P. Robert, and P. Perez, Clustering Point Trajectories with Various Life-Spans, in: *Conference for Visual Media Production, 2009. CVMP '09*, 2009, pp. 7-14.
20. W. S. W. Samsudin, K. Sundaraj, A. Ahmad, and H. Salleh, "Initial assessment of facial nerve paralysis based on motion analysis using an optical flow method," *Technology and Health Care,* vol. 24, no. 2, pp. 287-294, 2016.

Introduction of Forehead Lesion Assessment with House-Brackmann Score for Facial Nerve Paralysis Evaluation

Wan Syahirah W Samsudin[1*]
Faculty of Electrical & Electronics Technology Engineering,
Universiti Malaysia Pahang, 26600 Pekan, Pahang, Malaysia.

Rosdiyana Samad
Faculty of Electrical & Electronics Technology Engineering,
Universiti Malaysia Pahang, 26600 Pekan, Pahang, Malaysia.

Kenneth Sundaraj
Faculty of Electronics & Computer Engineering,
Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia.

Mohamed Rizon
[3]Faculty of Engineering, Technology and Built Environment,
UCSI University, No 1, Jalan Menara Gading, Taman Connaught, 56000 Kuala Lumpur, Malaysia.

Mohd Zaki Ahmad
Department of Otorhinolaryngology,
Hospital Tuanku Ampuan Afzan, Jalan Tanah Putih, 25100 Kuantan, Pahang, Malaysia.
E-mail: wansyahirahwsamsudin@yahoo.com, rosdiyana@ump.edu.my, kenneth@utem.edu.my,
MohdRizon@ucsiuniversity.edu.my, drmohdzaki@yahoo.com.my

Abstract

This paper present a novel approach of forehead lesion assessment using Gabor filter method to identify the type of facial nerve paralysis and proposed a forehead-HB score table to assist the clinicians in assessing the facial nerve function quantitatively, quicker and in a convenient way. A 100% of accuracy has been obtained in identifying the two type of facial lesion, Upper Motor Neuron (UMN) and Lower Motor Neuron (LMN) lesion. Increasing more data may enhance the performance of the accuracy

Keywords: Facial Nerve Paralysis, Gabor Filter, Upper Motor Neuron (UMN) lesion, Lower Motor Neuron (LMN) lesion , House-Brackmann (HB)

1. Introduction

Facial nerve paralysis is a loss of the voluntary muscle movement due to temporary or permanent injury to any of facial nerves. Since the pathway of the facial nerve is in a long distance, starts from the motor cortex, hence various causes may resulting in this facial paralysis. The most reported cases are Bells' Palsy which is 50 % of the cases. Bells' Palsy is an idiopathic disease which named after Sir Charles Bell (1774-1842).[1-3]The patient with the

*Wan Syahirah W Samsudin, Faculty of Electrical & Electronics Technology Engineering, Universiti Malaysia Pahang, 26600 Pekan, Pahang.

paralysis will have difficulties in smiling, puckering, closing eyes tightly, whistling and speech problems. Besides, the taste sensation also affected.[4] Because of these reason, the patient will suffer from serious functional, cosmetic and also physiological problems. Even it is not a fatal disease, the facial paralysis might cause extensive destruction in patients' life because the face is a symbol of beauty, attractiveness and interest with each other.[5]

Facial paralysis can be categorized into two types, central facial paralysis and peripheral facial paralysis. Central facial paralysis is known as upper motor neuron lesion (UMN) because of resulting from the lesion of the central neuron in the brainstem and often caused by a stroke. The forehead are not distorted in UMN due to the bilateral innervation of the facial nuclei supplied to the forehead[4] but the lower half of the face is affected. In other hands, peripheral facial paralysis or mostly known as Lower Motor Neuron (LMN) lesion is caused by the lesion on the peripheral neuron where it affected the upper, middle and also the lower part of the face.[6] The most common cause of acute LMN is Bells' Palsy.

In the facial paralysis cases, the localization of the facial lesion is crucial to determine specific clinical features[4] and it is the most essential procedure for suggesting the further treatment and rehabilitation to the patient.[7] The localization of the lesion whether it is in the brainstem or cerebral cortex can be determined as soon as the upper motor neuron lesion is suspected.[8] The cortical lesion will affects the contralateral limbs and promote the involuntary movements of the face. At this stage, urgent referral is needed to the secondary care (neurology or acute medical unit).

Even though these type of lesions are the main concern to the facial nerve evaluation, to date, only a single work has been proposed. Barbosa et al. [9-10] have introduced an approach on iris segmentation to classify the type of facial paralysis. Based on this work as a benchmark, the study has found the prominent features to differentiate the lesion which is the forehead area other than nasolabial fold and asymmetry of the mouth.[8,11] Most surprisingly, even though it is critical, none of the published work have quantitatively analyzed the forehead region in determining the UMN and LMN lesion. For LMN cases, the House-Brackmann scale[12] is the most used in clinical application to evaluate the severity of facial nerve paralysis but not the lesion.

Based on the above discussion, the purpose of this paper is to propose a classification method of facial lesion whether it is Upper Motor Neuron (UMN) lesion or Lower Motor Neuron (LMN) lesion of facial nerve paralysis based on the forehead assessment and provide the House –Brackmann (HB) scale for the assessment. The Gabor filter method with different scales and orientations was used and investigated by experiments to attain potential parameters to be used in the final implementation of the system. This method is succeeded in classifying the lesion into UMN and LMN and it hopes that this study will encourage more research to be done in future in assisting the clinicians in the assessment of facial nerve functions.

2. Materials and Methods

2.1. *Ethical Statement*

For this study, all the procedures conducted were approved by the Medical Research and Ethics Committee (MREC), Ministry of Health, Malaysia (Ref No.: KKM/NIHSEC/800-2/2/2/P13-96). As involving the human subjects, the study need to be conducted in accordance to Declaration of Helsinki. The consent were obtained from each subjects prior the experiment.

2.2. *Experimental Setup*

At a distance of 1.0 meter, a handy camera (SONY DCR-SX45E) with a 640 x 480 resolution, was placed in front of the subjects to capture the frontal view of the subjects' face with a rate of 30 frames per second. Additional lighting system was developed in order to make the quality of the image capturing were adequate to be processed. The data was collected at the Physiotherapy and Rehabilitation Department of Hospital Tuanku Ampuan Afzan (HTAA), Kuantan, Malaysia.

2.3. *Data Acquisition and Image Pre-Processing*

The data that satisfy the study was not found and not accessible anywhere. Therefore, a database was developed with 62 data of normal subjects and 40 data of patients. The patients for LMN lesion in this study presented with the same etiology which is Bells' Palsy and the data used for the UMN lesion is from the normal subject that exhibit the same criteria as UMN patient where the forehead is spared and able to lift the eyebrows symmetrically due to the limited sources of the database. The data acquisition was performed through facial image capturing. The data consists of two types of facial images; when the subject is at rest and the other is at maximal condition where the subject raise the forehead. All the facial images are then cropped at the forehead region to be further processed.

2.4. *Computation of Gabor Filter and Features Extraction*

The framework of the proposed lesion assessment method is presented in Figure 1. The development will be explained in details throughout this section.

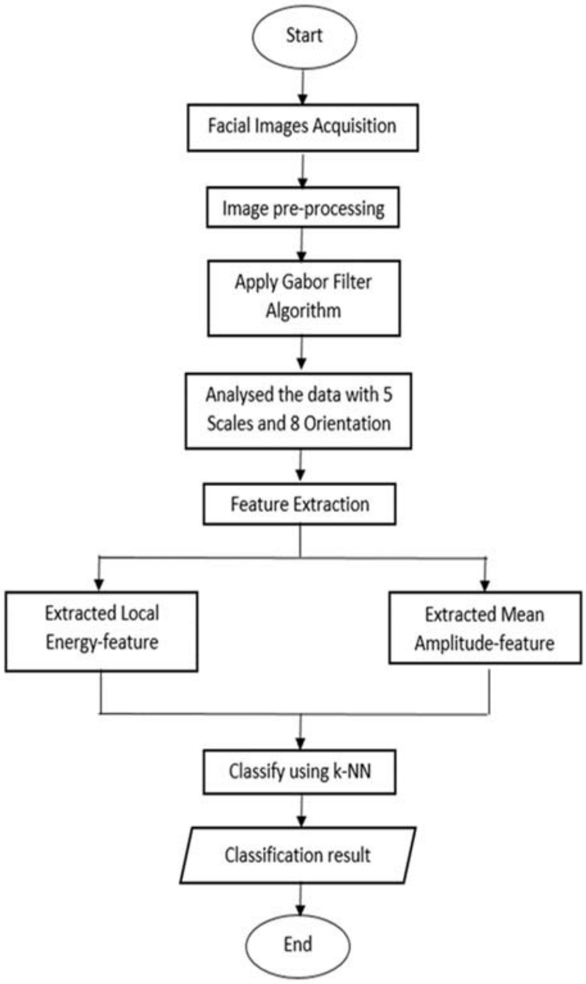


Fig. 1. The development of forehead lesion assessment

The Gabor filter was named after a physicist, Dennis Gabor and has been used enormously in many applications[13], such as for face recognition, edge detection, texture segmentation and classification and image analysis. Various salient properties can be manipulated such as spatial localization, orientation selectivity and spatial frequency characteristics from this filter. The wavelets from the Gabor filter is being acknowledged as one of the most favorite one for the face representation[14.] Hence, the excellent performance is considered for the study and the Gabor feature was chosen to exemplify the wrinkles of forehead during raising eyebrows movement. For this study, only five scales with eight orientation were selected in order to reduce the complexity and computation costs. The best scales and orientation will be selected at the end of the study to be implemented in the final system.

Gabor filter consists of a real and an imaginary component, which represents the orthogonal directions. The Gabor features are then computed by using the convolution operation of Gabor filter bank ψ (x, y) with the input image *I (x, y)*.

As a result, the Gabor response can be computed, and two main features have been extracted, Local Energy (LE) and Mean Amplitude (MA). Both features were extracted for both sides of face to find the differences between them. The computation of percentage difference was initiated based on hypothesis made prior the experiment. For the patient subject (LMN lesion), the difference should be higher compared to the normal subject (UMN lesion) because of the eyebrows cannot be lifted symmetrically. The flow to get the percentage difference is shown in Figure 2.

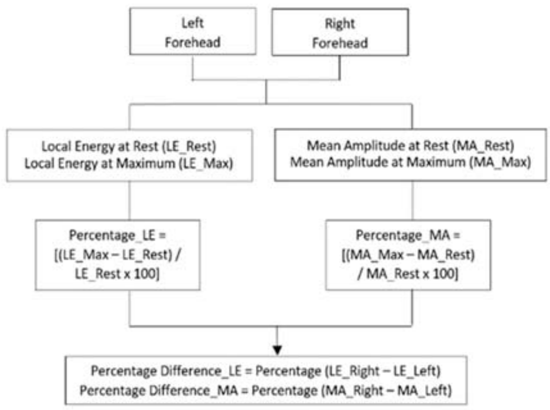


Fig. 2. Flow of Percentage Difference Computation[15]

K-nearest neighbor (*k*-NN) classifier is used in the study to classify the subjects into Upper Motor Neuron (UMN) and Lower Motor Neuron (LMN). The percentage difference value as discussed earlier will be the input features to the classifier. The *k*-NN is classifying the objects based on the closest training examples in the feature space. Many previous works have proved it is the simplest algorithm all other machine learning algorithm a possible to classify various types of image data.[16-17]

3. Results and Discussions

The values of *k* of *k*-NN classifier is varied from 1 to 5 to find the best results in classifying the UMN and LMN lesion. As our previous research[15], the Scale 2 and Orientation 5 of Gabor filter was chosen because of getting the best accuracy compared to the other scales and orientations after the numerous experiments have been done.

Based on Figure 3, 100% of percentage accuracy was obtained in classifying the lesion whether it is UMN or LMN using $k = 1$ till $k = 4$ for Local Energy feature and 96.7% of accuracy acquired by applying the Mean Amplitude feature as the input to the *k* –NN classifier. From these results, the study has shown a great potential to aid the clinicians in determining the lesion and suggesting the suitable treatment to patient.

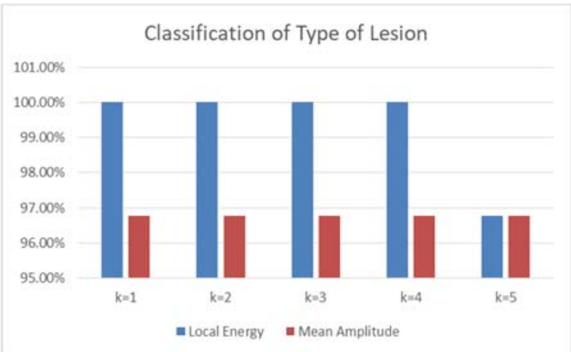


Fig. 3. Classification performance of Forehead Lesion Assessment

As for clinicians, the using of House-Brackmann scale is very convenient because it has been used for decades in clinical application. Therefore, this study has proposed a forehead lesion score which also gives out the House-Brackmann score to assist the clinicians in referring the severity level of paralysis and lesion at the same time as presented in Table 1. This Forehead-HB Score is developed based on our evaluation on numerous experiments.

Table 1. Forehead Lesion – HB Score Table

Difference Percentage	*Type of Lesion*	*House-Brackmann Score*	*Characteristic of Forehead*
0-5	UMN	I	Normal
6-8	LMN	II	Mild asymmetry
9-11	LMN	III	Slight movement
12-14	LMN	IV	No movement
15-17	LMN	V	No movement
> 17	LMN	VI	No movement

The Level I is only for Upper Motor Neuron (LMN) lesion since the forehead exhibit the same characteristic with the normal facial function. Level II till Level VI is for Lower Motor Neuron (LMN) lesion which is Bell's palsy cases. The table is presented with the hope in assisting the clinicians to identify the type of facial paralysis in the initial assessment. Nevertheless, a lot of works are still remain to be conducted in the research area. The future works will be the classification process of the level of lesion and tested them with many patients. Besides, real patient of UMN should be involved in the experiment rather than using normal subjects. Hence, the number of patients should be increased by collaborating with the other research sites and hospitals to exchange many more information of patients and these may enhance the accuracy of the system.

Acknowledgements

The authors would like to thank the Medical Research Committee (MREC) of Malaysia for the approval of the data collection (Ref. No.: NMRR-12-1195-14375). Thanks also goes to the Ministry of Higher Education of Malaysia (MoHE) for the financial support.

References

1. J.S. Brach and J.M. VanSwearingen, Physical Therapy for Facial Paralysis: A Tailored Treatment Approach, *Physical Therapy* **79** (1999), 397-404.
2. JR. JE. Benecke, Facial paralysis, *Otolaryngol Clin North Am* **2** (2002), 357-65.
3. C.G. Jackson, P.G. Doersten, The facial nerve. Current trends in diagnosis: treatment, and rehabilitation, *Med Clin North Am* **1** (1999), 179-95.
4. C. Scully, "20- Sensory and motor changes", *Oral and Maxillofacial Medicin*e, 3rd edn. (Churchill Livingstone, 2013), pp. 145-153.\
5. J.G. Neely, K.X. Wang, C.A. Shapland, A. Sehizadeh, A. Wang, Computerized objective measurement of facial motion: normal variation and test-retest reliability, *Otol Neurotol* **9** (2010), 1488-92.
6. J. Barbosa *et al.*, Efficient quantitative assessment of facial paralysis using iris segmentation and active-contour-based key points detection with hybrid classifier, *BMC Medical Imaging,* vol. 16 (2016), pp. 23.
7. J. Finsterer, Management of peripheral facial nerve palsy, *European Archives of Oto-Rhino-Laryngology*. **265**(7) (2008) 743-752.
8. L. Masterson, M. Vallis, R. Quinlivan, and P. Prinsley, Assessment and management of facial nerve palsy, *BMJ, British Medical Journal,* 351, 2015.
9. J. Barbosa, W.-K. Seo, and J. Kang, *"paraFaceTest: an ensemble of regression tree-based facial features extraction for efficient facial paralysis classification," BMC medical imaging***,** **19**(1) (2019), 30.
10. J. Barbosa et al., *"Efficient quantitative assessment of facial paralysis using iris segmentation and active contour-based key points detection with hybrid classifier," BMC Medical Imaging,* 16 (2016), p. 23.
11. H.S. Hira, Clinical Examination: a practical guide in medicine (Jaypee Brothers Medical Publisher, New Delhi, 2011).
12. J.W. House and D.E. Brackmann, Facial Nerve Grading System, *Otolaryngology-Head and Neck Surgery*, 93(2) (1985), pp.146-147.
13. J.K. Kamarainen, Gabor features in image analysis, *Image Processing Theory, Tools and Applications (IPTA),* 2012 3rd International Conference (2012), pp. 13-14.
14. W. Zhao, J. Park, and S. Lee, Fully Automatic face detection and facial feature points extraction using Local Gabor filter bank and PCA, *in 2011 International Conference on Machine Learning and Cybernetics,* 4 (2011), pp. 1789-1792.
15. W.S.W. Samsudin, R. Samad, M.Z. Ahmad, K. Sundaraj, Forehead Lesion Score for Facial Nerve Paralysis Evaluation, *in 2019 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS),* 2019, pp. 102-107.
16. P. Heroux, S. Diana, A. Ribert, and E. Trupin, "Classification method study for automatic form class identification," in *Proceedings. Fourteenth International Conference on Pattern Recognition (Cat. No. 98EX170)*, 1998, vol. 1, pp. 926-928: IEEE.
17. I. Ismail, M. Ramadan, T. El Danf, and A. Samak, "Multiresolution Fourier-wavelet descriptor for fingerprint recognition," in *2008 International Conference on Computer Science and Information Technology*, 2008, pp. 951-955: IEEE.

Intelligent Wearable Biofeedback Fuzzy Logic Based Device for Monitoring and Treatment of Voice Loudness

Ali S. AlMejrad
Department of Biomedical Technology, King Saud University
P.O. Box 10219, Riyadh 11433, Kingdom of Saudi Arabia
(amejrad@ksu.edu.sa)

Abstract

Development of medical devices are expanding rapidly. This is occurring as results of development of high tech electronics, information technology and wireless communication technology. Integrating these technologies helps to provide good, less cost and more effective healthcare services. In this paper a novel intelligent medical biofeedback device is developed for accurate monitoring of the voice loudness using fuzzy logic compared to proportional control. This accurate monitoring device then enhanced treatment of people who are unable to control excessive or low loudness due intellectual disability or aggressive behavior causing loudness levels rejected by others. These people with this behavior may often be unaware of their problems and hence an intelligent wearable biofeedback device is necessary to be used by the patients outside the clinic for self-controlling excessive and low loudness without the need to the continuous therapeutic follow-up in speech pathology clinic. The use of the intelligent biofeedback device proved to be an effective due to its enhanced features. The technical design considerations, enhanced features of the device and evaluation will be presented.

Keywords: Biofeedback Device, Fuzzy Logic, Voice Loudness.

1. Introduction

There are many voice disorders such as vocal cord nodules, vocal loading and many other speech disorders[1]. These disorders are caused by excessive loudnes or speaking loudly due to aggressive behavior or intellectual disability[2]. This vocal abuse may result in vocal cord pathologies such as vocal nodules. This pathology occurs in high percentage of children with voice disorders[3]. On the other side inadequate volume, weak or soft voice due to reduced vocal intensity are classified as voice disorders such as in Parkinsonian patients who are unable to regulate the volume of their speech [4].

Usually the patients with the above behaviors may often be unaware of their problems. So intervention is required to limit these negative outcomes. This intervention using simple verbal prompting by therapist can help patients if instructed to regulate the speaking volume by controlling their loudness but unfortunately such improvement generally is not long-term one. So, voice monitoring is necessory in addition to the use of biofeedback training. The use of biofeedback control instead is more effective compared to the simple verbal control. Feedback using microcomputer based system and some hardware one kilogram weight was used for speech therapy with Parkinson's disease patients as in Ref.5. Schliesser used the biofeedback training employing electromyography to voice disorders[6]. Another limited functions system using biofeedback analog device is developed by McGillivray et al. to reduce loudness for treating a child with soft nodules on her vocal cords via interrupting child conversation by the therapist with a loud tone when her speech exceeded threshold intensity level[7]. Lancioni as in Refs 8 used a portable auditory-feedback and vibratory-feedback devices for reducing excessive loudness in a deaf woman. with mental retardation.

To limit these negative outcomes and enhance the features of the feedback device, a novel intelligent biofeedback device is developed for accurate monitoring and feedback of the voice loudness using fuzzy logic compared to proportional control[9]. This accurate monitoring and feedback with variable alarrmdevice will enhance treatment of people who are unable to control excessive or low loudness accurately. This device is portable and can be used at home without the need to the continuous therapeutic follow-up in speech pathology clinic.

2. Materials and Methods

2.1 Device design hardware

The intelligent Wearable Biofeedback Fuzzy Logic Based Device for Monitoring and Treatment of Voice Loudness in this work consists of many components including: throat microphone, signal conditioning including amplifiers, filters, rectifier, unity gain buffer, in Psoc microcontroller; auditory alarm feedback. The device is battery operated as shown in Figure 1.

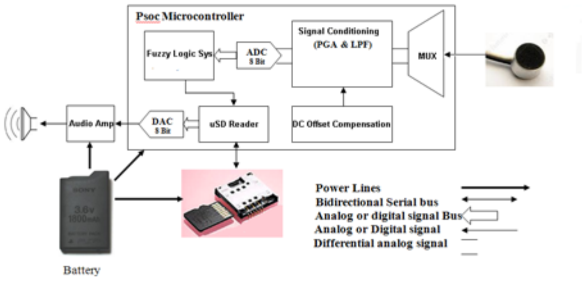


Fig. 1. Block diagram of Wearable Biofeedback Fuzzy Logic Based Device for Monitoring and Treatment of Voice Loudness

The signal conditioning as shown in Figure 2. consists of preamplifier, band pass filter to eliminate unwanted noise, main Amplifier with an adjustable gain to make the signal bigger, rectifier to turn negative signal values to positive and integrator to evaluate voice intensity over a set time period. A 500 Hz corner frequency low pass filter is considered because this frequency is low enough to include the volume or intensity of the voice. It cuts off the high frequencies associated with pitch, quality, or uniqueness of a particular speech. The gain of the amplifier is 100 since the typical values expected are in the order of 10 mV[10].

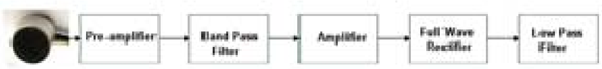


Fig. 2. Block diagram of the signal conditioning of voice signal

After conditioning the voice signal using analog circuitary in the Psoc microcontroller, it passes through 8-bit A/D converter to digitize the voice signal for intelligent processing using fuzzy logic system (FLS). The feedback is consisting of an audio alarm with duty cycle determined by FLS depends on the low or high loudness levels. The purpose of the feedback device is to warn the patient when his or her speech intensity had gone below or above the specified threshold as input to the fuzzy logic system based on therapist suggestion. With auditory feedback, the patient would be notified that vocal changes either decreasing or increasing voice level as regulation of voice is needed to be made. Finally, the fuzzy logic in Psoc microcontroller is programmed to process the voice data acquired and to determine when feedback to the patient is necessary as will be discussed in the following section.

2.2. Design of loudness fuzzy logic based system

Using graphical programming method, fuzzy logic based system can be designed with qualitative description for the system and it avoids the tedious method of developing tuning parameters for the system via emulating the human linguistic expressions using IF-Then rules[11].

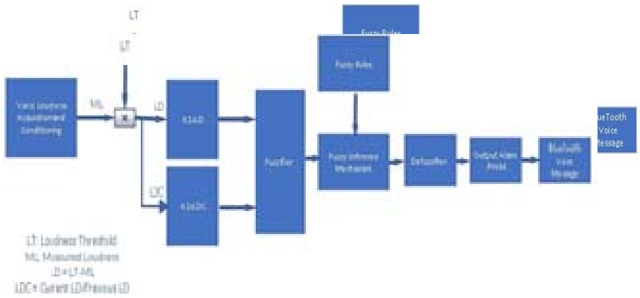


Fig. 3. The loudness biofeedback fuzzy logic based monitoring system

First, fuzzy based system was designed using a graphical user interface with the fuzzy rules stored in the fuzzy system editor as shown in Figure 4.

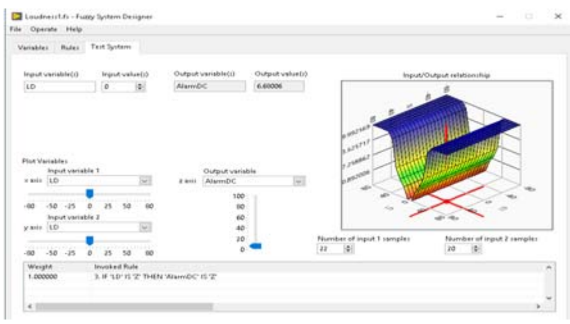


Fig. 4. Fuzzy System Editor

The designed fuzzy logic system is a Multiple input and Single output system (MISO) where the difference between the measured intensity (LD) and threshold intensity level and change of intensity to normal level (LDC). The output variable is alarm of up-normal cases calculated as duty cycle (AlarmDC). Our objective is to check that intelligent system behaves better than conventional system. The design of fuzzy logic system was done to control the loudness level. The fuzzy system is designed with two input variables and one output variable. For the classical fuzzy logic system, the input variables are LD and LDC and the output variable is AlarmDC. To control loudness level Mamdani's method is used to create the linguistic rules which is considered easy to interpret by human operator. For assigning the values to membership functions, triangular and trapezoid shaped membership functions are used for input variables and the output variable. The universe of discourse of intensity difference, change of intensity and alarm duty cycle are (-80 80), (-100 100) and (0 100%) respectively.

The membership functions for LD and LDC inputs consist of negative large (NL), negative small (NS), zero (Z), positive large (PL) and positive small (PS) as shown in Figures 5 and 6.

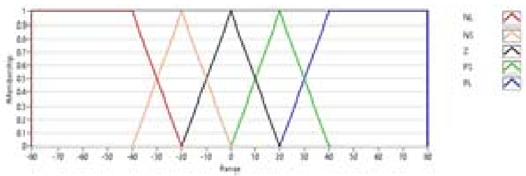


Fig. 5. The membership functions for LD input

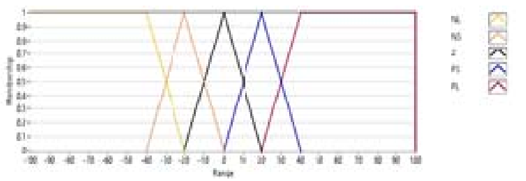


Fig. 6. The membership functions for LDC input

The output AlarmDC variable value was chosen in the range of 0-100% with variable frequency output based on percentage of high or low loudness. The membership functions for AlarmDC output as zero (Z), small (S), medium (M) and large (L) as shown in Figure 7.

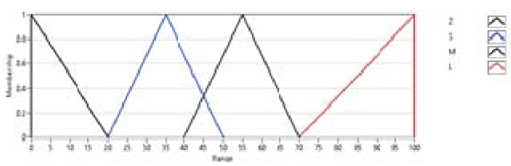


Fig. 6. The membership functions for LDC input

In the design of the fuzzy inference system was formulated which does the mapping from the given input to an output using fuzzy logic as shown in Figure 8. For evaluating the linguistic rules of the fuzzy logic Max-Min inference system has been used. In developed FLS, the centroid defuzzification method has been used.

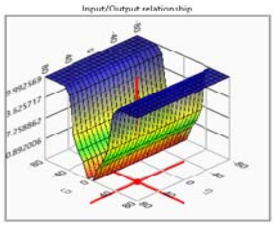


Fig. 8. Input/Output Relationship of FLS

3. Simulation and Validation

To simulate and validate the fuzzy logic system, a user interface virtual instrument (VI) is designed using graphical language as shown in Figure.9 which shows the front panel of VI in testing mode. In the simulation loop there exists one input node and one output node to configure parameters interactively to determine the alarm output level represented as duty cycle based on the loudness level..

Fig. 9. Front Panel of the Loudness Monitoring Controller

The VI has a simulated voice level as an input and threshold level preset by a speech therapist and set prior to use, based on the patient's requirements. These two inputs are used as inputs to fuzzy logic system. The softer or louder the patient speaks, the lower or higher the intensity measured by FLS. The FLS then utilizes the specified rules to process the input data and decides whether or not the led alarm and audio feedback should be sent to the patient in addition to the length of the feedback alarm based on the up normal level.

First, when the start button is pushed, the code will determine the threshold level above the background noise. This alarm threshold is the loudness level the patient needs to speak above or below.

After that, it will start processing using fuzzy logic rules. If the simulated voice signal level received is below or above the speaking threshold, the output audio alarm is activated with message instruction. During times in which the user is not speaking, the counter is held constant, while speaking below the threshold, the counter is tracking the number of both events low and high alarms in addition to the level of loudness. The test continues by taking another simulated input sample to define the new level and then determine whether to set the alarm or not. This loop repeats until the user stops the test and exits.

4. Results Analysis and Discussion

A complete simulation for evaluation was performed using the initial designed, simulated and developed and assembled using a breadboard to evaluate the connections to each unit of the system. Each stage of the intelligent system was implemented, tested and calibrated separately. The processing stage using the programmable microcontroller was simulated firstly using different levels of speech intensity of speech-like signals. Secondly it was tested by different normal subjects who were asked to speak for few minutes using different levels of speech intensity and test to speak softly or loudly a few seconds and check if the biofeedback activates the audio alarm. To make accurate tests a sound level meter for voice monitoring was used to monitor the volume of voice loudness to adjust the biofeedback to be activated when low or excessive speech intensity. After calibrating the device, the device was tested for different levels of voice intensity which proves for the different tests that the system was successful.

5. Conclusion

The Intelligent Biofeedback Fuzzy Logic Based Device for Monitoring and Treatment of Voice Loudness developed as a prototype achieved the objective of this work such as the accuracy, ease of programmability, portability due to its small size, patient alert, automatic tracking of up normal occasions during therapeutic sessions in addition to low power consumption. Our intelligent voice monitoring system performed satisfactorily as a prototype system for monitoring, to provide a patient with information about voice loudness level accurately outside the clinic and for controlling of the treatment of upnormalvocal loudness levels without the need to the continuous therapeutic follow-up in speech pathology clinic. The device is able to collect voice loudness levels through a throat microphone and able to discriminate between voice and silence and is only active when the subject is speaking. The device activities an audio alarm as a biofeedback to the subject whenever the voice intensity level falls below or exceeds above adjustable single threshold for a preset time interval. The device has uSD external memory to store information concerning time dependent data of the user's voice intensity when alarms occurred in and counts of up normal events. These data can be displayed in real-time and offline by a speech therapist using the stored data to monitor the patient's progress and make changes to the therapy program.

Clinical evaluations of the device is required to be done for testing its feasibility as a monitoring and therapeutic device for voice disorders.

Acknowledgement

I would like to extend my appreciation to the Deanship of Scientific Research and the Research Center of College of Applied Medical Sciences at King Saud University for funding this research.

References

1. M. Greene, and M. Lesley M, *the* Voice and its Disorders, John Wiley & Sons; 6th Edition, 2001.
2. G. E. Lancioni et al, Reducing excessive vocal loudness in persons with mental retardation through the use of a portable auditory-feedback device, J Behav Ther Exp Psychiatry,1997,28(2):123-8.
3. C. R. Hooper, Treatment of voice disorders in children, Lang Speech Hear Serv Sch.,2004, 35(4):320-6.
4. B. J. de Swart et. al., Improvement of voicing in patients with Parkinson's disease by speech therapy, Neurology,2003,60(3):498-500.
5. J. E. Zicker et al., A portable microprocessor-based biofeedback training device, IEEE Trans Biomed Eng.,1980,27(9):509-15.
6. H. Schliesser, EMG biofeedback as a function of voice-quality change, J. Percept. Motor Skills,1987, 64, pp. 719-724.
7. R. McGillivray et al., Simple biofeedback device to reduce excessive vocal intensity, Med Biol Eng Comput,1994, 32(3):348-50.
8. G.E. Lancioni and S. Markus, a deaf woman learning to control her excessive vocal loudness through a portable feedback system, Percept Mot Skills,1999,88(3 Pt 2):1347-9.
9. A. AlMejrad, Design of intelligent system for speech monitoring and treatment of low and excessive vocal intensity, Journal of Artificial Life and Robotics, 2010, Vol.15, pp.320-324.
10. J. Webster, J., Medical Instrumentation Application and Design, 3rd Edition, Wiley,1997.
11. L. A. Zadeh, Fuzzy Sets, Information and Control , 8,1965.

The Development and Evaluation of Fig Leaf Syrup

Shang-Hui Li, Pei-Hi Zheng, I-Chih Chiang, Yu-Ting Su, Syue-Sheng Lin, Chen-Zong Ho
Department of Food and Beverage Management, Far East University, Taiwan (R.O.C.)
E-mail: a0956358700@gmail.com

Abstract

The figs sold in the market are only fruits and their processed products, the use of fig's leaves are relatively rare. But, the fig's leaves contain furanocoumarins, flavonoids, pectins, resins, sugar and vitamin C (Meng et al., 1996). Perhaps the special aroma of its leaves, or perhaps the astringency of the leaves, is less acceptable to consumers. This study used fig's leaves as the main ingredient with crystal sugar and different kinds of vanilla formula to developed if more consumers willing to accept fig's leaves syrup or not. Sensory evaluation was processed after the finished product was diluted. The obtained results was statistically analyzed by SPSS12.0, and analyzed by single factor variance and number allocation table analysis.110 students and teachers of Far East University, Tainan were selected for convenience sampling methods and 100 valid questionnaires were collected. The content of the evaluation was based on the appearance, aroma concentration, aroma preference, sweetness concentration, sweetness preference, astringency, aftertaste, overall acceptability, and willingness etc. And find out the most suitable formula of the fig's leaves syrup. In terms of appearance, the fig's leaf syrup with other vanilla is more acceptable than the original fig's leaf syrup. As for lemon fig's leaf syrup, although the aroma is the strongest, but the overall acceptance and purchase intention is the lowest due to its high astringency. Mint fig's leaf syrup, its aroma is accepted with the highest level. The results of the comprehensive data analysis found that the most suitable formula is fig's leaf 60gs, water 400gs, crystal sugar 200gs and mint 15gs.

Keywords: Fig's Leaf, Syrup, Sensory Evaluation, Purchase Intention

1. Research Motivation and Purpose

The figs can be seen as desserts such as fig cakes, fig fruit towers or dried fig fruits, but there are few products about fig leaves. Fig and its leaves are rich in selenium[1]. Selenoproteins has functions of anti-oxidation, regulates thyroid hormone metabolism and maintains vitamin C and other molecular reductions. It also affects the body's immune system and enhances the body's immunity [2]. [3] Pointed out that among the fig plants, the leaves contained the highest amount of selenium and were edible. This study used fig leaves to make natural syrup, hoping to retain ingredients that are beneficial to the human and make delicious syrup under using fig leaves with sugar and vanilla to make fig leaf syrup. The purpose is as follows: 1. Discuss the feasibility and acceptability of fig leaves for making drinks; 2. Explore consumers' willingness to purchase fig leaf syrup; 3. Identify the most acceptable fig leaf syrup formula for consumers.

2. Literature Review

The fig is native to Asia Minor and the Mediterranean coast introduced in Taiwan in 1915. Because it easily care for, early mature, and uneasy to disease, therefore, the farmers who planted in Taiwan are mostly based on Bojihong [4]. This study used the fig leaf of Bojihong as the material. This study used rock sugar as the ingredient to make syrup. Rock sugar is a crystal of sucrose after refining. Its characteristic is similar to those of white sugar[5]. The auxiliary of syrup are mint, lavender and lemon. Mint is native to Western Europe, Central Europe and the Mediterranean which adaptable to temperate and warm areas and is a perennial herb that spreads everywhere. It was first used in the Roman era and was an important culinary herb and medicinal plant [6] . Mint is perhaps the most well-known vanilla plant in the world. Lavender is found from Europe in northern and eastern Africa, the Mediterranean, southwest Asia and southeast India. It is a perennial shrub and enjoys a ventilated and open all-sun environment. The history of using lavender was very long. It is said that in the ancient Egyptian era, mummies were made using lavender oil. The origin of the lemon is in the western Himalayas. It belongs to the family of oranges and is a representative food rich in vitamin C [7].

The Institute of Food Technologists, IFT [8], in 1985, considered sensory evaluation as a scientific discipline that applies principles of experimental design and statistical analysis to the use of human senses, sight, smell, taste, touch and hearing for the purposes of evaluating consumer products [9].This study used a scoring test to explore consumer acceptance and willingness to purchase new flavored fig leaf syrup. Usually, consumer-based trials typically investigate new product developments and the quality of listed products, hoping to know if consumers like or accept them. Furthermore, the scoring test is to know the high or low level of the product by consumers.

The purchase intention refers to the consumer who is willing to purchase a certain product. The higher the perceived value of the product is, the greater the intention to purchase the product is [10] . In addition to measuring purchase intentions from multiple perspectives, some studies measured in a single facet only, such as Biswas and Blair [11], which directly measured purchase intentions with a single question.

3. Research Result

After preliminary tests, the results showed that the overall acceptance of fig leaf syrup with added rock sugar was higher. However, The average number did not reach 3. After discussion, it was decided to add three different vanilla plants with rock sugar fig leaf syrup for formal tests. The formal test was divided into four samples The formal test was conducted by convenient sampling. A total of 110 participants from Far East University were invited to attend the sensory evaluation. 5 points score system was taken, the 1 is the least favorite or the weakest, 5 is the most favorite or the strongest, and 100 valid questionnaires were collected. The collected valid questionnaires were statistically analyzed with SPSS 12.0, and the Duncan method was used to measure the differences between the samples using the one-factor variability.

4. Conclusion and Suggestion

The appearance of fig leaf syrup (Table 1), the Duncan analysis results showed significant differences among the three groups ($p < 0.001$), but all the participants could accept its appearance, the average number was greater than 3. In terms of aroma concentration, the participants considered that the lemon fig leaf syrup was the strongest (mean 3.60), and there was a significant difference among the three groups ($p<0.001$). In aroma preference, the mint fig leaf syrup had a higher aroma preference and was significantly different from the other three samples ($p<0.05$). In terms of sweetness concentration, the results showed that the lemon fig leaf syrup (mean 2.17) had the lowest sweetness, followed by the original fig leaf syrup (average 2.70), then, lavender fig leaf syrup (mean 3.02), and the three groups were significantly different ($p < 0.001$). Refer to sweet taste, lemon fig leaf syrup had a lower sweet taste and was significantly different from the other three samples ($p < 0.001$). In syrup astringent, the participants considered that the four syrups groups were not smashed, the lemon fig leaf syrup had a higher astringency (average 2.56) and was significantly different from the other three samples ($p < 0.001$). In the aftertaste, the participants considered that the lemon fig leaf syrup (mean 2.13) was significantly different from the other three samples ($p < 0.001$). As for overall acceptance, lemon fig leaf syrup (mean 2.12) was significantly different from the other three samples ($p < 0.001$). Lavender, plain and mint were barely acceptable to consumers. Compared with the purchase intention, the purchase for mint fig leaf syrup was the highest, and there was a significant difference among the three ($p<0.001$). The results found that the overall acceptance and purchase intention of the mint-added fig leaf syrup was the most favored by the participants. The most suitable formula was to add mint 15g, fig leaf 60g, water 400 g and rock sugar 200g of mint fig leaf syrup.

5. Conclusion and Suggestion

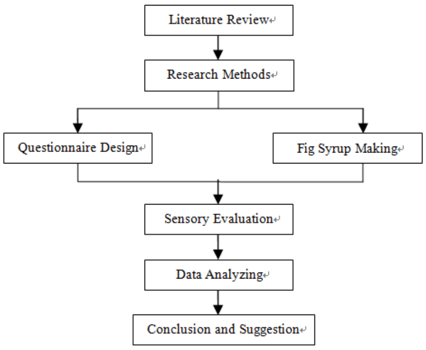


Figure 1 Research flow chart

This study was aimed at the development and evaluation of fig leaf syrup products. The conclusions are as follows: 1. In terms of appearance, fig leaf syrup drink with other aroma is more acceptable to the participant than the original fig leaf syrup. 2. The fig leaf syrup subverts the general impression that the fig can only be eaten its fruit, although the overall acceptance and purchase intention in the trial was only barely acceptable. 3. In terms of overall acceptance, it may be that the sour taste was too prominent after adding lemon which resulting in a conflict of taste and lead to the most unpopular acceptance by the participants. 4. The favor of mint leaf syrup which taste may similar to the common herbal tea on the market, so it is more popular with the participants.

In the future, the benefits of fig leaves can be promoted through the education or news media, so that consumers can further understand the fig leaf syrup products and purchase them. In addition, try some other herbs such as aromatic marigolds and geraniums, which can cover the special taste of fig leaves. Future researchers are advised to combine fig leaf syrup with other beverages such as milk to make it more developmental and commercially valuable.

References

1. Jiang. S. S. (2012). Eating on health food!Jiang.s.s teach you to be smart and eat really healthy. New Naturalism Press. Taipei.

2. Jang Y. D. & Yuan S.(2012). Application of selenium in the prevention and treatment of digestive diseases. *Clinical metaphor*, 27(3), 63-266.

3. Dong. S. S., Zhu Y., An, D. K., Wang, B. S., Gao, L. & Zhang, Q. F. (1992). Determination of Trace Element Selenium in Chinese Medicine Fig and Its Oral Liquid. *Chinese Journal of Modern Applied Pharmacy*. 2, 57-59.

4. Chen, J. Y. Ni, J. J. & Sie, C. C.(2011). Investigation on fig horticultural traits. National Chung Hsing University Thesis.

5. Luo,W. J. (2009). Natural and delicious healthy jam. KATE Publishing. Taipei County, Taiwan.

6. Chen, J. S. (2004). Sugar into the room . Scientific development. 384, 56-61.

7. Ishihara Yui. (2006). Dr. Ishihara tells you.Eat like this Eat more healthier. Sanyue Culture Books Co., Ltd. Taipei County, Taiwan.

8. Jellinek, G. (1985) *Sensory Evaluation of Food*. VCH Publishers. Florida. U.S.A.

9. Lin, Q. S. & Ye, R. Y.(2004). Application and practice of sensory evaluation。Rui yu Publishing Co., Ltd. Pingtung Taiwan.

10.Dodds, W. B., Monroe, K. B. & Grewal, D. (1991). Effects of price, brand, and store information on buyers'product evaluations. *Journal of Marketing Research*. 28(3), 307-319.

11. Biswas, A. & Blair, E. A. (1991). Contextual effects of reference prices in retail advertisements. *Journal of Marketing*, 55(3), 1-12.

Table 1 Difference Analysis of Fig Leaf Syrup

Item	Products	Mean/Standard deviation	F	significance
Appearance	Origin	2.79±0.988^{a}	11.861	.000***
	Lavender	3.16±0.813^{b}		
	Lemon	3.35±0.914^{bc}		
	Mint	3.52±0.915^{c}		
Aroma concentration	Origin	2.58±1.156^{a}	16.081	.000***
	Lavender	2.77±1.014^{a}		
	Mint	3.21±0.977^{b}		
	Lemon	3.60±1.371^{c}		
Aroma preference	Origin	2.67±0.900^{a}	4.122	.007**
	Lemon	2.69±1.237^{a}		
	Lavender	2.83±0.995^{a}		
	Mint	3.13±1.022^{b}		
Sweetness concentration	Lemon	2.17±1.207^{a}	12.898	.000***
	Origin	2.70±1.049^{b}		
	Mint	2.95±1.019^{bc}		
	Lavender	3.02±1.005^{c}		
Sweetness preference	Lemon	2.04±1.136^{a}	15.760	.000***
	Origin	2.80±1.119^{b}		
	Lavender	2.88±1.066^{b}		
	Mint	3.00±1.054^{b}		
Astringent	Origin	1.72±1.092^{a}	10.664	.000***
	Lavender	1.80±0.953^{a}		
	Mint	1.95±1.104^{a}		
	Lemon	2.56±1.452^{b}		
Aftertaste	Lemon	2.13±1.253^{a}	11.931	.000***
	Origin	2.74±1.134^{b}		
	Lavender	2.82±1.149^{b}		
	Mint	3.08±1.116^{b}		
Overall acceptance	Lemon	2.12±1.241^{a}	14.098	.000***
	Lavender	2.93±1.191^{b}		
	Origin	2.93±1.157^{b}		
	Mint	3.12±1.157^{b}		
Purchase Intention	Lemon	1.87±1.178^{a}	11.190	.000***
	Origin	2.45±1.149^{b}		
	Lavender	2.54±1.158^{bc}		
	Mint	2.81±1.245^{c}		

The Research and Development of Fruit Puffed Rice

Shang-Hui Li, Yi-Ru Wang, Yi Ting Liu, En-Yi Lu, Cheng-Han Li, Fang-Quan Zhang
Department of Food and Beverage Management, Far East University, Taiwan (R.O.C.)
E-mail: a0956358700@gmail.com

Abstract

In recent years, Agriculture and Food Agency, Council of Agriculture, Taiwan, has vigorously promoted rice food products such as rice bread, rice noodles, rice seed strips, rice instant noodles, rice cakes, rice puffs and other products. Puffed rice is one of the most common traditional snacks in Taiwan. The ingredients are mostly rice, maltose, brown sugar, water and salad oil. But snacks on the market have gradually diversified due to year by year, moreover, the commercially available puffed rice flavor is too monotonous and generally not popular with young people. In this study, to make puffed rice based on natural freshly squeezed fruit juice with four different formulas of original flavor, pineapple, lemon and grape for sensory evaluation.100 participants were selected from Far East University, Tainan, by means of a convenient sampling method for formal favorite sensory evaluation. 95 valid questionnaires were collected. The evaluation included rice flavor, sweetness, acidity, degree of adhesion, brittleness and overall acceptance, and the results were statistically analyzed with SPSS12.0.The study results showed that pineapple-scented puffed rice had the highest level on crispness, the overall acceptance and the willingness to purchase. However, lemon- scented puffed rice had a low scent, low overall aroma, less fragile, high acidity, and low on overall acceptance and willingness to purchase. The most suitable formula for this study was rice scent 300g, pineapple juice 90g, salt 3g, maltose 90g, salad oil 24g, and dark brown sugar 160g.

Keywords: Puffed Rice, Juice, Sensory Evaluation, Purchase Intention

1. Introduction

Rice is the staple food for Taiwanese. However, the population of rice in Taiwan is getting less, and the excess rice has existed for many years. Due to Taiwan imported rice from many foreign countries which had already affected to a considerable extent with the small farming, but costly rice industry. Despite the government's efforts to promote rice-processing industry, the excess rice remains unresolved. The fundamental problem is the transformation of dietary westernization, diversification or eating less starch [1]. In recent years, the AFA has vigorously promoted rice food and encouraged local people to eat more rice products.

The purpose of this study is to give a further innovation of combining fresh fruit juice with puffed rice as shown on figure 1. The research flow chart. Meanwhile, use sensory evaluation to explore consumer preferences and develop optimal formulas. Based on the motivation above, the purposes are as follows: 1. develop the best formula for new flavored puffed rice; 2. explore consumers' preferences for different flavors of puffed rice; 3. explore consumers' willingness to purchase new flavored puffed rice.

2. Literature Review

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. Regarding the origin of puffed rice, it was said that long time ago when the ancestors entered to Taiwan from Fujian, a coastal province of China, and built temples. Due to the rice utensils were made by bamboo wares, the bamboo and rice were grilled into a gray and cracked rice when met fire. However, the abbot considered the food was hard-won, and then stuck sugar to eat as a meal. According to the making process, it is known that the puffed rice is made up of high temperature and high pressure, and the instantaneous burst with sugar to shape, crisp and more delicious and delicious[2].

Wang, Guo An, a puffed rice producer, pointed out that japonica rice is more fragrant and viscous than other kinds of rice for making puffed rice. In addition to japonica rice, many other ingredients are used to make the puffed rice completely: 1. Dark brown sugar, a level between white sugar and light brown sugar, the color comes from a small amount of minerals and organic matter, with a strong hygrosc Tpicity, which is the most commonly used in general cooking seasoning[3]; 2. Maltose, or malt sugar, is rarely found in natural foods, mainly as a product of digestion or decomposition of starch[4]; 3. Salad oil has the function of providing a lubricious taste and enhancing the product's loose and crispy properties [5]; 4. Salt can enhance the special flavor and reduce the sweetness [6]; 5. Water helps to mix and dissolve materials and mediate the softness of food and adds water retention [7]. In this study, fresh fruit was added to the origin puffed rice trying to increase its flavor such as pineapple, grape and lemon are added separately, and they are rich in vitamin B and C [8].

Sensory Evaluation relies on the human’s eyes, nose, mouth, skin and ears, that is, the sense of smelling, tasting, touching and listening. It also measures and analyzes foods by psychology, physiology, physics, chemistry and statistics, and measures the characteristics of product quality and the degree of preference of people's senses [9].

[10] pointed out that there are two main methods for sorting out food sensory evaluations: the first one is a consumer test such as a preference or an affective test, the second is analytical tests including sensitivity testing, and analysis of differences and descriptions. The difference analysis is mainly to understand whether there are any differences and differences between products. This study used a preference test to fill out a questionnaire with a scoring method to find out consumers' preferences and purchase intentions for new flavors of puffed rice.

Considered intention to be the subjective probability of an individual engaging in a particular action [11]. Argued that the purchase intention refers to the likelihood that the consumer is willing to purchase the product, and the higher the intention to buy, the greater the chance of purchase [12]. Argued that purchase intention is a desire of a consumer to purchase a product or to visit a service store [13]. In addition to measuring purchase intentions from multiple perspectives, some studies measured in a single facet only [14], which directly measured purchase intentions with a single question.

3. Research Methods

Make the puffed rice prior to questionnaires. First, place an appropriate amount of rice into a pressure boiler and continue rolling heating for about 15 minutes. Followed, cook juice with salt, maltose, salad oil and dark brown sugar to become liquefied by medium heat. After that, pour it into the container and rapidly mix. Finally, pour the well-mixed rice into the wooden mold and flat evenly. The puffed rice can be cut into pieces after it slightly cool.

Perform two preliminary tests. The preliminary test was carried out twice, and experienced 20 teachers and students from Far East University were selected for the sensory evaluation. The results found that the dark brown sugar puffed rice was the most popular by the tester, so the formal test was based on that formula.

Questionnaires. Sample number order of the formal test was randomly selected and the formula is as shown in Table 1. A total of 100 participants from Far East University were invited to attend the sensory evaluation. 5 points score system was taken, the 1 is the least favorite or the weakest, 5 is the most favorite or the strongest, and 95 valid questionnaires were collected. The collected valid questionnaires were statistically analyzed with SPSS 12.0, and the Duncan method was used to measure the differences between the samples using the one-factor variability.

4. Research Result

Analysis results of four different flavors of puffed rice are shown as Table 2. In terms of appearance, there is a significant difference compared pineapple with original and grape flavor ($p<0.05$). However, the appearance of all types of puffed rice could be accepted by the participants, the average value is more than 3.

As for fragrance, the participants thought that the lemon-favored puffed rice was the lightest (average 2.44) and had a significant difference from the others ($p<0.001$). The participants thought that the lemon-favored puffed rice was the weakest and were significantly different from the others ($p<0.001$) on its overall aroma (mean 2.62), sweetness (mean 2.00), sourness (mean 4.26) and the brittleness (mean 2.75). In terms of sticky teeth, the results of Duncan analysis, the participants considered that the original flavor was the stickiest (mean 3.54). Regarding overall acceptance, the pineapple-favored puffed rice was the highest (average 3.91), which was significantly different from the others ($p<0.001$); Refer to purchase intention, the average of three kinds of puffed rice exceeded 3 except lemon-favored one (mean 1.88), The results showed that pineapple-flavored puffed rice was the most welcome by the participants, which was the most suitable for this study.

5. Conclusion and Suggestion

The study found that the participants could receive the taste of all new fruit flavor puffed rice except lemon-favored one.

(1). In detail analysis that pineapple-favored puffed rice has the highest crispness, the highest overall acceptance and high purchasing intention, and it is obvious that which has the possibility of marketing in the future. In the future, it could combine the promotion of rice food with government, or promote the new favored puffed rice through print media, social media such as Line, Facebook, etc.

(2) Lemon-favored puffed rice was in a low acceptance in brittleness, acidity, overall aroma, overall acceptance and purchase intention. It may be because the lemon itself is sour, which leads to low acceptance of the participants.

(3). The most suitable formula for this study was puffed rice 300g, pineapple juice 90g, salt 3g, maltose 90g, salad oil 24g, and dark brown sugar 160g. Due to limitation in time, budget and human power, it is impossible to make more research on different flavors of puffed rice. The future researchers can refer to this formula, and try to add more different flavors of fruit for improvement.

(4). This study conducted sensory evaluations limited on the participants of Far East University, it could expand to different cities, different age and different groups in the future.

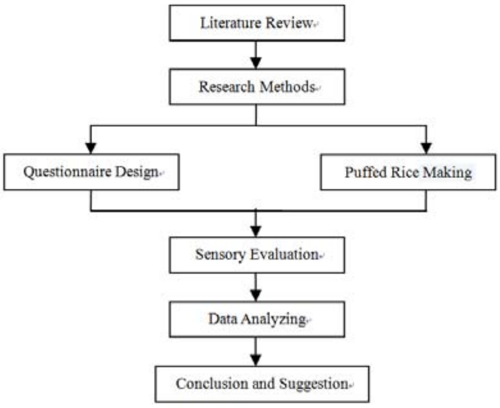


Table 1 Puffed Rice Test Formula

Items	Origin	Grape	Pineapple	Lemon
Dried Rice 300 g	ν	ν	ν	ν
Salt 3 g	ν	ν	ν	ν
Maltose 90 g	ν	ν	ν	ν
Salad Oil 24 g	ν	ν	ν	ν
Dark Brown Sugar 160 g	ν	ν	ν	ν
Water90 g	ν			
Grape Juice 90 g		ν		
Pineapple Juice 90 g			ν	
Lemon Juice 45 g				ν
Water 45 g				ν

Table 2 Difference Analysis of Puffed Rice

Item	Product	Mean/Standard Deviation	F	Significance
Appearance	Origin	3.69±0.787^{b}	4.242	.006 **
	Grape	3.81±0.842^{b}		
	Pineapple	3.40±0.749^{a}		
	Lemon	3.57±0.941^{ab}		
Fragrance	Origin	3.19±1.179^{b}	16.161	.000***
	Grape	3.33±0.844^{b}		
	Pineapple	3.33±1.036^{b}		
	Lemon	2.44±1.028^{a}		
Sweetness	Origin	3.49±0.886^{b}	46.492	.000***
	Grape	3.24±0.942^{b}		
	Pineapple	3.21±0.988^{b}		
	Lemon	2.00±1.011^{a}		
Sourness	Origin	1.36±0.824^{a}	210.098	.000***
	Grape	1.46±0.712^{a}		
	Pineapple	1.79±1.061^{b}		
	Lemon	4.26±1.054^{c}		
brittleness	Origin	3.27±1.036^{b}	12.211	.000***
	Grape	3.28±1.048^{b}		
	Pineapple	3.72±1.155^{c}		
	Lemon	2.75±1.176^{a}		
Overall acceptance	Origin	3.61±0.854^{b}	92.678	.000***
	Grape	3.48±0.713^{b}		
	Pineapple	3.91±0.787^{c}		
	Lemon	2.04±0.988^{a}		
Purchase intention	Origin	3.52±1.080^{bc}	69.062	.000***
	Grape	3.25±0.875^{b}		
	Pineapple	3.73±0.939^{c}		
	Lemon	1.88±0.988^{a}		

References

1. Song, H. Y. (2015). Multi-rice food promotion situation. Farming and Farming. 2019/9/5 taken from https://www.coa.gov.tw/ws.php?id=2503765.
2. Siao, P. C. (2010). Declining ancient flavor and Puffed Rice.2018/11/14 from http://sixstar.moc.gov.tw/blog/yiheli99/myBlogArticleAction.do?method=doListArticle ByPk&articleId=19813
3. Li, S. H. & Lin, J. L. (2011) .Chinese dessert. Ciying Culture Press Co., Ltd. Taipei, Taiwan.
4. Taiwan Sugar Production. (2015) .2019/5/10 taken from https://contest.cpmah.org.tw/award/2015 /web /C2015b026/layoutcategory.html.
5. Cheng, S. H. (2014). Current Principle of food science. Huaxin Publishing Co., Ltd. Taipei.
6. Hu, J. J. (2013). The second book necessary for baking novices. Happiness Culture Press Co., Ltd. New Taipei City. Taiwan。
7. Wu, J. C. (2007). Baking Food Production Internship. Guangmao Culture Press Co., Ltd. Taichung City.
8. Jonval, L. (2003).Bonjour! Fruit. Locus Culture Press Co., Ltd. Taipei. Taiwan。
9. Yau, L. C. (2012). Sensory evaluation and practical application-Tasting Technician Essentials. SCG Taiwan. Hsinchu City.
10. Qu, S. M. (2012). Food Sensory Evaluation and Exercise. Wagner Publishing Co., Ltd. Taichung. Taiwan.
11. Fishbein, M. & I. Ajzen. (1975). *Belief Attitude Intention and Behavior An Introduction to Theory and Research*, Addison Wesley Reading, Mass: Addision Wesley.
12. Dodds, W. B., Monroe K. B. & Grewal, D. (1991). Effects of prices, brand and store information on buyers' product evaluations. *Journal of Marketing Research*, 28, 307-319.
13. Shao, C. Y., Baker J. A. & Wagner, J. (2004). The effects of appropriateness of service contact personnel dress on customer expectations of services quality and purchase intention: The moderating influences of involvement and gender. *Journal of Business Research*. 57, 1164-1176.
14. Biswas, A. & Blair, E. A. (1991). Contextual effects of reference prices in retail advertisements. *Journal of Marketing*, 55(3), 1-12.

The Research of Health Western Cuisine – A Study of Aloe Vera in Cooking

HisYing Hsieh
Department of Food & Beverage, Far East University,
No.49, Zhonghua Rd., Xinshi Dist., Tainan City, 74448, Taiwan (R.O.C.)
E-mail: dianeh@mail.feu.edu.tw

Abstract

Aloe vera, is widely used in food, medicine and cosmetic products with its health component. For more choices in health cuisine, the study aims to understand the purchase intention of aloe vera in cooking for health western cuisine . The questionnaire will be distributed to consumers in Tainan City. Data analysis of description ,T-test, and Anova will be conducted by SPSS software. Based on the results of study, the future development and utilization of aloe vera cuisine are also discussed.

Key word: Aloe vera, purchase intention, health cuisine

1

1. Introduction

2000 years ago the Greek scientists regarded aloe vera as the universal panacea and the Egyptians called aloe "the plant of immortality." Today, the aloe vera plant has been used for various purposes in dermatology(Surjushe, Vasani, & Saple,2008). Aloe vera, as one of the green resources of high economic and medicinal value, has been explored and used widely increasingly. Aloe belongs to a health food. Its nutrition and safety is in line with the requirements of the UNFAO(Hamman, 2008) . Developing and utilizing its economic value has attracted more attention from the health food industry and aloe vera is planted on a large scale.

In the late twentieth century, the pressure of global competition, the opportunities and threats from the changes in the economic market circumstances and the coming aging population society could create a huge benefits, for example, the healthy aging business (the health and organic food, green hotels and the recreational sports) .

In Taiwan various health food products are very popular. Recently health cuisines are made from health plant in Chinese cuisine, but not too many in western cuisine. The study tries to find out the acceptance and purchase intention of aloe vera in western cuisine.

2. Main Text

2.1. Aloe vera

Most of the whole aloe vera leaf is water and, there are more than 200 chemical substances in the dry matter constituting the leaf (Rodríguez, Martín, & Romero,2010). Aloe vera is being cultivated in other areas with different climatic conditions and such as onion, garlic, and asparagus, which are known to have medicinal properties (Lawless and Allen, 2000). Most of these plants originated in the dry regions of Africa, Asia, and Southern Europe, especially in the Mediterranean regions (Urch, 1999). Aloe vera is the most commercialized aloe species and processing of the leaf pulp has become a large worldwide industry. In the food industry, it has been used as a source of functional foods and as an ingredient in other food products, for the production of gel-containing health drinks and beverages. In the cosmetic and toiletry industry, it has been used as base material for the production of creams, lotions, soaps, shampoos, facial cleansers and other products.

Despite the strategic importance of these three issues, there is surprisingly little scientific research available to support food companies in making these decisions. For example, besides some general rankings of most popular health claims and health concerns (Young, 2000),

2.2. Health food and western cuisine

The media has provided various coverage of health related topics which has made consumers increasingly health conscious and desirous of a better quality of life. Also, the soaring environmental costs have given rise to "green" consumers (Peattie and Ratnayaka, 1992) and health concern customs. In a geographical context, the consumers of Central and Eastern European countries are demanding safe and nutritious food items on account of increasing globalization, education and income levels (Anderson, 2000; Zakowska-Biemans, 2011). At the same time, Asian consumers are also gravitating towards organic food due to apprehensions about the commercially grown food as being risky (Williams and Hammitt, 2001) and partly in charge of the high risk of lifestyle diseases, such as obesity and diabetes and cardiovascular disease (Cummins, 2001; Schmidt, 1999; Birchard, 2001). Usually the western food menu is including salad, soup, main course, desert and drink.

2.3. Purchase intention

Seyfang (2006) conducted the first empirical study of Eostre Organics, an organic food producer cooperative in the United Kingdom, to show that 65.2% of respondents cited a desire to support and strengthen the local economy and community, such as greater selfreliance and independence from global corporations and supermarkets. These aspirations enhance some prominent consumers to purchase organic food to set an example and inspire others to modify their consumption patterns for societal benefit (Canavari and Olson, 2007). Social influence is a broad area that affects one's emotions, opinions and behavior (Wood and Hayes, 2012), especially in purchase intention.

3. Method

Data were analyzed with descriptive analysis, t-test, and one-way analysis of variance. The study will conduct the questionnaire survey, 250 questionnaires and employer SPSS 18 to analysis data; purchase intention, consumer preference of aloe avra cuisine and demography individuals. It is understood that each consumer behaves differently and consumes products that suit their personality.

4. Conclusion

The participants in this study were Tainan's consumers, To explore the impact factors of consumer behavior, product exception, and Purchase intention were issued by Random sampling surveys of 270 questionnaires distributed from October 1st. to October 31st, 2017, and used SPSS12.0 software. The results of the analysis hope to provide to the industry reference to make a marketing strategies.

5. Implication

In the future, according to results the study will propose the practical implication and future research suggestion in market of health cuisine. The aloe vera in cooking will bring more economic benefits in western cuisine. The consumers who concern personal health will have more intention to purchase the aloe vera food in western restaurants.

6. Reference

1. A.Surjushe, R. Vasani, R., and D. G. Saple, Aloe vera: a short review. *Indian journal of dermatology*, *53*(4), 2008, pp. 163.

D. Urch, "Aloe vera-Nature's gift." *Great Britain: Blackdown Publications* 15 (1999).

2. E. R. Rodríguez, J. D. Martín, and C. D. Romero , Aloe vera as a functional ingredient in foods. Critical Reviews in Food Science and Nutrition, 50(4), 2010, pp. 305-326.

3. J. H. Hamman, Composition and applications of Aloe vera leaf gel. Molecules, 13(8),2008, pp.1599-1616.

4. J. Lawless, and J. Allen . Aloe vera: Natural wonder care.(1st Edn, Thorsons,2000).

4. K. Birchard, Irish Ombudsman finds medical records "atrocious". *The Lancet*, *358*(9275), 2009, pp48.

K. Peattie & M. Ratnayaka, Responding to the green movement. *Industrial Marketing Management*, *21*(2), 1992, pp103-110.

5. M. Canavari ans K. D. Olson, *Organic Food*. Springer Science+ Business Media, LLC (2007).

6. M. Schmidt, "Luminosities and space densities of gamma-ray bursts." *The Astrophysical Journal Letters* 523.2 (1999): L117.

7. S. Żakowska-Biemans, Polish consumer food choices

and beliefs about organic food. *British Food Journal*, *113*(1), 2011,pp 122-137.
8. T. Reynolds . Aloe vera: what are the active ingredients? Proceeding of the Active Ingredients Conference.1998, pp. 85–93. Paris.
9. W. Wood, & T. Hayes, Social Influence on consumer decisions: Motives, modes, and consequences. *Journal of Consumer Psychology*, *22*(3)2012, pp. 324-328.

The Influence of Attitude, Subjective Norm, Perceived Behavior Control on Purchase Intention – A Study of Green Restaurants in Taiwan

Hsiying Hsieh
Department of Food & Beverage, Far East University,
No 49, Zhonghua Rd., Xinshi Dist., Tainan City, 74448, Taiwan (ROC)
E-mail: dianeh@mail.feu.edu.tw

Abstract

Green food is widely promoted for personal health and environmental benefits. Thus, decision making on the consumers' purchase intention of green restaurants is a critical issue. The purpose aims to understand how attitude, subjective norm and perceived behavioral control influence purchase intentions of green restaurant. The questionnaire surveys were conducted to customers of green restaurants in Taiwan. Based on the finding, attitude, subjective norm and perceived behavioral control were significantly affect purchase intention. Implication and suggestion are then discussed.

Keywords: green restaurant, attitude , subjective norm, perceived behavioral control , purchase intention, theory of planned behavior

1. Introduction

There is public acceptance that unmitigated economic development is not only environmentally harmful, but it can also be detrimental for society. As a result, environmental protection became prominent issue and acquainted with public and (Schubert, Kandampully, Solnet, & Kralj, 2010). Environmentally friendly food has risen quickly over recent decades. Increasing consumer interest in eco-friendly food has brought a number of changes to the restaurant industry (LaVecchia, 2008). Understanding the key determinants of behavior has been a paramount goal for many theorists in the social and decision-making sciences. The underlying psychological assumption driving the connection between intentions and behavior is that most human behavior is under volitional control (Ryan, 1970).

2. Main Text

2.1 Green Restaurant

Considering such high consumption rates, it is patent that unmitigated restaurant practices would contribute to the depletion of the natural environment. Such practices include: the construction of restaurant facilities that destroy the natural environment; excessive usage of resources such as water, electricity and gas.

According to Gilg, Barr and Ford (2005), green restaurant emphasize on 3Rs(reduce, reuse, recycle)and two Es (energy and efficiency). Engaging in green practices can have critical implications for a restaurant in terms of cost management, market differentiation, as well as environmental protection.

The purchase of eco-friendly products such as sustainable food, nontoxic cleaning and chemical products, and bio-mass agricultural waste products not only helps the restaurant to build a good corporate image, it also expands green practices vertically in the supply chain. For example, it encourages farmers or manufacturers to embrace in green practices in order to supply green products to be used in such restaurants.

2.2 Theory of planned behavior

Fishbein and Ajzen (1975) have defined intention as a person's location on a subjective probability dimension involving a relation between himself and some action". Intentions are the single best predictor of planned behavior and intentions are also an unbiased predictor of action (Bagozzi, Baumgartner, & Yi, 1989). The behavioral intention models have received robust support in numerous behavioral domains (Ajzen, 2001; Eagly & Chaiken, 1993) and are considered to be some of the most widely applied theories in social psychology (Greve, 2001). As mentioned earlier in introduction, the Theory of Planned Behavior (TPB) proposed by Ajzen (1991) explains the consumers' food choice behavior convincingly. Thus, the TPB approach is employed to predict the consumer's purchase intentions of green restaurants.

The behavioral intention (i.e., purchase intention), based on the TPB (Ajzen, 1991), is basically determined by three factors: the attitude that the person holds toward engaging in the behavior (i.e., purchasing attitude), the degree of social pressure felt by the person with regard to the behavior (i.e., subjective norm), and the degree of control that the person feels he or she has over performing the behavior (i.e., perceived behavioral control). The first twofactors reflect the perceived desirability of performing the behavior, while the third reflects perceptions of whether the behavior is personally controllable or not. These three factors predict intention and the ensuing predicts behavior.

2.3 Attitude

The construct of attitude to the purchase of one product is akin to the perceptions of the personal desirability of performing a particular behavior. This attitude depends on expectations of and beliefs in the personal impacts on the outcomes resulting from that behavior. Organic foods are perceived as much more healthy, natural, nutritious, and sustainable than conventional foods (Chen, 2007). Thus, the consumer's attitude to green restaurants purchase is believed to be positively related to the attitude to green restaurants. The hypotheses are proposed as follows.

H1: Attitude will positively affect the purchase intention of green restaurants.

2.4 Subjective norm

Subjective norm deals with a consumer's motivation to perform a behavior which is constructed to incorporate the expectations of what the important people in his or her life (e.g., family, friends, and significant others) (Eagly & Chaiken, 1993; Mowen, 1993) think about performing that particular behavior. If consumers believe that those people important to them think foods at green restaurant are good, then they will have more intention of purchasing food at green restaurant . The hypotheses are proposed as follows.

H2: Subjective will positively affect the purchase intention of green restaurants.

2.5 Perceived behavioral control

One may have a positive attitude towards performing a behavior but may not intend to perform it when faced with a perceived impediment. Perceived behavioral control refers to the consumer's perceptions of personal control over what to buy and eat, which he or she believes to influence the judgment of risks and benefits of green foods in a purchase situation. The hypotheses are proposed as follows.

H3: Perceived behavioral control will positively affect the purchase intention of green restaurants.

3. Methodology

3.1 Research model –

The theory of planned behavior (Ajzen, 1991)

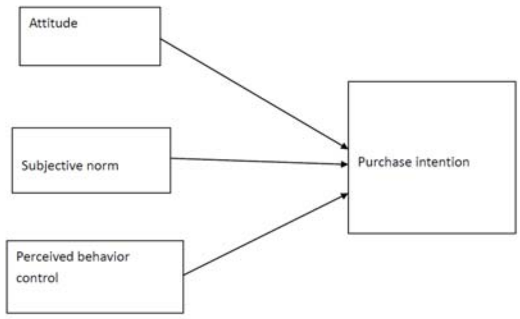


Figure 1 research model- The theory of planned behavior (Ajzen, 1991)

3.2 Measures

The items of attitude were adopted from Sidique et al. (2010) and Koufaris (2002) to access the the concept , the items of subjective norm were adopted from Rook & Fisher (1995) and Tonglet et al. (2004) to access the concept, the items of perceived behavior control were adopted from Taylor & Todd (1995) to access the concept and the items of purchase intention were adopted from Taylor & Todd (1995) to access the concept.

3.3 Sampling

The sampling survey was conducted for consumers who have been to green restaurants in Taiwan and obtained 250 valid questionnaires.

4. Conclusion

Among these respondents, most of them were between the ages of 21 to 40 (60.3%). More than half of the respondents held a bachelor's degree (50.9%) and 51.3 % work in service industries, and more than two-thirds (65.3%) of the female. Hypotheses are all supported as follows. Attitude will positively affect the purchase intention of green restaurants. Subjective will positively affect the purchase intention of green restaurants. Perceived behavioral control will positively affect the purchase intention of green restaurants.

5. Implication

The study proposed practical implication as (1) to enhance positive attitude of consumer on social enterprises product. (2)to promote perceived behavioral control on social enterprises, product.(3) to invite endoser credibility to attract cosumers.

6. References

1.A. Gilg, S. Barr, & N. Ford. Green Consumption or Sustainable Lifestyles? Identifying the Sustainable Customer. *Futures*, 2005, 37.

2. F. Schubert, J. Kandampully, D. Solnet, and A. Kralj, Exploring consumer perceptions of green restaurants in the US. *Tourism and Hospitality Research*, *10*(4), 2010, 286-300.

3. G. LaVecchia, Green: the new gold. Restaurant Hospitality 92 (4), 2008.

4. F. Schubert, *Exploring and predicting consumers' attitudes and behaviors towards green restaurants.* The Degree Master's of Science in the Graduate School of The Ohio State University, 2008.

5. I. Ajzen, The theory of planned behavior. *Organizational behavior and human decision processes*, *50*(2), 1991, 179-211.

6. S. Taylor, and P. A. Todd,. Decomposition and crossover effects in the
theory of planned behavior: A study of consumer adoption intentions.*International Journal of Research in Marketing*, *12*(2),1995a, pp.137-155.

7. S. Taylor, S. & P.A. Todd, . Understanding information technology usage: A test of competing models. *Information Systems Research*, *6*(2),1995b, 144-176.

A Rack and Pinion Driven Mechanical Footstep Power Genarator

Ammar Abdulaziz Al-Talib, See Kai Sheng

Faculty of Engineering, Technology and Built Environment, UCSI University,
No.1, Jalan Menara Gading, UCSI Heights
56000 Cheras, Kuala Lumpur, Malaysia
Email: ammart@ucsiuniversity.edu.my, ksheng.sks@gmail.com
www.ucsi.edu.my

Abstract

The kinetic energy produced by human footsteps during movement is commonly wasted. By using a non-conventional way, it is possible to convert this energy into electrical energy. Due to the exponentially growth in populations, the energy consumption has reached its crisis level while the fossil fuel is depleting significantly for years. Hence, the energy which is going to be wasted when humans are walking, can be harvested to provide a promising solution to this issue economically and affordably. In such a scenario, a mechanical footstep power generator has been proposed to transfer the kinetic energy of the human footsteps into useful electrical energy. In this research, some modifications have been implemented to further refine the performance and the efficiency of a footstep power generator studied previously and works on the same principle of rack and pinion. To justify and evaluate the feasibility of the proposed concept, a prototype which only applying the rack and pinion mechanism principle has been fabricated and tested. In this system, the power is generated when force is applied as a result of the weight of a person walking on the top plate, and causing the rack and pinion mechanism to rotate a dynamo and to store the produced electrical power in a battery. Individuals with weight ranging from 35kg to 75kg have been invited to participate in the experiments by stepping and jumping onto the top plate. The results are acquired to compare with the theoretical results. When it is compared to the conventional energy sources, the power generation from human footsteps can be considered continual, environmental friendly and also it can be easily accessed.

Key words: Human Footstep; Renewable Energy; Footstep Power Generation; Rack and Pinion

1. INTRODUCTION

In the era of globalization, non-conventional energy is highly demanded, as the conventional energy sources such as fossil fuels are depleting spectacularly nowadays for the sake of the growing energy consumption. The development of national economy may be negatively impacted due to the potential shortage and rising price of oil and gases and have escalated the concerns about security of energy level in the future [1]. The utilization of the wasted energy from human footsteps is highly needed to replace the conventional energy [2]. A footstep power generator is a device that can transfer the energy exerted by footsteps into useful electricity, supplying the renewable energy in very low cost [3]. The footstep power generator is an environmental friendly energy source as it consumes zero fuel to generate electricity. Thus, it does not release greenhouse gas or any harmful waste that pollutes the environment.

To generate this type of renewable energy, a footstep power generator has been proposed because it utilizes the source of energy from human footsteps. It simply relies on human locomotion to generate electricity. Walking is a normal daily activity, easy to fit into our daily routine and is a simple way to create a nonstop energy that can be used to develop electricity whenever the human walks. When a person start walking, energy is discharged to the ground surface in the form of vertical force, and this force can be converted into vertical displacement to generate the proposed electrical energy. This device can be installed in the footpath at crowded areas to harvest footsteps and convert its energy into an electrical form [2].

Piezoelectric and Mechanical footstep power generators are the two fundamental types of footstep power generators known in the current market. The piezoelectric footstep power generator uses piezoelectric transducer to transfer the force that oltage. Moreover, it is essential to choose the appropriate ferroelectric material used for the piezoelectric transducer since this power generator uses piezoelectric transducer as the main component to convert human motion into electricity. Owing to the fact that the efficiency of piezoelectric transducer to transform force applied into voltage depends on the type of ferroelectric material used [4]. Besides that, most of the mechanical footstep power generators uses gear, pulley, rack and pinion, and chain with sprocket to transfer human motion into electricity [5].

In this research, mechanical footstep power generator studied by Kit,[11] is opted to be modified based on the concept of rack and pinion so as to improve the conversion of the energy from human footstep into electrical energy. The mechanical footstep power generator usually consists of springs, gearwheel arrangement, rack and pinion, shafts, bearings, DC generator and a battery. It converts the force applied from human footsteps in the form of mechanical energy into useful electrical energy. In this paper, Section II will describe the design and fabrication of the mechanical footstep power generator whereas Section III and Section IV explicate the experiments conducted and discuss the results obtained respectively. Lastly, Section IV will provide the conclusion of the entire work.

2. DESIGN AND FABRICATION

By utilising the rack and pinion combination, the mechanical footstep power generator has been designed to remove the complexities from the power transmission system. Figure 1 shows the assembly of the mechanical footstep power generator built for the current research.

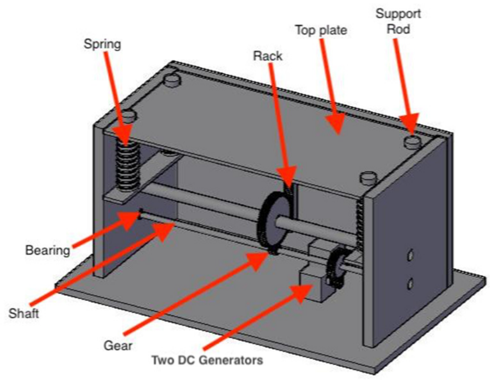


Figure 1: The Proposed assembly for the mechanical footstep power generator used in the current study.

As shown in Figure 1, the working principles of the modified mechanical footstep power generator are:

1. Through human footstep on the top plate, force is vertically applied to the system.
2. The rack moves downward together with the top plate and rotates the first gear,
3. The first gear transfers the rotational motion to the second gear that attached at the lower shaft,
4. The third gear is connected to the same shaft while rotating at the exact same speed and direction,
5. There are two DC generators coupled with the fourth and fifth gears respectively, and are driven by the third gear.
6. The rotation of the last gears is transferred to the two DC generators to convert kinetic energy into electrical energy
7. The power generated is to be stored in a battery through an electronic circuit.

The specifications of gears and springs of the proposed mechanical footstep generator are illustrated in Tables 1 and 2 respectively. The values shown are used for the calculation of the theoretical results.

Table 1: Specifications for gears in the proposed mechanical footstep power generator

	Gear 1	Gear 2	Gear 3	Gear 4	Gear 5
Number of teeth	72	12	36	12	12
Module	1				
Outside diameter (mm)	74	14	38	14	14
Root diameter (mm)	69.5	9.5	33.5	9.5	9.5
Base circle diameter (mm)	67.658	11.276	33.829	11.276	11.276
Tooth thickness (mm)	1.571				

Table 2: Specifications for spring in the proposed mechanical footstep power generator

	Spring
Wire diameter, d (mm)	3
Mean diameter, D (mm)	26.5
Outer diameter, Do (mm)	29.5
Inner diameter, Di (mm)	23.5
Total coil, Nt (turns)	15
Free length, Lo (mm)	95
Spring constant, k (N/mm)	3.32

3. EXPERIMENTAL SETUP AND DATA COLLECTION

The experiments on this prototype have been carried out by associating the individuals without foot disabilities to step onto the prototype directly in one test sets, and then to jump onto the prototype directly in another test sets, in order to operate the mechanism which will generate the electricity. The participants whom involved in the study are weighted between 35kg to 75kg to conduct the test. A multi-meter has been used to measure the output from the two DC generators which are installed in the system. In addition, a weighing balance is prepared to measure the weight of the participants before carrying out the experiments. The multi-meter is then connected to the output of the DC generators by using crocodile clips to measure the voltage and electrical current generated from each of the footsteps.

The participants have been asked to step onto the prototype and then to make the top plate fell to the maximum depth. It has been repeated for at least 10 times for each individual, hence, the voltage and current have been measured separately to obtain the average value. While conducting the test, the multi-meter has been set at the range of 20V to measure the voltage output and the current has been measured by setting the range at 20mA. Moreover, another test has been conducted by having the participants to jump onto the top plate of the prototype from a height of 15cm and the readings from the multi-meter have been also recorded. The average reading of voltage is multiplied with the average current to obtain the power generated for each of the steps and jumps.

4. RESULTS AND DISCUSSION

Results for the experiments conducted on the prototype used in the study will be presented in this section,

4.1 Experimental Results Vs Theoretical Results for Normal Stepping:

Figure 2 shows a comparison between the experimental power generated and the theoretical power calculated. From Figure 2, it is clear that the power generated from the two 12V DC generators has increased when the mass of a person stepping on it increased. The average percentage of error between experimental and theoretical power has been calculated to be within 9.59% only. Then, the results obtained are considered reasonable. It is obvious that the experimental power for each stepping weight are all lower than its theoretical power. This can be referred to the losses due to the friction and may be imperfect alignment of some components. The theoretical power is calculated by using the formulae shown in Eq. (2) and Eq. (3).

$$Theoretical\ power = (weight - spring\ forces) \times \frac{displacement\ of\ top\ plate}{seconds} \quad (2)$$

$$Theoretical\ power = \{mg - 4k(x)\} \times \frac{x}{60} \quad (3)$$

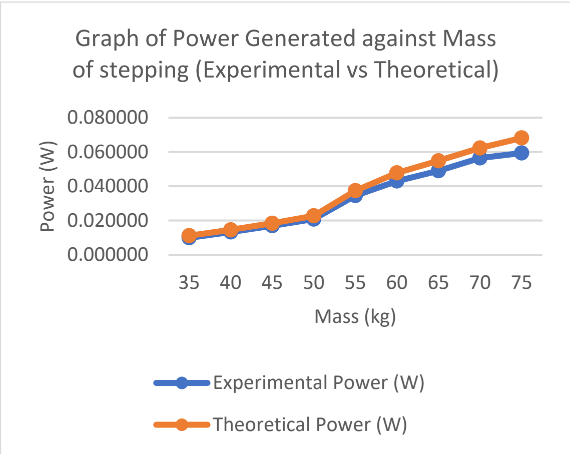


Figure 2: Graph of experimental and theoretical power generated against stepping mass for normal stepping

4.2 Experimental Results Vs Results from Kumawat,[6],Research Paper for Normal Stepping:

According to Kumawat,[6], the results have been recorded by having the setup using piezoelectric sensors to generate power, while in this current project, only 2 DC generators have been used to generate the power. The experimental power generated from the current prototype have been compared with the results acquired by Kumawat,[6] who have used piezoelectric sensors arrangement. . The comparison has shown that the current mechanical footstep power generator had better performance and producing more power as shown in Figure 3.

To get another set of results in the current study, the participants have been asked to repeat the test by jumping onto the top plate of the mechanical footstep generator from a height of 15cm. Figure 4, can demonstrate that the greater the mass of participants, the greater the power to be generated from the mechanical footstep power generator. The average percentage of error is slightly higher at the value of 12.72% when it is compared to the test involving only normal stepping.

When a person jumped onto the top plate, the person may not fully exert the force to the middle of the top plate. In this situation, the force applied will be concentrating either close to the left side or right side of the top plate which have affected the performance. In the stepping test, the participants have stepped onto the top plate by one foot after another, which means that they can be focusing their feet at the middle of the top plate, but for the case of jumping, the participants have reached the top plate on both feet in which they did not have a guide to focus the force in the middle of top plate. In general, the percentage of error has been found to be less than 15% .

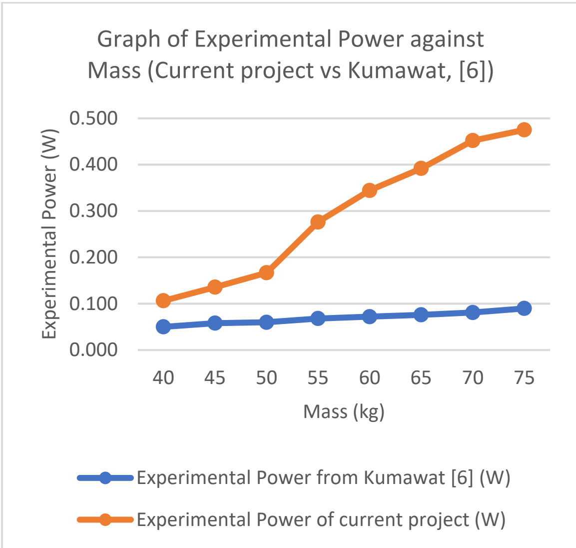


Figure 3: Graph of experimental power generated against mass for the current project and the research conducted by Kumawat[6]

4.3 Experimental Results Vs Theoretical Results for Jumping

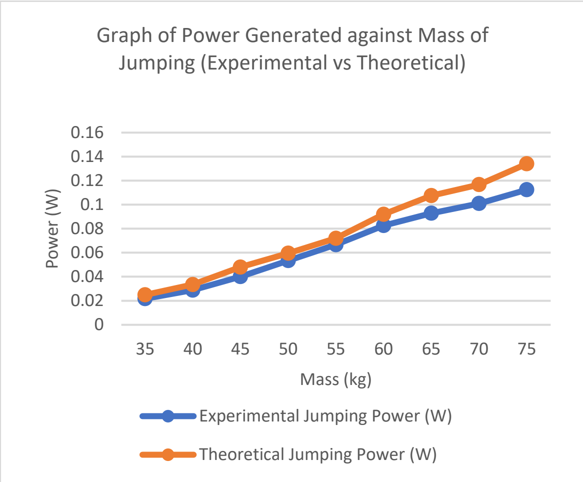


Figure 4: Graph of experimental and theoretical power generated versus mass for jumping

4.4 Experimental Results of Normal Stepping Vs Jumping

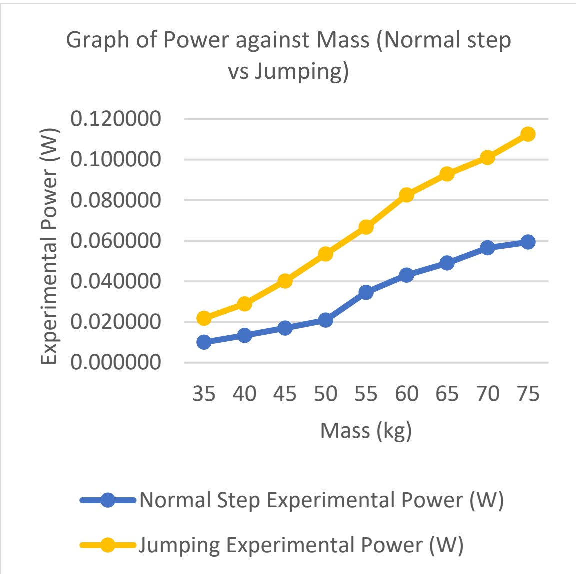


Figure 5: Graph of experimental power generated against mass for the normal stepping and the jumping condition

According to Figure 5, the performance of mechanical footstep power generator can be improved by increasing the force applied onto the top plate by jumping instead of normal stepping. In this case, the participants jumped from a height of 15cm above the top plate. The participants at that height have possessed more gravitational potential energy to be added to the driving force on the system. It is concluded that the force applied to the top plate of the mechanical footstep power generator under the study will be greater by having the participants to jump rather than just stepping on it.

5. CONCLUSION

This research has been successfully proved that the concept of power generation from human footsteps wasted energy is possible. It can be considered as a new source of non-conventional energy techniques. The use of only rack and pinion combination for this mechanical footstep power generator has simplified the complexity in any other mechanisms studied before, like the rack and pinion with pulley system [7], the rack and pinion with chain drive system [8-9], the rack and pinion with flywheel system [10], and also, the one generator rack and pinion system [11]. The mechanical footstep power generator can operates better in crowded areas, and it can make a noticeable impact in the countries that have high populations such as India and China to produce high electricity from human footsteps in crowded areas like airports, bus stations and all shopping malls entrances. In a nutshell, all the experiments have proven the proper performance and feasibility of the proposed design and the results obtained are promising.

Nomenclature

Symbol	Description	Dimension / Unit
m	Mass of participant	kg
g	Gravitational acceleration	ms^{-2}
k	Spring stiffness	N/mm
	Displacement of top plate	mm

Acknowledgement

The authors wish to express the gratitude towards the management and Faculty of Engineering, Technology and Built Environment (FETBE) of UCSI University for the guidance and support towards the completion of this research.

REFERENCES

[1] Sarat Kumar Sahoo, Shubham kumar, Pankaj Kumar Yadav, Rishav Kumar, 2016. Foot Step Power Generation. International Journal of Mechanical Engineering and Technology (IJMET), 7(2), pp. 187-190.
[2] R.S.Pokale, T.R.Deshmukh, 2016. Design and analysis of a Mechanical Device to harvest energy from human footstep motion. International Journal of Innovative and Emerging Research in Engineering Volume 3, Special Issue 1, ICSTSD 2016
[3] HR, Nyan,2015. Power Generation Using Piezoelectric Material. Journal of Material Science & Engineering, 4(3), pp. 1-4.
[4] Krishna Jaiswal, August 2013. A brief comparison of renewable and non-renewable resources of energy. Research Gate
[5] B.Munaswamy, Ch.Prudhvi, V.Srikanth, B.Kirankumar, E.Pradeep Kumar, April 2018. International Journal of Engineering Trends and Applications (IJETA), 5(2), pp. 291-298.
[6] Suryaprakash Kumawat, Altaf Rawoot, Ajinkya Wakade, December 2019. Foot Step Power Generation. International Journal of Science and Research (IJSR), 7(12), pp. 879-882.
[7] Amel Farouk, and Ahmed F. Abdel Gawad, October 2015. Utilization of Human Footsteps for Power Generation. Research Gate, pp. 1-13.
[8] Abhishek N, Shivasharana Yalagi, May 2016. POWER GENERATION BY FOOT STEPS USING RACK. International Journal of Engineering Research And Advanced Technology (IJERAT),2(1), pp. 10-14.
[9] N,Modi, P.Shrivastava, R.Bhardwaj et al, 2016. Generation of Electricity through Footstep, Page852-854.
[10] Ali Muhammad, Asif Alib, Sumera Bibic, Nadeem Ahmed Tuniod, 2016. Foot Step Power Generation System. Energy, Environment and Sustainable Development 2016 (EESD 2016).
[11] Chun Kit Ang, Ammar A. Al-Talib, Sook Meng Tai, and Wei Hong Lim, 2019. Development of a footstep power generator in converting kinetic energy to electricity. E3S Web of Conferences 80.

Table Tennis Using Arduino For Seniors' Healthcare

Reem Ali Abdullah, Mastaneh Mokayef, Miad Mokayef, Sew Sun Tiang, Wei Hong Lim
Faculty of Engineering, UCSI University
Kuala Lumpur, Malaysia
E-mail : mastaneh@ucsiuniversity.edu.my, tiangss@ucsiuniversity.edu.my. limwh@ucsiuniversity.edu.my

Abstract
Virtual reality (VR) has been shown to function well as an assistive technology to physical therapy for senior users. Seniors, from retirement home residents, form a unique user group in this field, due to their characteristics and demands. A VR controller was implemented using Arduino, MPU9565 and OpenCV to control a Table Tennis game made on Unity3D. The satisfactory results have been obtained in terms of tracking accuracy.
Keywords: Virtual Reality; Seniors; Arduino; mpu9265; OpenCV.

1. Introduction

The ever fast increasing of the elderly population, will lead to the increment of the number of mental and physical disordered people. Many seniors with lack of social environment tend to decline in terms of phycological and physical health. It can be explained when seniors do not have any friends or relatives around, and it can cause severe boredom [1]. Intense severity of boredom will most certainly cause lack of motivation to keep up with general activities for health maintaining. However, the only possible solution for people in such age is to provide them with entertainment and social interaction. Hence, if social aspect of person's life is limited due to various circumstances, the entertainment would be then a promising solution for such problem. Now a days the wide range of digital products can help many people to experience unusual sensations. Video gaming industry is one of the brightest examples that is very developed due to its high demands and it is still under develop to bring even brighter experience in a near future. While younger generation has no difficulty with gaming experience, the same cannot be said for the elder generation as they are not so familiar with the technology. One of the ways to overcome this problem is to bring the perception of such technology closer to the perspective; which in this project, this has been done by implementing the VR concept to games that represents the real-life sport activities to the elderlies. However, economically, VR devices are not the easiest approach to provide the new experience though there are alternative devices and controllers provided by Arduino that can be made to lower down the price of such entertainments. On the other hand, earlier research in [2] illustrates the succeed of utilizing the VR in improving the balance and mobility of elderlies and the results show the improvement in the reaction time on seniors, in other word, seniors reaction time decreased after VR training. Moreover, research [3,4] proved that the table tennis is a "reaction/speed game "with anaerobic motions. Hence, this work has been done by combining the VR and table tennis game ideas and resulted to the proposed VR-based table tennis game for seniors' healthcare.

2. Benefits of virtual reality for seniors

Using VR through active video games is increasingly being used as a supplementary instrument in health rehabilitation. Extensive research has been conducted into the efficacy and feasibility of virtual reality and gaming systems for home use by seniors [1].A major health concern is falling in the home for older adults, which has serious implications, including fear of leaving home, which contributes significantly to social isolation [2]. Virtual Reality can help train motor and cognitive skills of seniors. Research have been done and showed that the elderly can not only improve but build new motor and mental skills even at an old age. It can lead to improvement in balance, gait, and neural connections. The use of VR technology also helps in chronic pain management and distraction therapy, when going into an immersive world the brain is distracted from pain without medication.

VR can be use in reminiscence therapy for patients with dementia, by using nostalgic feelings from photos or objects as a therapeutic tool. Nostalgia produces positive emotions, reduces stress, and can strengthen cognitive memory components. VR can help to make the real feeling in their experience, imitating the feeling of "being there." It triggers therapeutic memories more effectively than an old photograph by providing highly immersive nostalgic experiences. [3]

3. Components and software

3.1 Arduino Uno

The Arduino development board is the motherboard for development of systems such as Uno, Leonardo, Due, Mega, and Nano. Consider the Arduino Uno development board as an example. It includes an ATmega328 microcontroller chip, a USB port, and an input/output pin. The ATmega328 microcontroller chip is the main processing chip. The USB port is a port through which the power supply is provided, and the developed PC program is uploaded to the development board. The input/output pin can be connected to the required sensor. Owing to its affordability and convenient development, Arduino Uno is one of the most popular development boards.

3.2 MPU-9265

The MPU9265 is a motion tracking device or MEMS (Micro Electromechanical System). A SIP (System in a Package) combining two components the MPU6500 & AK8963. The MPU6500 which is contains 3 axis gyroscope, 3 axis accelerometer, and onboard digital motion processor (DMP) capable of processing complex sensor fusion algorithm. 2018.*14Core.*

AK8963 is 3-AXIS electronic compass integrated circuit (IC) with highly sensitive hall sensor technology on a small package. The AK8963 integrates magnetic sensor for detecting terrestrial magnetism in the X-AXIS, Y-AXIS, and Z-AXIS, a sensor driving circuit, signal amplifier chain, and arithmetic circuit for processing the signal from each sensor, it has self-test function integrated. This device is suitable for navigation applications and gaming joystick.

IMU sensors usually consist of two or more parts. Listing them by priority, they are the accelerometer, gyroscope, magnetometer, and altimeter. The MPU 9265 is a 9 DOF (degrees of freedom) or a nine-axis IMU sensor, which means that it gives nine values as output: three values from the accelerometer and three from the gyroscope and three from magnetometer. This chip uses I2C (inter-integrated circuit) protocol for communication. [4]

Serial plotter on Arduino was used to check the x,y,z of the accelerometer, gyroscope and magnetometer. Every time the mpu-9265 sensor moves it tracks the movement and displays it on the serial plotter.

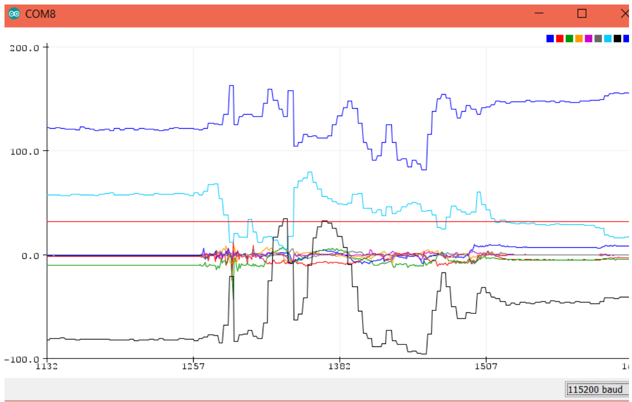


Figure 1 Serial plotter results

3.3 UNITY 3D and Arduino Communication

The table tennis game was made using UNITY3D to portray a first-person point of view of a table tennis racket. Visual aspects of the game do not play an important role, however, it is desired to have simple graphics in order to make it with very low spec requirements. Additionally, the position of camera in-game is aimed towards first-person mode though at current stage it is maintained at third person mode due to compatibility problems. In order to make the game respond to the controller, a communication between unity and Arduino needs to be established, this can be done by using Uduino.

Uduino is a Unity plugin simplifying communication between Arduino and Unity. It has all the mapping features most developers need for their interactive installations. It makes serial communication between unity and Arduino easier, it automatically detects the Arduino board once the library and code is included in Arduino ide, the way that this communication works is that Arduino sense serial print and prints the data from the IMU sensor, when its printed a separating character and the values, in unity it reads and uses the path command ,

The IMU sensor is using quaternions for the rotation instead of Euler because Euler Angles are limited by a phenomenon called "gimbal lock," which prevents them from measuring orientation when the pitch angle approaches +/- 90 degrees. Quaternions provide an alternative measurement technique that does not suffer from gimbal lock. The script used for the mpu-9265 uses the Uduino library to detect which port is the Arduino connected to without inputting it manually inside the code.

3.4 OpenCV for object recognition

OpenCV was used in Python to detect a circle-shaped blue object in a webcam stream. UDP server was used to send the data from python to unity. This enable the camera to track the movement of the object specified when it goes to x-axis, y-axis and z-axis direction. To find the direction of z-axis, the camera detects the radius of the circle. The bigger the circle radius the close it goes to the z-axis. The port has to be set up according to the camera connected, since it was an external camera, the port was 4747, To call the camera for video capturing, a function cv.videocapture( ) from the OpenCV library is used with specifying the address of the camera, in this case its 0 since the PC has no default camera. Figure 1 shows the detection of the blue circular object, the yellow circle line indicates the object is beginning successfully detected.

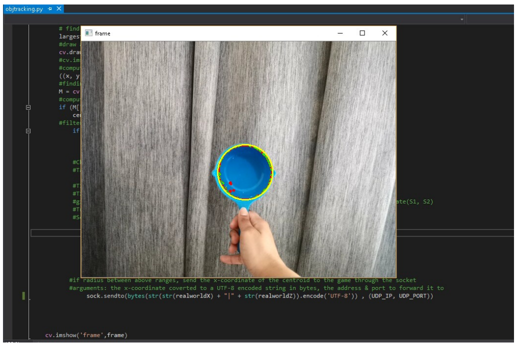


Figure 2 Object detection using OpenCV

4. Results

4.1 Project fabrication

The MPU 9265 communicates with the Arduino through the I2C protocol. The MPU 9265 is connected to Arduino as shown in the following Figure 3.. The 3.3V of the IMU sensor is connected to the 3.3V pin of the Arduino. Next, the GND of the Arduino is connected to the GND of the MPU 9265. To set up the I2C lines,the SDA pin on the MPU 9265 is connected to the Arduino's analog pin 4 (SDA). Tthe SCL pin on the IMU sensor is connected to the Arduino's analog pin 5 (SCL).

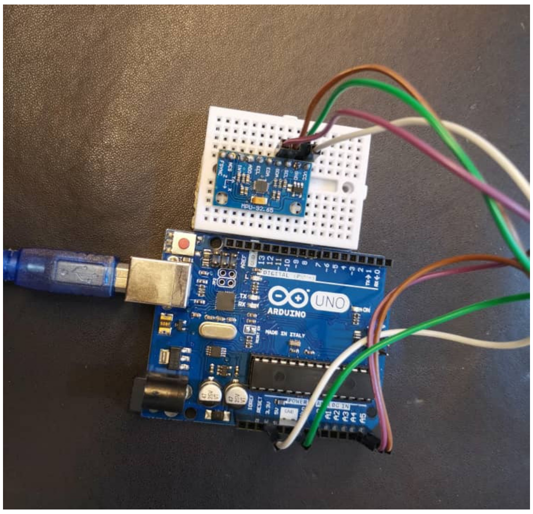


Figure 3 Hardware connection

4.2 Table tennis game

The game has been made using unity3D.It is a simple first-person single player game. The blue racket represents the Artificial intelligence (AI) player. It is programmed to track the ball; and it will follow the same direction as the ball. there is a time factor thus it is possible to outplay the AI player. Both user racket and AI racket are locked in to a specific x,y,z position to avoid going into the wrong direction. If the user exceeds the limit of the set position, it will stop at the specified point. Script of AI player was made using unity and visual studio and was uploaded to the inspector tab of the AI racket. The user player contains both scripts of the OpenCV and the MPU-9265.

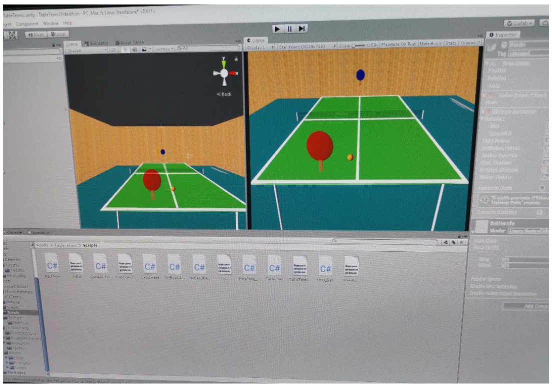


Figure 4 Table tennis on unity

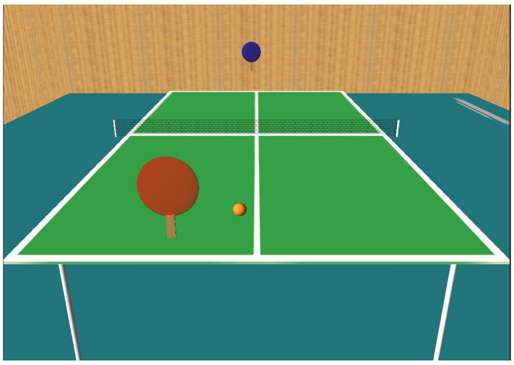


Figure 5 Final game

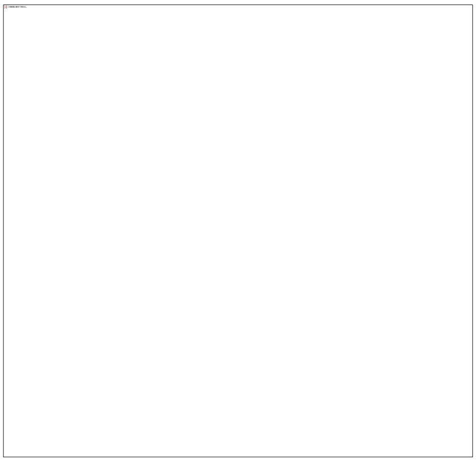


Figure 7 Preview of sent data

1. Testing of the table tennis game

The testing has been done using the MPU-9265 to track the rotation angle of the object. The camera tracking has been used to detect the direction of movement. Although, there is a slight delay in the direction detection, the rotation according to hand movement was captured in the game. In Figure 5, the detection of the circle is shown. The socket library from python is used to specify the IP address and port of the camera to send data from python to unity. The data sent is in bits which then converted to string values to be read by unity. The data from the camera port is sent to the PC local IP address. The unity then receives through this IP address. The data is then extracted from the code and it will try to identify any blue circular object to capture its motion.

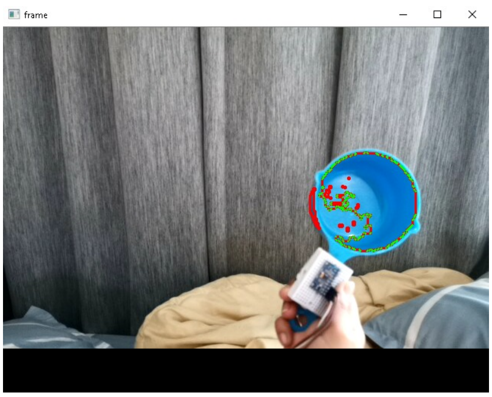


Figure 6 Direction detection of object and IMU

2. Limitations

This project only uses one camera and it does not support a 360 view, there are some errors in detecting the object when there is a fast movement or noise around it. This can cause inaccuracy in the game when the movement of the object is undetectable.

5. Conclusion

In this paper, a controller using Arduino and OpenCV motion capturing was used to control a 3D game made with unity, and how it was implemented to portray a virtual experience of a table tennis game. It also discussed how can sports and virtual reality combined can help seniors to be in a better health.

References

1. Miller, K.; Brooke, A.; Pearce, A.; Said, C.; Morris, M.,2014. Effectiveness and feasibility of virtual reality and gaming system use at home by older adults for enabling physical activity to improve health-related domains: A systematic review. Age Ageing, 43, 188–195.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
2. E. Bisson, B. Contant, H. Sveistrup, and Dr. Y. Lajoie. Functional Balance and Dual-Task Reaction Times in Older Adults Are Improved by Virtual Reality and Biofeedback Training. CyberPsychology & BehaviorVol. 10, No. 1, 17 Feb 2007.

3. Miad Mokayef, Mehrzad Moghadasi, Reza Nuri. Effect of cold water immersion on blood lactate levels of table tennis players. Volume 2 Number 9 (September-2014) pp. 115-123
4. Miad Mokayef, Parisa Shahini. Comparison of contrast water immersion, active recovery and passive recovery on blood lactate and CRP levels in table tennis players. Journal of Physical Activity and Hormones Vol 1, No. 2, Ser. 2 (June 2017), 017-028
5. Bruun-Pederson, J.R.; Pederson, K.S.; Serafin, S.; Kofoed, L.B., 2014. Augmented exercise biking with virtual environments for elderly users—A preliminary study for retirement home physical therapy. In Proceedings of the VR 2014—Workshop on Virtual and Augmented Assistive Technology (VAAT), Minneapolis, MN, USA.K.
6. Molina, N.; Ritchie, J. Virtual reality using games for improving physical functioning in older adults: A systematic review. J. Neuroeng. Rehabil. 2014, 11, 156.
7. Rendon, A.; Lohman, E.; Thorpe, D.; Johnson, E.; Medina, E.; Bradley, B., 2012. The effect of virtual reality gaming on dynamic balance in older adults. Age Ageing
8. Baecker, R.M.; Moffatt, K.; Massimi, M.,2012. Technologies for aging gracefully. Interactions, 19, 32–36.
9. Kim, A.; Darakjian, N.; Finley, J., 2017.Walking in fully immersive virtual environments: An evaluation of potential adverse effects in older adults and individuals with Parkinson’s disease. J. Neuroeng. Rehabil.
10. Brent (n.d.). MPU-9250 Hookup Guide.[Online] Available from https://learn.sparkfun.com/tutorials/mpu-9250-hookup-guide/all [Accessed 21 March 2019].

A Study on Prevention of Predictive Failures

Tsutomu Ito, Hiroshi Sakai
Hamura Plant, Hino Motors, Ltd.,
5-2-1 Midorigaoka, Hamura City, Tokyo, 205-8601, Japan
E-mail: fw.eldorado.500cuin@gmail.com

Takao Ito
Graduate School of Engineering, Hiroshima University
1-4-1 Kagamiyama, Higashi Hiroshima, 739-7527
E-mail: itotakao@hiroshima-u.ac.jp

Seigo Matsuno
Department of Business and Administration, NIT, Ube College,
2-14-1, Tokiwadai, Ube, Yamaguchi, 755-0096, Japan
E-mail: matsuno@ube-k.qac.jp

Makoto Sakamoto
Faculty of Engineering, University of Miyazaki,
1-1 Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan
E-mail: sakamoto@cc.miyazaki-u.ac.jp

Abstract

Electric motors play an important role as the primary power for conveyor and drop lifts in manufacturing systems especially in automobile manufacturing factories. Many electric motors equipped with brake systems have sensors that determine where they should stop. This brake system needs to be maintained to keep proper brake gap consistent. Measuring the brake gap of electric motors is one of the main jobs of maintenance department, but this is a difficult task because most of the motors are located in unsafe spaces, like the upper end or under the lift pit . This paper proposes a new approach to measure the brake gap of motors to reduce maintenance risks, thus saving costs. The method suggests using vibration acceleration to measure the motor gap in running production system based on mechanical model. The effectiveness of this research is indicative by experience results.

Keywords: motor brake, brake gap, acceleration, maintenance.

1. Introduction

As the main power source of conveyors and lifts in manufacturing systems, electric motors play an important role in automobile assembling lines as well as other machines. Known as motors with brakes inbuilt, their function is to improve the accuracy of machinery stopping when required. Most of these mechanisms have brake linings, which wear out over long-term use. To maintain the braking function, the necessitates measuring and tuning the brake. The dominant approach of measuring the brake gap is usually reassembling the cover and measure the gap directly by maintenance personnel using thickness gage. From viewpoint of securing the working place, most of the electric motors are housed in the pit (under-ground) or above the machinery. Thus, it makes maintenance process dangerous and inconvenient. To solve this problem, manuscript proposes a new remote approach to measure the brake gap without reassembling the motor based on mechanical model.

This paper is structured as follows. Section 2 reviews the relevant literature on brake measurement. Section 3 explains the structure of motor brake. Section 4 shows calculation and discussions. Section 5 identifies experiment apparatus and condition. Experimental results and discussion are provided in Section 6. Section 7 explains directions for future research.

2. Background

Motor braking tests is a primary determinant of motor performance, which has been the subject of many investigations. For example, Bhau, Patil and Samant presented methodologies used for automotive service brake testing for two wheelers. The main contribution of their study was to compare and contrast three main brake testing standards, viz., Indian Standards, Federal Motor Vehicle Safety Standards and European Economic Commission Standards [1]. Seyfert, in his study, introduced technical tips about brake maintenance [2]. Sangtarash et al. explained simulation procedures for series braking with optimal braking feel, series braking with optimal energy recovery, and parallel braking strategies using CRUISE software [3]. All these studies focused on different aspects of motor braking systems, but studies on industrial motor maintenance have been sparse. In assembly line, many electric motors equipped brake system are required to determine where they should stop. This brake system need to be maintained to keep proper brake gap. Measuring the brake gap of electric motors is one a key job of the maintenance department, but entail significant difficulties because most of the motors are housed in unsafe spaces, such as at the upper end or even under the lift pit. In this study, the authors propose a new approach to measure the brake gap of motors to reduce maintenance risks and save costs. The method is using vibration acceleration to measure the motor gap in running production system based on mechanical model. To our best knowledge, this is the study that measures the brake gap without machinery stopping.

3. Structure of motor brake

Major motor brake consist of a brake lining, friction plate and actuator. An example of motor brake is shown in Fig.1.

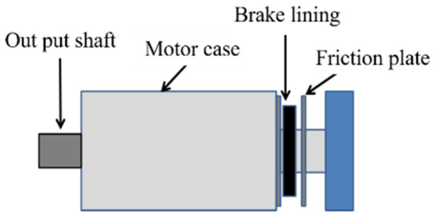


Fig. 1. A Typical image of Motor with Brake.

The brake lining is attached to main shaft of motor. Brake lining spins with main shaft of motor freely when brake is free. If brake is activated, friction plate pushes the brake ling by actuator. Typical type of actuator uses coil spring to push the friction plate, and electrical magnet to release brake. Fig.2 and Fig.3 show an image of brake free and brake locking relationship between brake lining and friction plate respectively.

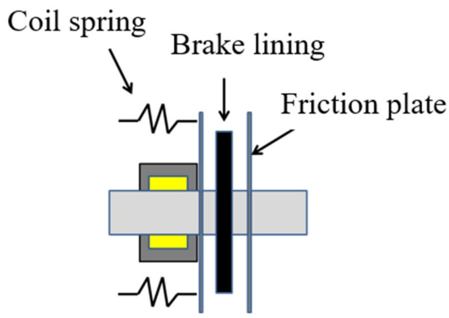


Fig. 2. An Image of Free Position Relationship between Brake Lining and Friction Plate.

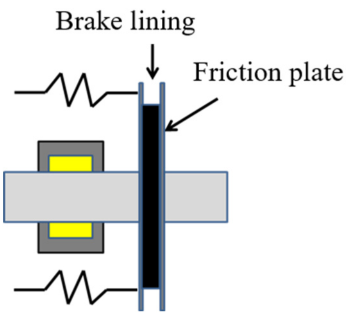


Fig. 3. An Image of Locking Position Relationship between Brake Lining and Friction Plate.

4. Calculation and Discussions

Motor brake is activated by coil spring in actuator to push the friction plate. The structure of our models is shown in Fig.2 and Fig.3, respectively.

On real motors, some coil springs are used, but they can be constructed as one modelled spring with spring rate k[N/m], m[kg] for math of friction plate and brake lining, and Δx[m] for brake gap. The energy this system dispenses is formulated in equation (1).

$$\frac{1}{2}k\Delta x^2 = \frac{1}{2}mv^2 \quad (1)$$

$$\mathrm{v} = \sqrt{\frac{k}{m}}\Delta x \quad (2)$$

This energy system vibrates the motor when the brake is activated. Equation (2) shows the relationship of Δx and v[m/s]. Here v[m/s] indicates the instant velocity when the friction plate hits the brake lining. But this velocity is difficult to measure. However, based on momentum conservation law, the relationships of acceleration to the brake gap is calculated in the equations below.

$$\mathrm{m}v' - mv = F\Delta t \quad (3)$$

$$v' = 0$$

$$\mathrm{F} = \frac{m}{\Delta t}\sqrt{\frac{k}{m}}x \quad (4)$$

$$\ddot{x} = \frac{1}{\Delta t}\sqrt{\frac{k}{m}}x \quad (5)$$

$$\ddot{x} = \beta x \quad (6)$$

Δt signifies a very small amount of time, so it is hard to measure. But by transforming it as equation (6), the linear relationship between brake gap and acceleration becomes evident. This suggests that measure the vibration acceleration can be used for brake gap.

5. Experiment Apparatus and Condition

Experiment apparatus and conditions are shown below, respectively.

5.1. *Experiment Apparatus*

A set of experimental apparatus is shown in Table 1.

Table. 1 Experimental apparatus

	Maker	Type
Motor	SUMITOMO HEAVY INDUSTRES GEARMOTORS	INDUCTION MOTOR 0.2kw 4P 3φ
Data logger	KEYENCE	NR-600
Amplifire of acceleration pick up	KEYENCE	NR-CA04
	KEYENCE	NR-U65
Data logger software	KEYENCE	Wave Logger pro
Acceleration pick up	SHOWA SOKKI	2304A , Sensitivity 47.6pC/G , 4.86pC/(m/s2) Capacitance 780pF
Shim tape	TRUSCO	TFG-0.05M1

5.2. *Experiment Condition*

The authors attached the acceleration pick up to motors fitted with magnets. First, an experiment was conducted to confirm the effect of the presence (or absence) of a fan cover on the measurement result of acceleration. From a practical viewpoint, this is expected if it is possible that the fan cover attached is able to be used. Second, an experiment to find out the linearity of vibration acceleration and brake gap was conducted. To demonstrate the several gap, the authors used shim tape to make the clearance between bake lining and friction plate. (A typical picture is shown in Fig.4.)

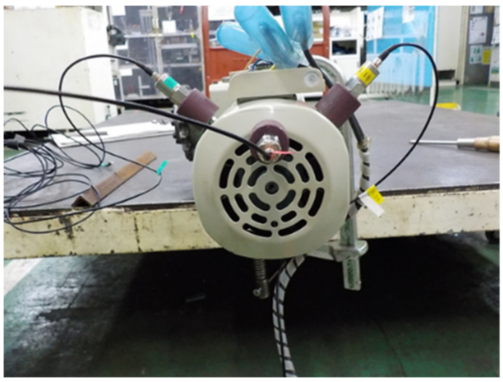


Fig. 4. Experimental Condition.

Vibration acceleration data were collected as plus acceleration and minus acceleration, thus the absolute number and average of 6 times of vibration—as is the norm—was employed. Fig. 5 shows an image for data collection.

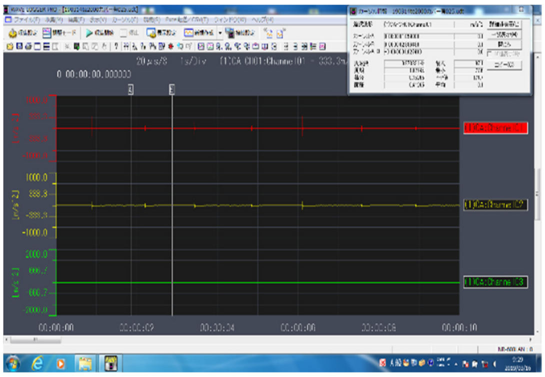


Fig. 5. An Image for Data Collection.

6. Result and Discussion

6.1. *Comparative study with Experiment with Cover*

To clarify the effect of the presence (or absence) of a fan cover on the measurement result of acceleration, two experiments were conducted. The first measured the acceleration with fan cover and the second experiment was conducted without the fan cover. Fig.6 shows the result of these experiments.

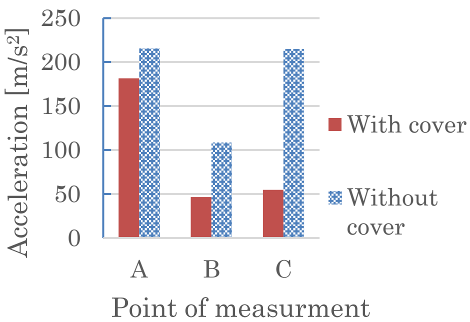


Fig. 6. Difference of Acceleration between Experiments with Cover and without Cover.

Measuring points, A and B are circumferential directions, whereas C is axial direction. It is expected that measuring acceleration without the cover would be higher than that with the cover. Thus, it is better to measure the acceleration without cover. On the other hand, a hard point of the motors case transmit vibration directly to the outside of the motor case. Result of point A shows the possibility of measuring vibration without reassembly the fan cover.

6.2. *Relationship of Acceleration and Brake Gap*

To illustrate the linear relationship between measured acceleration and brake gap, an experiment with brake gap between 0.26[mm] to 0.76[mm], difference of 0.05[mm] conducted. Based on the findings shown in Fig.7, a linear relationship between acceleration and quality is observed.

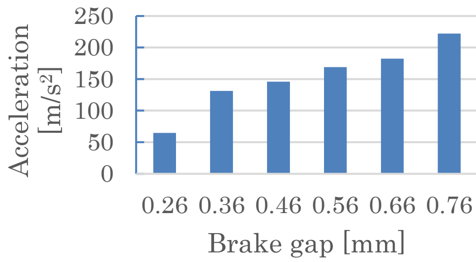


Fig. 7. Relationship between Acceleration and Brake gap.

Calculated coefficients of k [N/mm] for each gap are reported in Table 2.

Table. 2 Coefficient K[N/m] of each Brake Gap

Gap[mm]	0.26	0.36	0.46	0.56	0.66	0.76	Average
K[N/m]	248.5	364.3	317.5	301.7	276.3	292.2	300.1

The relationship between experimental result and calculated with the coefficient k [N/m] is shown in Fig. 8.

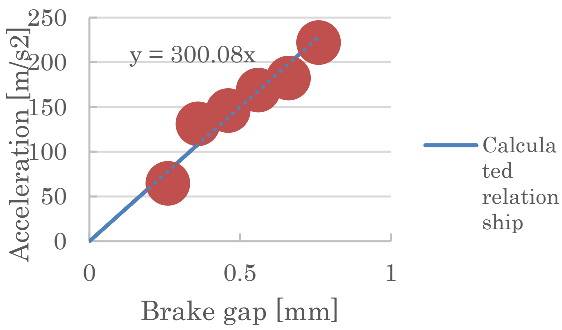


Fig. 8. Relationship between Theoretical Result and Experiment Result.

Additionally, a t-test was added to the experimental result and shows a statistically significant difference of the acceleration between 0.2[mm] of brake gap.

7. Conclusion

In this study, a new approach to measure the brake gap was proposed. According to the results of the experimental data analysis, the method to measure brake gap with accuracy 0.2mm was observed. The implications are that electrical measurement systems should be operated among IoT system in the future.

References

1. Bhau Kashinath Kumbhar, Satyajit Ramchandra Patil, Suresh Maruti Sawant: A Comparative Study on Automotive Brake Testing Standards, Journal of The Institution of Engineers (India): Series C, Volume 98, Issue 4, pp 527–531, August 2017, https://link.springer.com/article/10.1007/s40032-016-0289-y retrieved November 17, 2019.
2. Karl Seyfert, Preventive Maintenance Tech Tips: Brake Maintenance, https://www.motor.com/magazine-summary/preventive-maintenance-tech-tips-brake-maintenance-march-2004/ retrieved November 17, 2019.
3. Farhad Sangtarash, Vahid Esfahanian, Hassan Nehzati, Samaneh Haddadi, Meisam Amiri Bavanpour and Babak Haghpanah, Effect of Different Regenerative Braking Strategies on Braking Performance and Fuel Economy in a Hybrid Electric Bus Employing CRUISE Vehicle Simulation, www1.coe.neu.edu/~bhaghpan/papers/Effect%20of%20Different%20Regenerative%20Braking%20Strategies%20on%20Braking%20Performance%20and%20Fuel%20Economy%20in%20a%20Hybrid%20Electric%20Bus%20Employing%20CRUISE%20Vehicle%20Simulation.pdf retrieved November 17, 2019.

Did a mismatch between the ASBJ's standard-setting and its organizational structure occur?

Kensuke Ogata
Graduate School of Business, Osaka City University, 3-3-138 Sugimoto, Sumiyoshi, Osaka, Osaka, 558-8585, Japan

Abstract

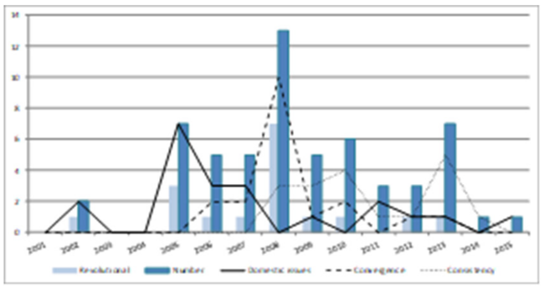
This paper examines the relationship between standard development and organizational structure of the ASBJ, Japanese accounting standard setter. According to the previous literature, the standard setter (re)structured its organization in order to execute strategies and goals that it sets by itself or is expected to achieve. In fact, the ASBJ slowed down the development of revolutionary and improved accounting standards during the period of 2009-2012. However, the ASBJ formed the organization to move ahead on developing such standards, that is, the accounting-profession-centric organization, based on organization analysis in this paper. This paper indicates that an exogeneous shock changed interests in some domestic actors, and consequently caused this mismatch.

Keywords: standard-setting, ASBJ, accounting regulation, network theory, profession

1. Introduction

Now, International Financial Reporting Standards (IFRSs) produced by the IASB have become the focus of attention in global financial community. Some countries introduced IFRSs as domestic or regional financial reporting standards; and others have converged domestic standards with IFRSs to invite foreign investment. In the latter case, national standard setters (NSSs) have to fulfill very difficult role in developing domestic standards, in terms both of balancing the conflicting interests of diversified domestic stakeholders and of achieving harmonization with international standards.

However, to which aspect a specific NSS attach weight can vary at different periods. According to Graph 1 which shows standard development activities of Japanese standard setter, Accounting Standards Board of Japan (ASBJ), from 2001 to 2015, the ASBJ attached importance to settlement of domestic problems in the initial four years; it spent much efforts into the global convergence of Japanese standards from 2005 to 2008. During the next four years from 2009 to 2012, the ASBJ slowed down in terms of the quantitative aspect, and stopped developing standards for the convergence.



Graph 1. Standard Setting of ASBJ from 2001 to 2015

Previous literature mentioned that changes in standard development approach could be caused by reorganization of and changes in composition of the standard setters (Refs.1-3). Following this, it is supposed that the ASBJ reorganized or changed its membership composition to alter its standard development approach. Accordingly, this paper aims to clarify under what organizational structure, the ASBJ changed to the standard development approach that it slowed down the development pace to stop setting standards for the convergence. In addition, this paper illustrates organizational structure of ASBJ using Social Network Analysis (SNA).

2. Studies on Accounting Standard Setting

2.1. *Previous Literatures*

It is generally said that processes of accounting standard development are political forum for various stakeholders who can be economically influenced by newly-creations, revisions or removals of any specific standards (Ref.3). Most studies on the standard development processes have been conducted on empirical basis. These studies however have been criticized as not having a full picture of the process and not explain the dynamics of power, interests and resources in the process (Ref.1). Recently, several interesting studies focusing on intra-board political activities began to appear (Refs.4-6).

On the other hand, some studies on accounting regulation illustrated that standard setters reorganized their structure through changes in the composition of board members, to pursue specific strategies and aims (Refs.2,3,6, and 7). In line with that, the relationship between board members' professional and political characteristics and properties of accounting standards was examined (Ref.8). Parts of the above studies incorporated social network analysis (SNA) as organizational analysis to define organizational feature of standard setters (Refs.6,7,9 and 10).

2.2. *Clarifying Organizational Structure Using SNA*

Fundamental feature of SNA is to quantify relationships among actors in network. Another feature is to describe the network structure using a graph in which the relationship is drawn with nodes and edges (Ref.11). Thus, this paper characterizes organizational structure of ASBJ using SNA.

To capture the network structure of ASBJ, this paper adopts a continuous coreness analysis. This analytical method is used to identify a set of actors who have a high density of ties among themselves (the core) by having many events in common and another set of actors who have a low density of ties among themselves (the periphery) by having few events (Ref.12). Herein as "common" events, this paper adopts organizations which provided their employees as members to the ASBJ and the associated institutions. I suppose that their backgrounds or careers could determine their underlying ideas, minds and preferences on accounting standards. Organizational analysis using background-data on members of the ASBJ and the associated institutions can highlight dominant actors in the standard-setting bodies in terms of human resources. As a result, I give a full account of which actors dominate and what values are predominant in the ASBJ standard-settings.

3. Organizational Structure of the ASBJ

3.1. *Using Data*

This paper uses matrix data-sets, composed of the organizations which provided their employees to intra-bodies of Financial Accounting Standards Foundation (FASF) from 2009 to 2012 in rows and these intra-bodies of the FASF by year in columns, including Board of Directors, Board of Councillors, ASBJ and Standards Advisory Council. It indicates that the larger numbers and the longer the organizations provided, the more dominant they were in the FASF.

As the next step, I transformed these data-sets into organization-to-organization data-sets through affiliation process. Based on the affiliated data-sets, this paper performed coreness analyses, contained in SNA software, *Ucinet VI* (Ref.13) and drew graphs using *NetDraw* software (Ref.14).

A matrix data-set for the period of 2009-2012 comprises 68 organizations. Another data-set for the period of 2005-2008 to relativize the organization during the period of 2009-2012 includes 77 organizations.

3.2. *The Characteristics of Organizational Structure of the ASBJ in the period of 2009-2012*

3.2.1 Coreness Result

Coreness scores of Top 16 organizations during the period of 2009-2012 shows in Table 1. According to this table, the top five organizations' scores left others' ones far behind. Included in the top five were: Tohmatsu (coreness: 0.505, rank: #1), FASF (0.470, #2), Azusa (0.433, #3), ShinNihon (0.413, #4), and Arata (0.249, #5). The above four organizations, except for the FASF, were all large accounting firms in Japan that respectively have formed alliances with large global accounting firms. The following second group was preparers of financial accounts, especially non-financial sectors, including Central JR (0.096, #8), Tokyo Electric Power (0.074, #12), NTT Docomo (0.071, #13), Fujitsu (0.071, #13), and Sumitomo Electric (0.051, #15). The third group was financial sectors, such as Nomura Securities (0.103, #7), Mitsui Sumitomo Insurance (0.096, #8), and Meiji Yasuda Life (0.096, #8).

Table 1. Coreness of Top 15 Organizations from 2009 to 2012

#	Organization	Attribute	Coreness
1	Tohmatsu	Pro	0.505
2	FASF	Reg	0.470
3	Azusa	Pro	0.433
4	ShinNihon	Pro	0.413
5	Arata	Pro	0.249
6	Tokyo Stock Exchange	User	0.145
7	Nomura Securities	Fin	0.103
8	Central JR	NonFin	0.096
8	Bank of Japan	Reg	0.096
8	Mitsui Sumitomo Insurance	Fin	0.096
8	Meiji Yasuda Life	Fin	0.096
12	Tokyo Electric Power	NonFin	0.074
13	NTT Docomo	NonFin	0.071
13	Fujitsu	NonFin	0.071
15	Sumitomo Electric	NonFin	0.051
15	Tohmatsu Tax	Pro	0.051

In order to relativize the organizational structure of ASBJ in that period, I show the top 15 high-coreness organizations for the period of 2005-2008, as shown in Table 2. In this table, the following four organizations got strikingly-high scores: Azusa (coreness: 0.545, rank: #1), Tohmatsu (0.444, #2), ShinNihon (0.410, #3), and FASF (0.373, #4). The following second group was preparers of financial statements, such as, Tokyo Electric Power (0.139, #6), Central JR (0.116, #10), and NEC (0.116, #10); and the third group was financial sector actors, such as, Nomura Securities (0.127, #7), Mitsui Sumitomo Insurance (0.121, #8), and Meiji Yasuda Life (0.120, #9).

Table 2. Coreness of Top 15 Organizations from 2005 to 2008

#	Organization	Attribute	Coreness
1	Azusa	Pro	0.545
2	Tohmatsu	Pro	0.444
3	ShinNihon	Pro	0.410
4	FASF	Reg	0.373
5	Tokyo Stock Exchange	User	0.155
6	Tokyo Electric Power	NonFin	0.139
7	Nomura Securities	Fin	0.127
8	Mitsui Sumitomo Insurance	Fin	0.121
9	Meiji Yasuda Life	Fin	0.120
10	Central JR	NonFin	0.116
10	Bank of Japan	Reg	0.116
10	NEC	NonFin	0.116
10	ARC	Pro	0.116
14	Bank of Tokyo-Mitsubishi UFJ	Fin	0.085
15	Arata	Pro	0.084

Comparing to these two organizations, I present the ratios of each total coreness score acquired by attribute group to the whole score in the network and the difference, as shown in Table 3. Attribute groups are divided into seven groups as follows: academic community (Aca); financial community (Fin); non-financial community (Non-Fin); accounting professions (Pro); regulators (Reg); users of financial statements (User); and others (Other).

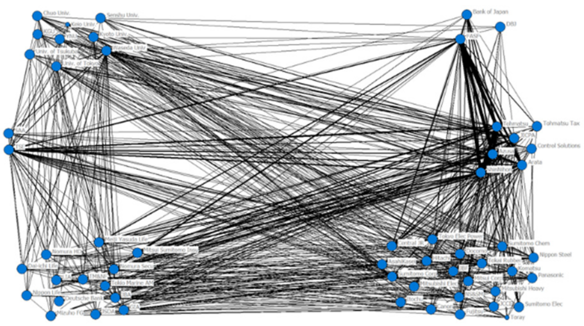
Table 3. Difference of Composition Ration between the Two Periods by Each Attribute Group

	2005-2008 (1)	2009-2012 (2)	Difference (2)-(1)
Aca	5.2%	4.1%	-1.1%
Fin	16.2%	13.4%	-2.8%
Non-Fin	18.1%	16.9%	-1.2%
Pro	43.5%	45.8%	2.2%
Reg	11.4%	15.2%	3.8%
User	4.7%	4.5%	-0.2%
Other	0.9%	0.2%	-0.7%
Total	100.0%	100.0%	0.0%

With regard to this result, I have two points to be referred. The one is that the ASBJ compositions between the two periods had almost the same by attribute group. The other is that accounting professions and regulators slightly increased the ratios. In sum, the ASBJ continuously formed accounting-profession-centric organizations in recent periods.

3.2.2. Graph Representation

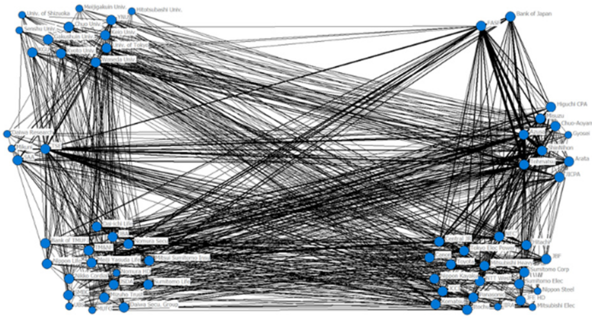
Next, I show a network graph of the ASBJ from 2009 to 2012 in Graph 2. It implies that the stronger the relationship is, the wider the line is drawn using a ten-scale level. Also, this graph plots regulators on the top right; accounting professions on the right centre; non-financial businesses on the bottom right; financial businesses on the bottom left; user actors on the left centre; and academics on the top left. According to Graph



Graph 2. Network Graph of ASBJ in the Period of 2009-2012

2, Japanese Big 4 accounting firms and the FASF had stronger connections than others.

To extract organizational characteristics of the ASBJ for the period of 2009-2012, I make another graph for the period of 2005-2008 (*see* Graph 3). Comparison between both graphs indicated that ASBJ remained forming the organization which put accounting professions into core positions, as in the last period.



Graph 3. Network Graph of ASBJ in the Period of 2005-2008

4. Conclusions

According to coreness analyses and graph presentation, the ASBJ remained the same organization structure from the last period, that is, accounting-profession-centric organization. This indicates that the ASBJ employed standard development approach which did not seek globalization of Japanese standards under the structure which was best suited to facilitate the global convergence. Why did this mismatch happen?

In considering facts that the ASBJ expressed the strategy for accelerating the convergence with IFRSs earlier in the period of 2009-2012. To do so, the ASBJ was likely to construct accounting-profession-centric organization to make use of not only the expertise of accounting professions, but the experience on implementation of global standards through their global networks, as in the period of 2005-2008. However, due to the turnaround of the US Securities and Exchange Commission in the convergence program, the Japanese government and the ASBJ came less need for pursuing the convergence emergently. In this regard, it could be interpreted that this mismatch occurred due to the temporary gap between the rapid change of standard development policy or approach and organizational structure which cannot be changed instantly.

References

1. R. G. B. Walker and P. Robinson, A Critical Assessment of the Literature on Political Activity and Accounting Regulation, *Research in Accounting Regulation*, 7, 1993, pp.3-40.
2. W. Mattli and T. Büthe, Accountability in Accounting? The Politics of Private Rule-Making in the Public Interest," *Governance: An International Journal of Policy, Administration, and Institutions*, 18(3), 2005, pp.399-429.
3. A. S. Zeff, The Evolution of U. S. GAAP, The Political Forces behind Professional Standards, Part II, *The CPA Journal*, 75(2), 2005, pp.19-29R.
4. M. Klein and R. U. Fülbier, Inside the Black Box of IASB Standard Setting: Evidence from Board Meeting Audio Playbacks on the Amendment of IAS 19 (2011), *Accounting in Europe*, 15, 2018, pp.1-43.
5. J. Morley, Internal Lobbying at the IASB, *Journal of Accounting and Public Policy*, 35, 2016, pp.224-255.
6. J. Perry and A. Nöelke, International Accounting Standards Setting: A Network Approach, *Business and Politics*, 7(3.5), 2005. pp.1-32.
7. S. Botzem, *The Politics of Accounting Regulation, Organizing Transnational Standard Setting in Financial Reporting*, 2012, Edward Elgar Publishing.
8. A. M. Allen and K. Ramanna, Towards an Understanding of the Role of Standard Setters in Standard Setting, *Journal of Accounting and Economics*, 55(1), 2013, pp.66-90.
9. A. J. Richardson, Regulatory Networks for Accounting and Auditing Standards: A Social Network Analysis of Canadian and International Standard-Setting, *Accounting, Organizations and Society*, 34, 2009, pp.571-588.
10. K. Ogata, K. A Study of Accounting Standard-Setting Using Graph Theory, *Artificial Life and Robotics*, 15, 2010, pp.279-283.
11. S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*, 1994, Cambridge University Press.
12. S. P. Borgatti and M. G. Everett, Models of Core/Periphery Structures, *Social Networks*, 21, 1999, pp.375-395.
13. S. P. Borgatti, M. G. Everett and L. C. Freeman, L. C. *Ucinet for Windows: Software for Social Network Analysis*, 2002, Harvard, MA: Analytic Technologies.
14. S. P. Borgatti, *Netdraw: Network Visualization*, 2002, Harvard, MA: Analytic Technologies.

Effectiveness of Data Augmentation in Pointer-Generator Model

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: touchi@mei.kpu.ac.jp, tabuse@kpu.ac.jp

Abstract

We propose a new data augmentation method in automatic summarization system, especially Pointer-Generator model. A large corpus is required to create an automatic summarization system using deep learning. However, in the field of natural language processing, especially in the field of automatic summarization, there are not many data sets that are sufficient to train automatic summarization system. Therefore, we propose a new method of data augmentation. We use Pointer-Generator model. First, we determine the importance of each sentence in an article using topic model. In order to extend the data, we remove the least important sentence from an input article and use it as a new article. We examine the effectiveness of our proposed data augmentation method in automatic summarization system.

Keywords: automatic summarization, data augmentation, Pointer-Generator Model.

1. Introduction

In recent years, the amount of information on the net of the world has increased exponentially, and it is expected to become 44 ZB in 2020[1]. Under the circumstance, the technique of automatic summarization becomes indispensable for efficient selection of information. However, in the field of automatic summarization, there are not many data sets that are sufficient to train automatic summarization system. Therefore, we considered whether the technique of data augmentation commonly used in image processing can be applied to natural language processing.

Several studies have applied data augmentation to natural language processing. A method of adding noise to an intermediate vector[2], and a method of data augmentation by generating a dialect sentence correspond to each given standard sentence[3]. However, both methods are adapted to document classification, yet none have been adapted to automatic summarization. The nature of automatic summarization is that there is little influence on the generated summary whether there is redundant part or not. Further, in image processing, data may be expanded by changing the background. This assumes that the background has no direct impact on the target task. Therefore, we focused on change of background which is one of data expansion of image processing, and expanded data by deleting redundant sentences of original article in automatic summary.

Automatic summarization is classified into two approaches. One is an extractive summarization in which words, phrases or sentences of input articles are directly extracted and combined to create a summary. The other is an abstractive summarization in which an input article is temporarily converted to an intermediate vector and then a summary is generated based on the intermediate vector. In extractive summarization, there is a problem of how to connect extracted words, phrases or sentences. And there is also the possibility that in extractive summarization it is unknown what the directive word is pointing at. On the other hand, in abstractive summarization, since it is generated from an internal semantic representation, there is a possibility that it becomes one integrated summary. It also learns grammatical information, so it is easy to connect words and phrases in contrast to an extractive summarization. However, the abstract summarization needs to be learned with a large number of pairs of input articles and generation summaries, which takes time and effort to prepare the corpus.

In the previous research[4], we showed data augmentation is effective in abstractive summary. But in that research, using model is only a union of Encoder-Decoder model and attention mechanism. Therefore, in this study, we confirm the effect of data augmentation using the latest model, the Pointer-Generator model[5]. The method of data augmentation used in this research are described in Section 2. Experiment and evaluation are described in Section 3, discussion and summary are given in Section 4.

2. Data augmentation Method

We measure the importance for each sentence of the input article and use the article generated excluding the lowest important sentence as new data.
Refer to [6] for how to determine importance for each sentence in the input article. First, a topic model is created from the CNN/Daily Mail data set, and the sentence weight is determined for each topic of the input article. Then, the sum of sentence weights for all topics is taken as the final weight of the sentence, and it was taken as the importance of the sentence.

$$b_{t,i} = \frac{W_{t,i}}{\sqrt{N_i}} \tag{1}$$

$$b_i = \sum_t b_{t,i} \tag{2}$$

Firstly the weight $b_{t,i}$ of the sentence i with respect to the topic t is examined. In Equation (1), $b_{t,i}$ multiplies the sum $W_{t,i}$ of an importance of the words constituting the sentence i by the inverse of the square root of the total number N_i of words as a coefficient. This coefficient is for weighting sentences that are not influenced by the sentence length. In Equation (2), we let the sum of $b_{t,i}$ be the final weight of sentence i.

3. Experiment and Evaluation

In this section, we evaluate the effectiveness of the data augmentation proposed in Section 2.

3.1. *Dataset*

We used CNN/DailyMail dataset [7]. This dataset has 287,226 training pairs, 13,368 validation pairs and 11,490 test pairs. In the experiment, 10,000 articles, 45,000 articles, 90,000 articles, 180,000 articles, and 287,226 articles are prepared as training data.

3.2. *Evaluation method*

ROUGE [7] is used as a general evaluation method of automatic summarization, which is evaluated by matching of reference summary and generated summary word. We used ROUGE-1, ROUGE-2, ROUGE-L. ROUGE-1 and ROUGE-2 evaluate unigram and bigram coincidence, respectively. ROUGE-L evaluates matching the longest common subsequence. Each ROUGE evaluation method has F value, recall, and precision. “Recall” means how many words of the generated summary are included in the reference summary, and “precision” means how many words of the reference summary are included in the generated summary. The F value is obtained by the following formula.

$$F = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{3}$$

3.3. *Parameter settings*

The program used in this research uses PyTorch. It has been verified that this program can achieve the same result as [5]. In this program, the number of words used to encode an input article is limited to 400. However, in the case there is no sentence with the lowest importance within 400 words from the beginning when we use the data augmentation method, since the sentence is omitted there is no effect on learning. Therefore, I found the article with the most 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 show 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 2380, 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 the research of [5], after learning with the copy mechanism, it learns with the coverage mechanism. Therefore, we use copy mechanism and coverage mechanism in this research. In this research, the batch size is set to 8. Table 2 shows how many times batch-sized articles is learned.

Table2 the number of iteration times each articles was learned

	iteration in copy mechanism	iteration in coverage mechanism
10,000 articles	20,000 times	300 times
45,000 articles	80,000 times	1,000 times
90,000 articles	160,000 times	2,000 times
180,000 articles	320,000 times	4,000 times
287,226 articles	460,000 times	6,000 times

The Pointer-Generator program used in this research takes about 0.27 seconds to learn 1 batch using GeForce GTX 1080 Ti GPU. Therefore, it takes about an hour and a half to learn 10,000 articles. Similarly, it takes about 6 hours, 12 hours, 24 hours, 35 hours to learn 45,000 articles, 90,000 articles, 180,000 articles, 287,226 articles. While adding extended articles, the total epoch count is halved, therefore the learning time is almost the same.

3.4. *Result*

The ROUGE evaluation is shown below.

Table 3 Results on 10,000 articles. "original" is original data and "extended" is extended data.

	ROUGE-1-f	ROUGE-1-r	ROUGE-1-p	ROUGE-2-f	ROUGE-2-r	ROUGE-2-p
original	0.3226	0.3063	0.3665	0.1209	0.1142	0.1389
extended	0.3335	0.3181	0.3761	0.1275	0.1210	0.1452

	ROUGE-L-f	ROUGE-L-r	ROUGE-L-p
original	0.2994	0.2841	0.3403
exended	0.3085	0.2940	0.3482

Table 4 Results on 45,000 articles

	ROUGE-1-f	ROUGE-1-r	ROUGE-1-p	ROUGE-2-f	ROUGE-2-r	ROUGE-2-p
original	0.3398	0.3322	0.3731	0.1304	0.1269	0.1446
extended	0.3506	0.3556	0.3698	0.1378	0.1393	0.1464

	ROUGE-L-f	ROUGE-L-r	ROUGE-L-p
original	0.3155	0.3082	0.3468
extended	0.3249	0.3293	0.3428

Table 5 Results on 90,000 articles

	ROUGE-1-f	ROUGE-1-r	ROUGE-1-p	ROUGE-2-f	ROUGE-2-r	ROUGE-2-p
original	0.3538	0.3507	0.3830	0.1404	0.1387	0.1531
extended	0.3627	0.3565	0.3945	0.1495	0.1467	0.1636

	ROUGE-L-f	ROUGE-L-r	ROUGE-L-p
original	0.3288	0.3258	0.3562
extended	0.3372	0.3313	0.3671

Table 6 Results on 180,000 articles

	ROUGE-1-f	ROUGE-1-r	ROUGE-1-p	ROUGE-2-f	ROUGE-2-r	ROUGE-2-p
original	0.3758	0.3884	0.3893	0.1587	0.1635	0.1655
extended	0.3827	0.4027	0.3874	0.1628	0.1711	0.1655

	ROUGE-L-f	ROUGE-L-r	ROUGE-L-p
original	0.3475	0.3588	0.3604
extended	0.3515	0.3698	0.3561

Table 7 Results on 287,226 articles

	ROUGE-1-f	ROUGE-1-r	ROUGE-1-p	ROUGE-2-f	ROUGE-2-r	ROUGE-2-p
original	0.3918	0.4121	0.3958	0.1703	0.1784	0.1733
extended	0.3941	0.4288	0.3859	0.1721	0.1869	0.1693

	ROUGE-L-f	ROUGE-L-r	ROUGE-L-p
original	0.3593	0.3776	0.3633
extended	0.3625	0.3941	0.3551

3.5. *Discussion*

Table 3~7 show that the data augmentation is effective in all articles. However, as the number of articles increases, the data augmentation effect fades. Table 8 shows the value obtained by subtracting the F value of ROUGE of 'normal' from that of 'extended'.

Table 8 the value obtained by subtracting the F value of Rouge of 'normal' from that of 'extended'

	10,000articles	45,000articles	90,000articles	180,000articles	287,226articles
ROUGE-1-f	0.0109	0.0108	0.0089	0.0069	0.0023
ROUGE-2-f	0.0066	0.0074	0.0091	0.0041	0.0018
ROUGE-L-f	0.0091	0.0094	0.0084	0.0040	0.0032

Since all the values in Table 8 are positive, it is better to extend in each the number of articles. Also, looking at ROUGE-1, it gradually decreases as the number of articles increases. Looking at ROUGE-2, it goes up to 90,000 articles and decreases from there. Looking at ROUGE-L, it goes up to 45,000 articles and decreases from there.

The reason why the expansion effect is less likely to appear as the number of articles increases is as follows.

In a sufficient number of articles, learning became saturated and the expansion is not very effective. On the other hand, in the case where there were not many articles, the expansion is effective due to lack of learning. Therefore, it can be said that this method is more effective when the number of articles is small.

4. Conclusion

In this experiment, it is found that the value of ROUGE is improved by learning on adding an article from which the sentence with the least importance is extracted, rather than learning normally. Moreover, since this method has the same number of learning times as the case where it is not extended, it succeeded in improving the value of ROUGE without changing the learning time. In addition, compared to previous research [4], the overall value of ROUGE was much better, and it turns out that it can also be applied to a more accurate automatic summarization system. Moreover, the experiment of this research shows that the effect of expansion was small under the same conditions as the previous research [5]. However, it has also been found that reducing the number of articles increases the effect of expansion.

A future task is to increase the application range of this method. There are several models developed from the Pointer-Generator model[8,9]. I would like to examine whether proposed method is effective. Also, in this research, we used the topic model to determine the importance of sentences, therefore, I would also like to examine whether the data augmentation is more effective when we use ROUGE to determine the importance of sentences. Furthermore, in this research, the sentence with the lowest importance was extracted, therefore, it is possible to extract only the sentence with the highest importance and make it an extended article. I would like to consider such a new extension method as a future research subject.

References

1. Data Growth, Business Opportunities, and the IT Imperatives, http://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm
2. Dongxu Zhang, Zhichao Yang, “Word Embedding Perturbation for Sentence Classification” arXiv : 1804.08166vl (2018).
3. I. Saito, J. Suzuki, K.Nishida, K. Sadamitsu, S. Kobashikawa, Ryo. Masumura, Y. Matsumoto, J. Tomita “Improving Neural Text Normalization with Data Augmentation at Character- and Morphological Levels” Proceedings of the 8th International Joint Conference on Natural Language Processing, 257-262 (2017).
4. T. Ouchi, M. Tabuse “Effectiveness of Data Augmentation in Automatic Summarization System” ICAROB(2019)
5. Abigail. See, Peter J. Liu, et al. “Get To The Point: Summarization with Pointer-Generator Networks” arXiv:1704.04368(2017)
6. H.Sigematsu, I.Kobayashi “Generation of abstracts considering importance of potential topics” The Association for Natural Language Processing (2012) (In Japanese).
7. https://cs.nyu.edu/~kcho/DMQA/
8. Mike Lewis, Yinhan Liu, Naman Goyal, et al. “BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension” arXiv:1910.13461v1(2019)
9. Li Dong, Nan Yang, et al. “Unified Language Model Pre-training for Natural Language Understanding and Generation” arXiv:1905.03197v3(2019)

Mouse Cursor Control System Using Facial Movements

Masayoshi Tabuse [1], Manase Mizobe [2], Yasunari Yoshitomi [1], Taro Asada [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: tabuse@kpu.ac.jp
2 TORAY ENGINEERING Co.,Ltd
3-22 Yaesu Ryumeikan Building, Yaesu 1-chome, Chuo-ku, Tokyo 103-0028, 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 facial movement. Recognition of facial movement of a person makes it possible to operate a computer. Furthermore without the adjustment for a user and adjustment for the distance from a user, it is possible to reduce the burden on a user. We developed a system to resolve these problems. In our system, a web camera, dlib C++ library and OpenCV library are used to extract feature points of the face and obtain the face direction. Changing the face direction, we can move a mouse cursor. Recognizing an open mouth or closed eye, we can carry out an operation of mouse click. In this paper, we evaluated the effect on operability due to the face direction and recognition rate due to distance.

Keywords: Mouse cursor system, dlib C++ library, facial movements, Physically disabled people.

1. Introduction

Physically disabled people cannot move their limbs freely. It is difficult for a physically disabled person to use a computer. The importance of the computer with the information society increases as a current social background. Therefore, support of the computer operation is required. The interface for them has been developed so that they can use a computer recently. There are two types of the interface. One is the contact-type which operates while attaching the device to a body. Another is the noncontact-type which operates by recognizing the movement of a body. Although the contact-type has the advantage of easy detection of body movement, the user must attach it directly during use. On the other hand, noncontact-type must adjust parameters for a user. To resolve these problems reduce a burden for the user.

One of the possible physical movements of a physically disabled person is facial movement. This means the face direction and open and closed mouth or eyes. If a computer is operated with recognizing the facial movement, the physically disabled person can use it. We developed a system which operates a computer by the facial movement. In the previous study[1], we proposed a mouse cursor control system using Kinect. However Kinect got out of production. Therefore we developed a mouse cursor control system using a web camera instead of Kinect. Our system consists of a computer, a web camera, dlib C++ library[2] and OpenCV library[3]. It extracts feature points of the face and estimates the face direction. We can move a mouse cursor by changing the face direction. We can carry out an operation of mouse click by recognizing an open mouth or closed eye. This reduces the burden on the user.

2. Proposed System

2.1. *System overview*

In our system, a computer, a web camera, dlib C++ library and OpenCV library are used to extract feature points of the face and estimate the face direction. Dlib is an open source C++ library containing image processing and machine learning algorisms. Dlib can extract a face area in image and detect 68 facial feature points from the face area. OpenCV estimates a face direction using obtained facial feature points. Our system has three functions for a physically disabled person. Firstly, it can control a mouse cursor along face direction. Secondly, it can perform a mouse click by recognizing an open mouth. Finally, it can carry out a mouse click by judging the open or closed eyes. In the following subsections, these functions are explained in detail.

2.2. *Control of a mouse cursor by face direction*

Dlib extracts a face area in image and detect 68 facial feature points from the face area. We use 2 dimensional locations of 6 facial feature points (nose tip, chin, left eye left corner, right eye right corner, left mouth corner and right mouth corner). The solvePnP function in OpenCV estimates the rotation vector (the axis of rotation and the angle of rotation) of the face using these 6 locations, 3 dimensional face model and intrinsic camera parameters. The Rodrigues function in OpenCV converts the rotation vector to a rotation matrix. $pitch$ and yaw which express a value of the angle for vertical and horizontal face directions are calculated as follows:

$$pitch = \tan^{-1}\left(R_{32} / R_{33}\right)$$
$$yaw = \tan^{-1}\left(R_{21} / R_{11}\right),$$

where $R = (R_{ij})$ is a rotation matrix.

We define the face direction vector $\boldsymbol{a}$ as

$$\boldsymbol{a} = (\,yaw,\ pitch\,)\,.$$

The control of a mouse cursor is operated according to the face direction vector $\boldsymbol{a}$. The mouse cursor is moved in the same direction as the vector. The moving speed of a mouse cursor is changed according to the size of the vector. When $|\boldsymbol{a}| < 10$, we define the face is facing to the front and the mouse cursor does not move. When $10 \leqq |\boldsymbol{a}| < 15$, the mouse cursor moves slowly. When $15 \leqq |\boldsymbol{a}| < 20$, the mouse cursor moves normally. When $20 \leqq |\boldsymbol{a}|$, the mouse cursor moves fast.

2.3. *Recognition of open mouth*

We extract four feature points of mouth (top, bottom, left and right) using dlib, as shown in Fig. 1. The vertical length of the mouth is found by feature point coordinates of the top and bottom. The horizontal length of the mouth is found by feature point coordinates of the left and right. The rate of the open mouth R_m is expressed as $R_m = h_m / w_m$ where h_m and w_m denote the vertical and horizontal length of the mouth. Let Th_m be the threshold for judgment of opened or closed mouth. When $R_m > Th_m$, it is recognized that the mouth is opened. In the present study, the value of Th_m is set to 0.45, which was determined experimentally.

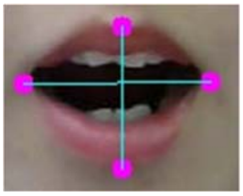


Fig. 1. Feature points of mouth extracted using dlib and the vertical and horizontal length.

2.4. *Judgment of opened and closed eyes*

We extract four feature points of eyes (top, bottom, left and right) using dlib, as shown in Fig. 2. The range of x-axis is the difference of x-coordinate between the left and right, and the range of y-axis is the difference of y-coordinate between the top and bottom. We define these ranges as eye region.

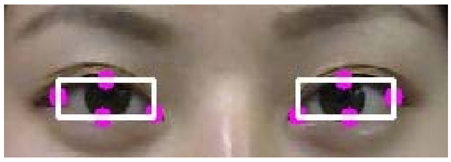


Fig. 2. Feature points of eye extracted using dlib.

The judgment of opened or closed eye uses binary images. For the binarization, at first, the detection of the eye region is performed in a RGB image. It is converted from the RGB image to a grayscale image. Then we binarize the grayscale image using discriminant analysis method[4]. The vertical and horizontal lengths of the eye are settled from the binary image, as shown in Fig. 3. We make the histogram of black pixels. The vertical length of the eye is the maximum on the histogram, and the horizontal length of the eye is the range of x-coordinate in the eye region. The rate of the open eye R_e is given by $R_e = h_e / w_e$ where h_e and w_e denote the vertical and horizontal length of the eye. Let Th_e be the threshold for judgment of opened or closed eye. When $R_e > Th_e$, the eye is judged to be opening state. When $R_e \leqq Th_e$, the eye is judged to be closing state. In the present study, the value of Th_e is set to 0.3, which was determined experimentally.

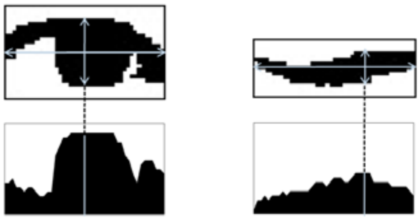


Fig. 3. The vertical and horizontal length and histogram of opened eye (left) and closed eye (right).

2.5. *Mouse click processing*

Mouse click processing is performed by recognizing the intentional movement of the user's mouth and eyes. The movement is the following conditions:

- When the duration time of opening mouth state reach for $t1$,
- When the duration time of closing eye state reach for $t2$,

where $t1$ and $t2$ are arbitrary times. In the present study, the value of $t1$ and $t2$ are set to 1.0 s and 0.8 s, respectively. Mouse click processing is carried out only in the case of $|\boldsymbol{a}| < 10$.

3. Experiments

3.1. *Conditions*

The experiment was performed in the following computational environment: the PC was a HP ENVY 700-560jp (CPU: Intel(R) Core(TM) i7-4790 CPU 3.60GHz, memory: 8.00GB); the OS was Microsoft Windows 8.1 Pro; the development language was Microsoft Visual C++ 2015 Express Edition and the libraries were dlib 19.16 and OpenCV 3.4.1. The image was produced by Logicool C270 HD Webcam which was placed on the top of the display, as shown in Fig. 4.

Fig. 4. Position of PC and a web camera in experiments.

3.2. *Method*

We conducted three kinds of experiments which were performed by six subjects in their twenties (2 males without glasses, 2 females without glasses and 2 females with glasses). For the first experiment, we measured times that a mouse cursor moved by face direction from a circle drawn on screen to another circle drawn in a distance of 20 cm. The subject sat at a distance of 0.8m from a web camera and moved the mouse cursor to eight directions (vertical, horizontal, and diagonal), respectively. This experiment was conducted to assess the operability by the difference of directions. For the second experiment, we examined the recognition rate of mouse click by open mouth. The distance between the web camera and the subjects was conducted at 0.6m, 0.8m, 1.0m, 1.5m, and 2.0m for evaluating influence to the recognition rate by the difference in distances. The subjects performed three times of the action of opening the mouth for one second or more. The recognition rate is the actual mouse click rate. For the last experiment, we conducted experiment to judge opened or closed eyes. This experiment was conducted at distance of 0.6m, 0.8m, 1.0m, 1.5m, and 2.0m between the web camera and subjects for evaluating influence to the judgment accuracy by the difference in distance. We evaluate the accuracy of judgment with F-measure. F-measure is expressed the following equation:

$$F = \frac{2PR}{P + R},$$

where P, R and F denote the rate of frames which is really closing eye among frames judged as closed state, the rate of frames judged as closing state among frames which is closing eye, and the harmonic mean of precision rate P and recall rate R. We performed judgment of closed eye by visual observation.

3.3. *Results and discussion*

4.3.1. Movement of mouse cursor by face direction

Fig. 5 shows the time to move the mouse cursor from a circle to another circle with 6 subjects. From the moving times of 6 subjects, it was the movement of top and bottom directions to take longer time than other directions. The subjects B and C wearing glasses took more time to move in all directions other than the bottom left compared to subjects without glasses. It is probably because the subject's glasses frame was misrecognized as subject's eye and this misrecognition interrupt subject's operation.

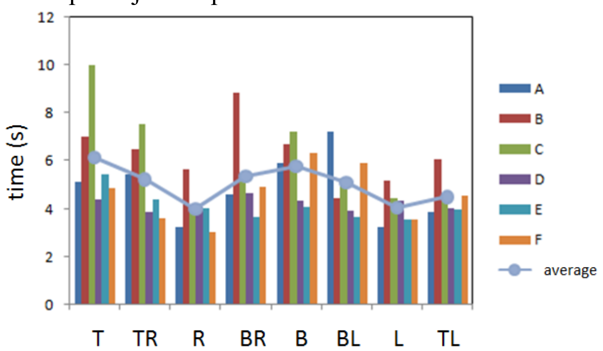


Fig. 5. Six subject A - F and Average moving time by face direction. T, TR, R, BR, B, BL, L and TL denote top, top right, right, bottom right, bottom, bottom left, left and top left directions, respectively.

4.3.2. Mouse click recognition by open mouth

In the previous study[1], there was a difference in the recognition rate of mouse clicks by opening the mouth depending on the distance, but in this study, mouse clicks could be recognized at any distance. The improved recognition rate is the difference in the angle of view between Kinect and the web camera and the brightness of the image. Since the angle of view of the web camera is smaller than that of Kinect, a face in the image can be obtained larger than that of Kinect and the accuracy of feature point detection is improved. In addition, the fact that the web camera was able to obtain in a brighter image compared to Kinect led to improved feature point detection accuracy.

4.3.3.Judgment accuracy of opened and closed eyes

Fig.6 shows the recognition rates of opened and closed eyes. Since the recognition rates exceeded 0.9 at all distances, it can be concluded that the judgment of opened and closed eyes was successful. This result is better than the previous study[1]. Since the angle of view of the web camera is smaller than that of Kinect, a face in the image can be obtained larger than that of Kinect and the accuracy of feature point detection is improved. In addition, the web camera was able to obtain in a brighter image compared to Kinect led and it improved feature point detection accuracy and reduced misrecognition of binarized eye areas.

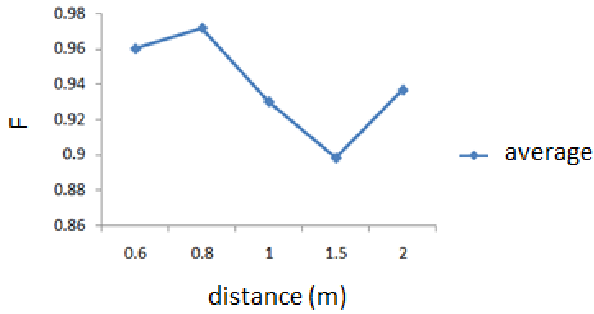


Fig. 6. F-measure of opening and closing eyes judgment.

4. Conclusion

We developed a mouse cursor control system for physically disabled people using a web camera, dlib and OpenCV. In this system, the user can move a mouse cursor by face direction and perform a mouse click processing by opening the user's mouth or closing the user's eye. As the results, the mouse cursor movements took more time to move up and down compared to right and left. Moreover it is somewhat difficult for a user with glasses to operate mouse cursor compared to a user without glasses. The mouse click operation performed by opening the mouth and closing the eyes did not differ in operability depending on the distance.

The future work is to improve the mouse cursor operability of a user with glasses and add other functions. Our system has only two functions, mouse cursor movement by face direction and mouse click by opening mouth and closing eyes. In order to improve practicality, it is necessary to enhance functions such as character input, double-clicking, and right-clicking.

References

1. J. Miyachi, M. Tabuse, "Development of Mouse System for Physically Disabled Person by Face Movement Using Kinect", Proceedings of the 2017 International Conference on Artificial Life and Robotics, (2017), 124-127.
2. Dlib C++ Library, http://dlib.net/
3. OpenCV, https://opencv.org/
4. Nobuyuki Otsu, "An Automatic Threshold Selection Method Based on Discriminant and Least Squares Criteria", *The IEICE Transactions*, *J63-D* (1980), 349-356.

Facial Expression Synthesis Using Vowel Recognition for Synthesized Speech

Taro Asada[1], Ruka Adachi[2], Syuhei Takada[3], 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: Software Service, Inc., 2-6-1 Nishi-Miyahara, Yodogawa-Ku, Osaka, Japan
3: Seika Town Hall, 70 Kitashiri, Minamiinayazuma, Nishi-Miyahara, Sagara-Gun, Kyoto, Japan

Abstract

Herein, we report on the development of a system for agent facial expression generation that uses vowel recognition when generating synthesized speech. The speech is recognized using the Julius high-performance, two-pass large vocabulary continuous speech recognition decoder software system, after which the agent's facial expression is synthesized using preset parameters that depend on each vowel. The agent was created using MikuMikuDanceAgent (MMDAgent), which is a freeware animation program that allows users to create and animate movies with agents.

Keywords: MMDAgent, Speech recognition, Vowel recognition, Speech synthesis.

1. Introduction

In Japan, the average age of the population has been increasing, and this trend is expected to continue. Because of this, researchers have been studying ways of applying information technology (IT) to improving the medical and/or mental support provided to older adults, including persons with extreme psychiatric disorders.

In our previous study[1], we developed a system for analyzing the facial expressions of a person obtained while answering interview questions posed by an animated agent. To accomplish this, we used MikuMikuDanceAgent (MMDAgent)[2], which is a freeware animation program that allows users to create and animate movies with agents.

In this study, to make the agent's performance on a personal computer (PC) screen more human-like, we have developed a system for agent facial expression generation that uses vowel recognition when generating synthesized speech.

2. Proposed System and Method

2.1. *System overview and outline of the method*

Figure 1 shows the processing flow of this system, which consists of six processing units:

- creating facial expression data, recording voice utterances, automatic WAVE file division,
- speech recognition by the Julius high-performance, two-pass large vocabulary continuous speech recognition decoder software[3],
- insertion of expressionless data, and
- the creation of facial expression motion.

The facial expression data are created in advance.

Creating facial expression data
↓
Recording agent voice
↓
Automatic division of WAVE files
↓
Speech recognition using Julius
↓
Inserting data of agent neutral facial expression
↓
Creating motion data for facial expression

Fig. 1. Processing flow of system.

3. Development of facial expression synthesis system

3.1. *Creating facial expression data*

Expression motions are generated by combining the expression data of each vowel for each utterance motion. Facial expression data were created with MikuMikuDance[4]. In this study, in order to realize more human-like agent facial expressions, facial expression data were created for the vowels / a /, / i /, / u /, / e /, and / o / (Fig. 2).

3.2. *Agent voice recording*

In our system, utterance contents are input as text and used by the MMDAgent to output synthesized voice that is then recorded by a stereo mixer inside a PC and saved as a WAVE file.

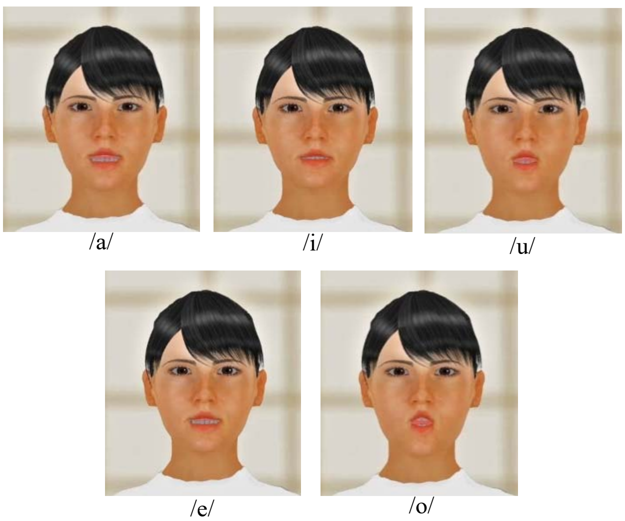


Fig. 2. Facial expression of the agent in uttering each vowel.

3.3. *Automatic division of WAVE files*

After all the utterances have been recorded, if there are multiple questions, the WAVE file is automatically divided for each question, and a new WAVE file is created for each question.

3.4. *Speech recognition using Julius*

The results of speech recognition using Julius are shown

```
sentence1:  食欲 は いかが です か 。
wseq1: <s> 食欲+名詞 は+助詞 いかが+形状詞 です+助動詞 か+助詞 </s>
phseq1: sp_S | sh_B o_I k_I u_I y_I o_I k_I u_E | w_B a_E | i_B k_I
cmscore1: 1.000 0.938 0.889 0.864 0.986 0.525 1.000
score1: 153.970123 (AM: 314.087799  LM: -160.117676)
=== begin forced alignment ===
-- phoneme alignment --
 id: from  to    n_score    unit
 ----------------------------------------
[   0    2]  -3.476162  sp_S+sh_B[sp_S]
[   3   13]  2.109969  sp_S-sh_B+o_I[sp-sh_B+o_B]
[  14   19]  1.689885  sh_B-o_I+k_I[ch_B-o_I+k_B]
[  20   27]  3.027793  o_I-k_I+u_I[o:_I-k_I+N_B]
[  28   32]  2.053262  k_I-u_I+y_I[k_B-u_I+y_B]
[  33   37]  1.453514  u_I-y_I+o_I[h_B-y_I+o_B]
[  38   44]  2.031667  y_I-o_I+k_I[y_E-o_I+k_B]
[  45   50]  2.800434  o_I-k_I+u_E[o:_I-k_I+N_E]
[  51   58]  2.530456  k_I-u_E+w_B[k_B-u_E+w_B]
[  59   64]  2.310862  u_E-w_B+a_E[h_E-w_B+N_E]
[  65   73]  1.006805  w_B-a_E+i_B[w_B-a_E+i:_B]
[  74   81]  2.154465  a_E-i_B+k_I[a:_E-i_B+k_B]
[  82   89]  2.765762  i_B-k_I+a_I[i:_B-k_I+a:_B]
[  90   96]  2.810837  k_I-a_I+g_I[f_E-a_I+g_E]
[  97  103]  2.868757  a_I-g_I+a_E[a:_B-g_I+a:_B]
[ 104  108]  3.200406  g_I-a_E+d_B
[ 109  114]  2.010101  a_E-d_B+e_I[a:_E-d_B+N_B]
[ 115  121]  1.140069  d_B-e_I+s_I
[ 122  125]  1.529968  e_I-s_I+u_E[e_I-s_I+h_E]
[ 126  131]  2.791662  s_I-u_E+k_B[s_B-u_E+f_B]
[ 132  139]  2.429546  u_E-k_B+a_E[h_E-k_B+a:_E]
[ 140  158]  0.035126  k_B-a_E+sp_S[f_B-a_E+sp]
[ 159  177]  0.900673  a_E-sp_S[sp_S]
re-computed AM score: 314.119537
=== end forced alignment ===
```

Fig. 3. Results of speech recognition by Julius.

0	o	14
1	o	27
2	u	28
3	u	37
4	o	38
5	o	50
6	u	51
7	u	64
8	a	65
9	a	81
10	i	82
11	i	89
12	a	90
13	a	103
14	a	109
15	a	114
16	e	122
17	e	125
18	u	126
19	u	139
20	a	140
21	a	177

Fig. 4. An output file listing vowels, and their start and end times.

in Fig. 3. From the recognition results, an output file is created that lists the vowels, and their start and end times (Fig. 4).

3.5. *Inserting neutral agent motion data*

In order to create more natural agent facial expressions, processing is then performed to insert a neutral facial expression when the vowel / a / is continuous. Figure 5 shows the insertion of neutral motion data executed when "いかがですか?" ("ikaga desu ka?") , which means, "How is it?", is spoken in Japanese. Since the vowel / a / is continuous when "ka" and "ga" are uttered, neutral data are inserted between them.

3.6. *Creating motion data for a facial expression*

Figure 6 shows the flow of creating a facial expression motion. The facial expression motion data, in the form of a Vocaloid Motion Data (vmd) file, is created by composing the vowel vmd file based on the speech duration (Sect. 3.4, Fig. 4). First, the number of bones of the wire frame model to be used is calculated from the number of vowels and neutral frames, and the total number is written in the vmd file header. Next, the vowel voicing time is converted to the number of frames (1 frame = 1/30 seconds) in order to set each bone of the facial expression data and is then written to the vmd file.

4. Experiment

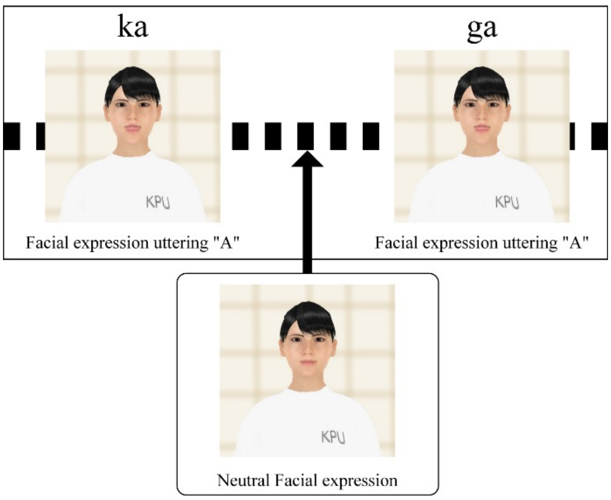


Fig. 5. Insert neutral data while vowel / a / is continuously uttered.

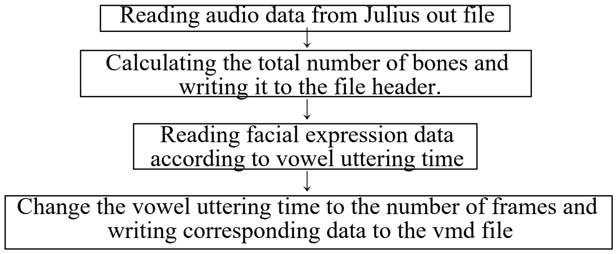


Fig. 6. Flow of facial expression motion creation.

4.1. *Conditions*

The experiment was performed on a Dell Inspiron 15 PC equipped with an Intel Core i7-6700HQ 2.2 GHz central processing unit (CPU) and 8.0 GB of random access memory (RAM). The Microsoft Windows 7 Professional operating system (OS) was installed on the PC, and Microsoft Visual C++ 2010 Express was used as the development language.

An animated agent that utters eight questions used in the initial diagnosis of depression by psychiatrists (Table 1) was created under two conditions (Condition 1: Created manually[3], Condition 2: Created with this system). In addition, an animated agent that asked nine questions (Table 2) on the Hasegawa Dementia Scale[5], which is used in the diagnosis of dementia, was also created using our system (Condition 3).

The content of the questionnaire is "How was the agent's mouth movements?" and the answer options were from a five-point scale (5: very natural, 4: natural, 3: normal, 2: unnatural, 1: very unnatural).

4.2. *Results and discussion*

Table 3 shows the average impression evaluation values of all subjects under Conditions 1 and 2. In Table 3, headings "1" to "8" indicate items by question no., and

Table 1. Depression diagnosis test.

No.	Question
1	Do you get depressed or feel gloomy in a daily life?
2	Do you feel less motivated or lacking energy?
3	Do you get enough sleep at night?
4	How is your appetite?
5	Do you enjoy any hobbies on your day off?
6	Do you feel worthless for yourself or hopeless on your alive days?
7	How about concentration and attention to work?

8	Do you work efficiently?

Table 2. Dementia diagnosis test.

No.	Question
1	How old are you?
2	What day of the month, what day, what day is it?
3	Where are we now?
4	Please tell me the three words you will say. Please remember it later.
5	Please subtract by seven in continuous order using 100 as the starting value.
6	Please tell me the numbers I will say from the reverse. 6-8-2.
7	Please say the words that you learned earlier.
8	I will show you five items. Tell me what happened because I'll hide it.
9	Please say as many vegetable names as you know.

the "average value" is obtained by averaging over all questions. In the average impression evaluation values shown in Table 3, Condition 2 was higher than Condition 1 for all questions, and the "average value" was about 18% higher. From these results, we confirmed that our system (Condition 2) has an advantage over manual work (Condition 1).

Table 4 shows the average impression evaluation values of all subjects under Condition 3. In Table 4, headings "1" to "9" indicate items by the question no., and the "average value" is obtained by averaging over all questions. From this table, it can be seen that because the animated utterances of the dementia diagnosis questions also obtained high impression evaluation values, the versatility of our system has been demonstrated.

Table 3. Average impression evaluation value of all subjects under Conditions 1 and 2.

Condition	1	2	3	4
1	2.29	2.71	2.86	3.57
2	3.29	3.71	3.29	3.86

Condition	5	6	7	8	Ave.
1	2.71	3.43	2.43	2.57	2.82
2	3.29	3.57	2.57	3.00	3.32

Table 4. Average impression evaluation value of all subjects under Condition 3.

Condition	1	2	3	4	5
3	3.57	3.86	3.57	4.00	3.57

Condition	6	7	8	9	Ave.
3	3.86	3.71	3.86	4.00	3.78

5. Conclusion

Herein, we reported on the development of a system for agent facial expression synthesis generation that uses vowel recognition when generating synthesized speech. The speech is recognized using the Julius high-performance, two-pass large vocabulary continuous speech recognition decoder software system, after which agent facial expression synthesis is performed using preset parameters depending on each vowel sound.

To create the agent, we used MMDAgent, which is a freeware animation program that allows users to create and animate movies with agents. To produce the agent's voice, we used the speech synthesis function setting built into MMDAgent. The impression evaluation values obtained from a questionnaire survey indicate that an agent produced by our proposed system is more natural than an agent created using preset parameters manually decided for each utterance. In the future, we plan to use this system for facial expression analysis and speech analysis experiments.

Acknowledgements

The authors would like to thank Professor J. Narumoto of the Kyoto Prefectural University of Medicine for his valuable support and helpful advice during the course of this research. We would also like to thank the subjects of our experiments for their cooperation. This research was supported by COI STREAM of the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

1. T. Asada, D. Kogi, R. Shimada, Y. Yoshitomi, and M. Tabuse, A System for Analyzing Facial Expression and Verbal Response of a Person while Answering Interview Questions by Agent, in *Proc. of Int. Conf. on Artif. Life and Robotics* (Beppu, Oita, Japan, 2018), 534-537.
2. MMDAgent, http://www.mmdagent.jp/ Accessed 6 July 2019.
3. Julius, http://Julius.osdn.jp/, Accessed 6 July 2019.
4. MikuMikuDance, https://sites.google.com/view/vpvp/, Accessed 9 December 2019.
5. Chiba medical association Revised Hasegawa's dementia scale (HDS-R) https://www.jpn-geriat-soc.or.jp/tool/pdf/tool_05.pdf, Accessed 6 July 2019.

Speech Synthesis of Emotions in a Sentence Using Vowel Features

Rintaro Makino[1], Yasunari Yoshitomi[2], Taro Asada[2], Masayoshi Tabuse[2]
1: SoftBank Corp., 1-9-1 Higashi-shimbashi, Minato-ku, Tokyo 105-7317, Japan
2: Graduate School of Life and Environmental Sciences, Kyoto Prefectural University,
1-5 Nakaragi-cho, Shimogamo, Sakyo-ku, Kyoto 606-8522, Japan
E-mail: {yoshitomi, tabuse}@kpu.ac.jp}, t_asada@mei.kpu.ac.jp,
http://www2.kpu.ac.jp/ningen/infsys/English_index.html

Abstract

We previously proposed a method for adding emotions to synthetic speech using the vowel features of a speaker. For the initial investigation in this earlier study, we used utterances of Japanese names to demonstrate the method. In the present study, we use the proposed method to construct emotional synthetic speech for a sentence formed from the emotional speech of a single male subject and produce results that are discriminable with good accuracy.

Keywords: Emotional speech, Feature parameter, Emotional synthetic speech, Vowel, Sentence.

1. Introduction

Recently, methods for adding emotions to synthetic speech have received considerable attention in the field of speech synthesis research.[1-8] To generate emotional synthetic speech, it is necessary to control the prosodic features of the utterances. Natural language is mainly composed of vowels and consonants. The Japanese language has five vowels. A vowel has a more dominant impact on the listener's impression than does a consonant, primarily because a vowel has a longer utterance time and larger amplitude in comparison to a consonant. We previously proposed a case-based method for generating emotional synthetic speech by exploiting the characteristics of maximum amplitude and utterance time for vowels, as obtained by using a speech recognition system, together with the fundamental frequency of emotional speech.[9]

In an earlier study,[10] we proposed an approach that further improved the method described in Ref. 9 by controlling the fundamental frequency of the emotional synthetic speech. The advantage of this earlier study[10] over prior research[1-8] was the usage of the vowel feature in emotional speech to generate synthetic emotional speech. In the previous study,[10] the speech included only Japanese names.

In the present study, we apply our previously proposed method[10] to creating emotional synthetic speech in sentence form, with new treatments for (1) smoothing the fundamental frequencies of sequential vowels and (2) suitably jointing the synthetic phonemes.

2. Proposed Method

In the first stage, we obtain audio data for emotional speech recorded as a WAV file. The subject is asked to speak in a way that reflects various emotional states: “angry,” “happy,” “neutral,” “sad,” and “surprised.” Then, for each type of emotional speech, we measure the time of each vowel utterance and the value of the maximum amplitude of the waveform while the vowel is being spoken, as well as the fundamental frequency of the emotional speech.[10]

In the second stage, we synthesize the phoneme sequence uttered by the subject. This stage consists of the following seven steps:

Step 1: For a vowel with a consonant appearing just before it in the synthetic speech with neutral emotion, the total phonation duration time of the vowel and the consonant is transformed into the time for speech with neutral emotion by the human subject. The synthetic speech obtained by this processing is hereinafter called “neutral synthetic speech.” [10]

Step 2: For a vowel with a consonant appearing just before it in the synthetic speech reflecting one of the intentional emotions ("angry," "happy," "sad," or "surprised"), the total phonation duration time of the vowel and consonant is set so that the ratio of this total to the corresponding total in the neutral synthetic speech is equal to the ratio of the phonation duration time of the vowel in the emotional speech to the phonation duration time of the vowel in the neutral speech. [10]
Step 3: The fundamental frequency of the synthetic speech obtained by the processing conducted up through Step 2 is initially adjusted based on the fundamental frequency of the emotional speech. [10]
Step 4: For a vowel with a consonant appearing just before it in the synthetic speech obtained by the processing conducted up through Step 3, the fundamental frequency is transformed into the average of three values of the sequential vowels that include it. For the first or last vowel in the sentence, two sets of fundamental frequencies are used for averaging.
Step 5: Amplitudes are transformed into final values by twice multiplying the ratio $\left(Max_{em}/Max_{ne}\right)$, where Max_{em} and Max_{ne} denote the maximum amplitude of the vowel in emotional speech and the amplitude of the vowel in neutral speech, respectively. The synthetic speech obtained by the processing up through Step 5 is hereinafter called "emotional synthetic speech."
Step 6: The fundamental frequency of the emotional synthetic speech obtained by the processing conducted up through Step 5 is further adjusted based on the fundamental frequency of the emotional speech.
Step 7: The synthetic speech for a sentence is generated by jointing all the synthetic speech phonemes constructing the sentence and adjusting the speech length to the phonation duration time of the emotional speech obtained in advance.

If no consonant appears before a vowel, the process described in Steps 1 through 6 applies only to the vowel.

3. Experiment

3.1. *Condition*

We used the speech recognition system known as Julius[12] to recognize the vowels and save the time position of the start and end of the vowels in each utterance. One male subject (Subject A) in his 50s was asked to speak the Japanese first names listed in Table 1 in a way that reflected various emotional states: "angry," "happy," "neutral," "sad," and "surprised." The audio data were recorded as WAV files. In preparation for generating the emotional synthetic speech, we measured the utterance time of the vowel, the maximum amplitude of the waveform, and the fundamental frequency while the vowel was being spoken. Tables 2, 3, and 4 show the phonation time average, the maximum amplitude average, and the fundamental frequency average, respectively, for each first vowel in each Japanese name in each emotion category as spoken by Subject A.

Voice Sommelier Neo (premium version; Hitachi Business Solution Co., Ltd., Yokohama, Japan)[13] was used as the speech synthesizer for each of the steps described in Section 2. In applying the method, the Male 1 (bright voice) mode in Voice Sommelier Neo was used. To convert the amplitude of each vowel and consonant described in Step 5 in Section 2, a digital audio editor was used. The method[11] using resampling was then used in Step 6 of Section 2.

To enable comparisons, subject A was asked to speak the sentence, 'このぬいぐるみかわいくない' (in

Table 1. Japanese first name used in experiments.

		First vowel				
		/a/	/i/	/u/	/e/	/o/
Last vowel	/a/	ayaka	shinnya	tsubasa	keita	tomoya
	/i/	kazuki	hikari	yuki	megumi	koji
	/u/	takeru	shigeru	fuyu	megu	noboru
	/e/	kaede	misae	yusuke	keisuke	kozue
	/o/	taro	hiroko	yuto	keiko	tomoko

Table 2. Phonation time average of each first vowel.

		Emotion				
		Angry	Happy	Neutral	Sad	Surprised
First vowel	/a/	62	110	72	112	58
	/i/	44	74	68	86	58
	/u/	48	120	130	86	44
	/e/	70	126	136	182	84
	/o/	70	104	118	176	84

(ms)

Table 3. Maximum amplitude average of each first vowel.

		Emotion				
		Angry	Happy	Neutral	Sad	Surprised
First vowel	/a/	1340	1714	714	573	1346
	/i/	362	385	287	199	400
	/u/	658	509	438	298	1017
	/e/	816	1079	794	575	1165
	/o/	748	1262	838	646	1479

Table 4. Fundamental frequency average of each first vowel.

		Emotion				
		Angry	Happy	Neutral	Sad	Surprised
First vowel	/a/	141	183	131	165	284
	/i/	114	192	134	191	286
	/u/	159	217	131	165	284
	/e/	122	186	178	229	276
	/o/	139	186	170	210	228

(Hz)

Japanese), which means, 'This stuffed toy is pretty, isn't it?" in a way that reflected each of the intentional emotions. ("Surprised" was not included here, as it was difficult to express in the sentence.) We used the recorded utterances as "emotional speech." We then generated synthetic speech for the same sentence ('このぬいぐるみかわいくない') using the method described in Section 2.

In all, 13 subjects participated in the experiment. These included one male in his 60s, one male in his 50s, one male in his 30s, five males in their 20s, and five females in their 20s. Each of the 13 subjects was asked to judge the emotional state of the speaker (angry, happy, neutral, sad) after listening to four utterances, one at a time, in the following order: the emotional speech by Subject A, the emotional synthetic speech. The 13 subjects were also asked which features were most important in judging the emotion of the sentence. There were seven choices available: three prosodic features, (1) utterance length, (2) height, and (3) volume; and four in-sentence positional features, (4) top, (5) middle, (6) last, and (7) total. For the three prosodic features, multiple answers were allowed.

3.2. *Results and discussion*

Table 5 shows the results of the subjective evaluations of the 13 subjects. As indicated, the mean accuracy of the categorizations for the emotional speech was 100%; for the emotional synthetic speech, the mean accuracy was 78.9%. The accuracy for the "happy" synthetic speech was highest, while that of "neutral" was lowest. Table 6 shows the results of the questions regarding the importance of various features in judging the emotional category. As shown, among the three prosodic characteristics, the "height of the voice" had the most influence, while among the four sentence positions, the "last position" appears to have influenced the subjects the most. Fig. 1 shows the waveforms of the emotional speech and the emotional synthetic speech. As illustrated in the figure, the waveform associated with the "happy" synthetic speech appears to be the most similar to its emotional speech counterpart among the waveforms of the four emotional synthetic speech types. Tables 7 and 8 show the results of the subjective evaluations for the names "Taro" and "Hiroko," respectively, as reported in our previous study.[10] In these two tables, the results for the emotional speech were calculated as the average of the values obtained in two sets of listening by all 18 subjects. Based on the values in Table 7, for "Taro," the mean accuracy for the emotional speech categorizations using the previously proposed method was 95.0%, while for the emotional synthetic speech, it was 90.0%.[10] For "Hiroko," the mean accuracy for the emotional speech categorizations

Table 5. Results of subjective evaluations.

(1) Emotional Speech

		Input			
		Angry	Happy	Neutral	Sad
Recognition	Angry	100	0	0	0
	Happy	0	100	0	0
	Neutral	0	0	100	0
	Sad	0	0	0	100

(%)

(2) Emotional Synthetic Speech

		Input			
		Angry	Happy	Neutral	Sad
Recognition	Angry	76.9	7.7	15.4	0
	Happy	0	92.3	0	15.4
	Neutral	23.1	0	69.2	7.7
	Sad	0	0	15.4	76.9

(%)

Table 6. Results of questions regarding the importance of various factors in judging the emotion category.

	Voice			Sentence position			
	Length	Height	Volume	Top	Middle	Last	Total
Number of vote	7	11	9	0	0	8	5
Ratio (%)	54	85	69	0	0	62	38

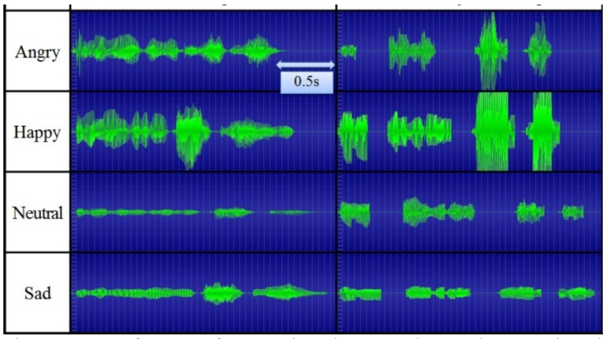


Fig. 1. Waveform of emotional speech and emotional synthetic speech for the intentional emotions of "angry," "happy," "neutral," and "sad " in the utterance of 'このぬいぐるみかわいくない' (in Japanese).

was 96.7%, while for the emotional synthetic speech, it was 77.8% (Table 8).[10] Combining the results for both "Taro" and "Hiroko," the average accuracy for the emotional speech categorizations was 95.9%, while the average accuracy for the emotional synthetic speech was 83.9%.

A comparison of the combined 83.9% average accuracy reported for the emotional synthetic speech categorizations in the previous study (Tables 7 and 8)[10] and the lower 78.9% average accuracy in the present study (Table 5(2)) suggests that it may be more difficult to effectively generate emotional synthetic speech in a sentence rather than for a first name. (It should be noted that there were four emotion categories in the more recent experiment versus five in the previous study[10]. As mentioned earlier, "surprised" was omitted in the current study.)

4. Conclusion

Recently, methods for adding emotion to synthetic speech have received considerable attention in speech synthesis research. We had previously proposed a method for speech synthesis with emotions using the vowel features of a speaker. In the present study, we used the proposed method to create emotional synthetic speech in a sentence based on the emotional speech of an individual and found that the results were discriminable with a good accuracy.

Acknowledgments

We would like to thank all those who participated in the experiments described here.

Table 7. Results of the subjective evaluation for "Taro."[10]

(1) Emotional speech

		Input				
		Angry	Happy	Neutral	Sad	Surprised
Recognition	Angry	97.2	0	0	2.8	0
	Happy	0	94.4	0	0	5.6
	Neutral	0	2.8	94.4	2.8	0
	Sad	0	0	5.6	94.4	0
	Surprised	2.8	2.8	0	0	94.4

(%)

(2) Emotional synthetic speech

		Input				
		Angry	Happy	Neutral	Sad	Surprised
Recognition	Angry	88.8	5.6	0	0	11.1
	Happy	5.6	88.8	0	0	0
	Neutral	0	5.6	94.4	11.1	0
	Sad	0	0	5.6	88.9	0
	Surprised	5.6	0	0	0	88.9

(%)

Table 8. Results of the subjective evaluation for "Hiroko."[10]

(1) Emotional speech

		Input				
		Angry	Happy	Neutral	Sad	Surprised
Recognition	Angry	91.7	0	0	0	8.3
	Happy	0	100	0	0	0
	Neutral	0	0	100	0	0
	Sad	0	0	0	100	0
	Surprised	8.3	0	0	0	91.7

(%)

(2) Emotional synthetic speech

		Input				
		Angry	Happy	Neutral	Sad	Surprised
Recognition	Angry	83.3	0	0	0	16.6
	Happy	0	66.7	0	22.2	5.6
	Neutral	0	27.8	94.4	11.1	0
	Sad	0	0	5.6	66.7	0
	Surprised	16.7	5.5	0	0	77.8

(%)

References

1. E. Donna, Expressive speech: Production, perception and application to speech synthesis, *Acoustical Science and Technology*, **26**(4), 2005, pp.317–325.
2. A. Iida, S. Iga, F. Higuchi, et al, A prototype of a speech synthesis system with emotion for assisting communication (in Japanese), *Trans. Human Interface Society*, **2**(2), 2000, pp.63-70.
3. N. Katae, S. Kimura, An effect of voice quality and control in emotional speech synthesis (in Japanese), *Proceedings of the autumn meeting the Acoustical Society of Japan*, Vol. 2, Iwate, Japan, Sept., 2000, pp.187-188.
4. T. Moriyama, S. Mori, S. Ozawa, A synthesis method of emotional speech using subspace constraints in prosody (in Japanese), *Trans. IPSJ*, **50**(3), 2009, pp.1181-1191.
5. I. R. Murray, J. L. Arnott, Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *J. Acoustical Society of America*, **93**(2), 1993, pp.1097–1108.
6. I. R. Murray, M. D. Edgington, D. Campion, et al, Rule-based emotion synthesis using concatenated speech. *Proceedings of Speech and Emotion, ISCA Tutorial and Research Workshop*, Newcastle, UK, Sep., 2000, pp 173–177.
7. S. Ogata, T. Yotsukura, S. Morishima, Voice conversion to append emotional impression by controlling articulation information (in Japanese). *IEICE technical report, human information processing*, **99**(582), 2000, pp.53-58.
8. M. Schröder, (2001) Emotional speech synthesis – A review. In: Dalsgaard P, Lindberg B, Benner H (Eds.), *Proceedings of 7th European Conference on Speech Communication and Technology*, Aalborg, Kommunik Grafiske Losninger A/S, Vol.1, pp 561–564.
9. K. Boku, T. Asada, Y. Yoshitomi et al, Speech synthesis of emotions using vowel features. in Roger L (Ed.), *Software engineering, artificial intelligence, networking, and parallel/distributed computing 2012*, Springer, Berlin, 2012, pp.129-141.
10. K.Boku, T.Asada, Y.Yoshitomi, and M.Tabuse, Speech synthesis of emotions using vowel features of a speaker, *J. Artificial Life and Robotics*, **19**(1), 2014, pp.27-32.
11. N. Aoki, in *Sound programming in C* (in Japanese), Ohmsha, Tokyo, 2008, pp.141-160.
12. Julius development team, Open-source large vocabulary CSR engine Julius. http://julius.sourceforge.jp/ Accessed 11 December 2019.
13. Hitachi Business Solution Co. Ltd. Voice Sommelier Neo. http://hitachi-business.com/products/package/sound/voice/ Accessed 12 March 2013.

Extraction of Irrelevant Sentences from Online Hotel Reviews

Shogo Watanabe
Department of Information Science, Okayama University of Science,
1-1 Ridaicho, Kita-ku, Okayama-shi, 700-0005, Japan

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: i16i092ws@ous.jp, hirota@mis.ous.ac.jp, oda@ice.ous.ac.jp

Abstract

Many reviews of hotels have been posted on review sites such as TripAdvisor and Yelp. Many tourists select a hotel to reserve using their ratings and reviews. Although containing useful information, those reviews may also contain useless information, which reduces their readability. Removing irrelevant sentences from those reviews can improve their readability. This paper proposes a method to extract irrelevant sentences from a review. Our approach uses a supervised learning method to classify sentences into relevant and irrelevant. We demonstrate the performance of our proposed method by evaluation experiment using the TripAdvisor dataset.

Keywords: Review analysis, Hotel reviews, Classification, Supervised learning method.

1. Introduction

Recently, the number of people who travel abroad have increased in many counry due to transportation and ecinomic growth. This has led to an increase in the number of them using accommodations such as hotels. As a result, reviews posted to review sites about them have increased. These reviews are used for many types of research, such as characteristic analysis of restaurant and hotel [1], and human sentiment analysis [2].

Online reviews may influence people who look for information about their next trip. When users decide in which hotel they want to stay, they inspect reviews to get information posted by other users. However, whereas these reviews include relevant information about hotels, they also include a large amount of information which is irrelevant to hotels, such as transportation to and from the hotel, nearby tourist sites, and affiliated facilities.

Table 1 shows some examples of sentences extracted from hotel reviews. It is difficult for people to get only the relevant information for them. We assume that users inspect reviews that are displayed at the top of the review site to lighten their burden. However, to get the relevant information, the users need to inspect many reviews. There is a necessity to remove review sentences not related to hotels from reviews and provide review sentences related to hotels for the user. This can lighten the user's burden and improve their satisfaction.

Researches [2] [3] that extract reviews with useful or helpful information have been conducted. However, such information depends on the user. Therefore, we propose a method to extract sentences in a review, which are irrelevant to hotels.

Table 1. Some examples of review sentences.

Class	Reviews
NOTRELATED	The shows at night are very nice especially Chicago and the Lion King.
ROOMS	Walls were a bit thin, but we didn't have any major issues.
SERVICE	Best of all – all the staff are so friendly, informative and helpful.

Table 2. The number of sentences in each class.

Class	The number of sentences
BUILDING	49
FOOD	104
LOCATION	106
ROOMS	161
SERVICE	170
VALUE	39
NOTRELATED	321

In this paper, we propose a method that extracts sentences not related to hotels from reviews, using a supervised learning method. Also, we classify the related sentences into various topics.

2. Related work

There are many kinds of reviews posted online, which were written about hotels, restaurants, travel sites, and merchandise. This section describes the related studies of tourist information using those reviews.

Timmaraju et al. proposed a method to predict sentiment on movie reviews using recursive and recurrent neural network architectures. The proposed feature achieved the almost same performance as hand-crafted features [2]. Wang et al. proposed a method to detect restaurants that should inspect a foodborne illness using a supervised learning algorithm [1]. Bhatt et al. proposed a method to classify whether reviews of an e-commerce site were written about the service or product. Additionally, they classified the reviews about the product into positive or negative [3].

Previous researches have extracted relevant sentences and classify review contents. In this paper, we proposed a method to extract sentences trrelevant to hotels from hotel reviews.

3. Method

In this section, we describe our method for extracting irrelevant sentences and classify relevant sentences into a topic. Our method includes the following steps.

1. Dividing reviews into sentences.
2. Applying the pre-processing to sentences.
3. Calculating feature vectors from text sentences.
4. Classifying sentences into relevant or not relevant using the supervised learning method.
5. Classifying the relevant reviews into topics using the supervised learning method.

3.1. *Class of review sentences*

We used a dataset which includes reviews from TripAdvisor[a] [4]. This dataset includes 1,576 review sentences with 11 classes manually annotated. In this dataset, the [OTHER] class includes miscellaneous sentences which are related to the hotels, but do not belong to other classes.

In this paper, we exclude the [BUSINESS] class and the [CLEANLINESS] class because those two classes have a small number of data. Also, because contents of the [OTHER] class and the [NOTRELATED] class are similar, we add the [OTHER] class to the [NOTRELATED] class. For the same reason, we add the [CHECKIN] class to the [SERVICE] class. In this paper, we deal with the [NOTRELATED] class as sentences not related to hotels and other classes as related to hotels.

3.2. *Pre-processing*

We use only review written in English. We apply langdetect[b] to detect the review language. If the language is not English, we do not apply further processing. In this paper, we divide the review into sentences based on symbols such as ".", "!" and "?". Also, other pre-processing is based on a pre-trained model for fastText

[a] https://www.triadvisor.jp/

[b] https://pypi.org/project/langdetect/

Table 3. Performance of the classification: Related or Not related.

	Classifier	Class	Precision	Recall	F1-score	Accuracy
fastText	SVM	RELATED	0.81	0.89	0.85	0.78
		NOTRELATED	0.68	0.54	0.60	
	Random Forest	RELATED	0.81	0.84	0.82	0.75
		NOTRELATED	0.61	0.56	0.58	
	XGBoosst	RELATED	0.80	0.80	0.80	0.73
		NOTRELATED	0.56	0.56	0.56	
	Classifier	Class	Precision	Recall	F1-score	Accuracy
BERT	SVM	RELATED	0.85	0.84	0.85	0.79
		NOTRELATED	0.66	0.68	0.67	
	Random Forest	RELATED	0.82	0.85	0.83	0.77
		NOTRELATED	0.64	0.59	0.61	
	XGBoosst	RELATED	0.83	0.84	0.83	0.77
		NOTRELATED	0.63	0.61	0.62	

[5]. We delete the numbers and lower all character strings. We delete stopwords such as "a" and "you" using nltk[c].

3.3. *Vectorization*

In this paper, we use two methods to generate feature vectors. First method is the pre-trained model of fastText. We extract 300 dimension vectors from each word included in the sentences. We calculate the average vectors of the vector of each word. We regard the average vectors as a feature vector of a sentence. The second method is Bidirectional Encoder Representations from Transformers (BERT) [6]. BERT can calculate a vector from a sentence. We extract 768 dimension vectors from a sentence and regard the vector as the feature of the sentence.

4. Experiments

4.1. *Dataset*

We used the dataset as described in Section 3.1 for learning classifiers. We divided the dataset into the training data and the test data at 8 : 2.

The number of sentences of [RELATED] class was 498. The number of sentences of [NOTRELATED] class was 262. We applied SMOTE [7] to the [NOTRELATED] class because that size was an imbalance. As a result, both classes were 498 data.

4.2. *Experimental conditions*

In this paper, we used three classifiers: Random Forest [8], Support Vector Machine (SVM) [9], and XGBoost [10]. We used Scikit-learn [11] to implement SVM and Random Forest. To generate the highest performance model, we performed a grid search with 10-fold cross-validation for each method.

We used Precision, Recall, F1-score, and Accuracy as evaluation criteria.

4.3. *Experiment results*

4.3.1. Removal of irreverent sentence

We describe the elimination of [NOTRELATED] sentences. Table 3 shows the performance of the classifier. The model with the best F1-score was SVM with BERT feature vectors. Here, we show two examples of [LOCATION] sentences misclassified as [NOTRELATED].

- We chose this resort because it, supposedly, had a first class golf course very near the resort.
- We spent our days going back and forth between the beach (absolutely beautiful clear water and white soft, sand) and the pool.

The misclassified sentences into [NOTRELATED] class include words related to hotels such as “beach” and “resort.” We assume that these words adversely affect the feature vectors of [NOTRELATED] class.

[c] https://www.nltk.org/

Table 4. Performance of the classification: topic classification.

		BUILDING	FOOD	LOCATION	SERVICE	ROOMS	VALUE	NOTRELATED
SVM	Precision	0.40	0.78	0.84	0.74	0.78	0.67	0.29
	Recall	0.22	0.82	0.94	0.68	0.81	0.33	0.37
	F1-score	0.29	0.80	0.89	0.71	0.79	0.44	0.33
Random Forest	Precision	0.00	0.88	0.73	0.81	0.76	0.67	0.38
	Recall	0.00	0.82	0.94	0.68	0.81	0.67	0.47
	F1-score	0.00	0.85	0.82	0.74	0.78	0.67	0.42
XGBoost	Precision	0.67	0.88	0.75	0.77	0.80	0.60	0.35
	Recall	0.22	0.82	0.88	0.68	0.89	0.50	0.42
	F1-score	0.33	0.85	0.81	0.72	0.84	0.55	0.38

The performance of removing irrelevant reviews is limited. Therefore, some [NOTRELATED] sentences are included in the dataset. The [NOTRELATED] class is included in the topic classification.

4.3.2. *Topic classification of sentences*

Table 4 shows the result of the multi-class classification. The F1-score of [NOTRELATED] class is low on thoes classifiers. The reason is that the classification performance for [NOTRELATED] sentences in the previous step is insufficient. Also, the performance of [BUILDING] class is low. Most of [BUILDING] sentences are classified into [ROOMS] class.

5. Conclusion

In this paper, we classify review sentences into related and not related to hotels. Also, we classify related sentence to topics. We compared the performances of some supervised learning method and feature vectors. We discussed the evaluation results of our experiments.

Future works will include applying the semi-supervised learning to improve performance of multi-class classification.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number 16K00157, 16K16158, and Tokyo Metropolitan University Grant-in-Aid for Research on Priority Areas Research on social big data.

References

1. Zhu Wang, Booma Sowkarthiga Balasubramani, Isabel F. Cruz, *Predictive Analytics Using Text Classification for Restaurant Inspections*, In Proceedings of the 3rd ACM SIGSPATIAL Workshop on Smart Cities and Urban Analytics. Redond Beach, CA, USA, 2017.
2. Aditya Srinivas Timma and Vikesh Khanna, *Sentiment Analysis on Movie Reviews using Recursive and Recurrent Neural Network Architectures*, Semantic Scholar, 2015.
3. Bhatt Aashutosh, Patel Ankit, Chheda Harsh, Gawande, Kiran, *Amazon review classification and sentiment analysis*, In International Journal of Computer Science and Information Technologies, 6(6), 2015, pp. 5107-5110.
4. Diego Marcheggiani, Oscar Täckström, Andrea Esuli, Fabrizio Sebastiani, *Hierarchical Multi-label Conditional Random Fields for Aspect-Oriented Opinion Mining*, Advances in Information Retrieval, 2014, pp.273-285.
5. Bojanowski Piotr, Grave Edouard, Joulin Armand, Mikolov Tomas, *Enriching Word Vectors with Subword Information*, arXiv preprint arXiv, 2016.
6. Devlin Jacob, Chang Ming-Wei, Lee Kenton, Toutanova Kristina, *Bert: Pre-training of deep bidirectional transformers for language understanding*, arXiv preprint arXiv:1810.04805, Oct. 2018.
7. Nitesh V Chawla, Kevin W Bowyer, Lawrence O Hall, W philip Kegelmeyer, *SMOTE: synthetic minority over-sampling technique*, Journal of artificial intelligence research, 2002, pp. 321-357.
8. Leo Breiman, *Random Forests*, Machine Learning, 45(1), 2001, pp.5-32.
9. Corina Cortes and Vladimir Vapnik, *Support-vector networks,* Machine Learning, 20(3), 1995, pp. 273-297.
10. Tianqi Chen and Carlos Guestrin, *Xgboost: A scalable tree boosting system*, In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, California, USA, Aug. 2016, pp. 785-794.
11. Pedregosa Fabian, Varoquaux Ga, Gramfort Alexandre, Michel Vincent, Thirion Bertrand, Grisel Olivier, Blondel Mathieu, Prettenhofer Peter, Weiss Ron, Dubourg Vincent, others, *Scikit-learn: Machine learning in Python*, Journal of machine learning research, Oct. 2011, pp. 2825-283.

Inferring Home Location of Foreign Tourists Based on Travel Routes Extracted from Social Media Sites

Chen Lugasi
Department of Information Science, Okayama University of Science
1-1 Ridaicho, Kita-ku, Okayama-shi, 700-0005, Japan

Masaharu Hirota
Department of Information Science, Okayama University of Science
1-1 Ridaicho, Kita-ku, Okayama-shi, 700-0005, Japan
E-mail: i16i090lc@ous.jp, hirota@mis.ous.ac.jp

Abstract

Tourists from certain regions tend to visit certain places when traveling abroad. The availability of large amounts of data from social media sites allows researchers to profile their tendencies, which could be useful for many applications such as route recommendations or advertisements. We propose a method to infer the home location of a tourist, based on their travel route, using the extracted tendency for each country. Our approach uses supervised machine learning methods to learn the quantized travel route of each user. In this paper, we use foreign travelers in Japan as a case study. We evaluate the performance of our proposed method by using photographs collected from their user accounts on Flickr.

Keywords: home location inference, user trajectory, machine learning, Flickr.

1. Introduction

With the rapid growth of social media services, it has become common among tourists to upload geo-tagged contents from their trips to online services, which led to an overflow of available information about tourist attractions. Users of those services are being constantly exposed to the same tourist attractions as their friends, which affects their decisions when planning their next trip, whether it is domestic or international.

Some places are widely popular tourist spots, like the Eiffel Tower in France and the Leaning Tower of Pisa in Italy. However, when traveling abroad, tourists from a particular country may be interested in specific places, which are not necessarily conventional tourist spots. For instance, places where a famous movie in their origin country was filmed, a restaurant whose owners are from their country, or a place that was featured on a local TV show.

By extracting only the differences in travel route between travelers from various origin countries, we assume that it is possible to create a tendency profile that can be used to predict the home location of a tourist based on their travel route, and vice versa.

Knowing a tourist's home location has many applications. For instance, it can be used to improve and personalize travel route recommendation and targeted advertisement for tourists. However, this information is not accessible for the most part. Therefore, there is a need to infer it based on other information. We propose a method to do that by finding the difference in traveling routes between tourists from different countries, based on their publicly-shared contents.

The remainder of this paper is organized as follows: Section 2 describes the work related to this topic. Section 3 presents the methodology used. Section 4 presents the experiment conducted in order to verify our methodology. Finally, Section 5 concludes this paper and proposes future research directions.

2. Related Work

A wide range of methods for inferring home location from social media content has been suggested in recent years.

Hironaka et al.[1] used data from Twitter to analyze users' home location based on their relationships with their friends. Hu et al.[2] suggested a method to infer home location from sparse and noisy Twitter data within 100 by 100 meter squares at high accuracy using users' trajectory in their home country. Jurgens et al.[3] evaluated several methods for geo-location prediction using data from Twitter.

While the above methods focus on inferring home location of social media users by using contents posted in their home country, in this study, we propose a new method for home location inference by analyzing the tendency profile of foreign tourists using travel routes in a country that is different from their home country.

3. Methodology

Our methodology relies on three consecutive steps; extraction of tourists contents (from now on will be referred to as "photo streams"), and dividing into individual visits (first step), clusters creation which will be used for quantizing individual users' trajectory (second step), and inferring the home location of a tourist from their trajectory using machine learning algorithm (third step).

3.1. *Extracting Tourists' photo streams*

To retrieve photographs from users who are not residents of the target country, we extract photo streams that are confirmed to be of a certain length of stay and show mobility patterns that indicate visits to different places in a short period of time.

Because a user may visit the same country several times in a given period, we, therefore, define one visit as being at least three days long and no longer than the standard for tourists' stay permit in that country. The interval between two visits is defined to be at least four weeks because it is unreasonable for a tourist who usually takes and uploads photographs to not take photographs for more than that period.

For each visit, it is required that it contains at least two photographs that have different GPS data (taken in different places).

3.2. *Creating Clusters*

In order to extract a user's trajectory, it is necessary to define areas that can be referred to as a single place between which the tourists can travel.

We create two types of clusters based on the obtained data; The first cluster combines photographs taken in the same region, generated from the photograph's longitude and latitude compared against a predefined regions list.

The second cluster is obtained by using grid clustering. The country is divided into an MxN grid with cells of about 6x6 kilometers width and height. Adjacent cells that meet a certain threshold are merged together up to a width of about 60x60 kilometers. The threshold definition is illustrated in Eq. (1).

$$threshold = \frac{\text{sum(unique users in a cell)}}{\text{total cells with data}} \quad (1)$$

3.3. *Inferring Home Location*

In this study, we use two types of classifiers; Long Short-Term Memory (LSTM) [4] and Support Vector Machine (SVM) [5]. We perform classification using all-versus-all method in LSTM and one-versus-all method in SVM. The feature vectors for those models uses quantized user trajectories as described in Section 3.2.

4. Experiment

4.1. *Dataset Preparation*

Flickr API[a] was used to collect metadata from publicly posted photographs with Global Positioning System (GPS) location data in the target country.

For the case study, data from Japan between April 1, 2009 and April 1, 2019 was used. We filtered the results to contain only users whose home location (owner

[a] https://www.flickr.com/services/api/

Table 1. Users, Visits and Photographs for each country used for the experiment.

Country	Users	Visits	Photographs
United States	1,490	2,222	168,680
Taiwan	966	2,044	351,333
United Kingdom	674	910	75,252
Australia	460	660	54,045
France	424	574	34,293
Canada	369	518	42,361
Italy	298	374	18,766
China	290	492	22,958
Spain	284	380	25,759
Germany	253	351	21,179
Singapore	226	381	32,298
Hong Kong	195	440	42,427
Netherlands	155	210	14,189

location, in Flickr API terms) is in a country other than the target country.

In order to decrease the negative influence of inaccurate data on our results, we only used photographs that contained GPS data accurate to the millionth decimal place, and photographs with the highest timestamps granularity.

Because "owner location" field in Flickr is an open text field, the data in it varies widely. To cope with that, we mapped all the abbreviations and major states/cities to the same country name. For instance, "usa", "u.s.a" and "Dallas, Texas" were mapped to "United States".

The final result yielded 1,084,645 photographs generated by 8,286 users from 102 countries. However, only countries with more than 200 visits were included, as indicated in Table 1.

4.2. Preliminary Experiment

In order to confirm the existence of variation between users from different locations, we first conducted a preliminary experiment.

The target country's whole area was divided into a 50x50 grid. Photographs of users from specific home locations were mapped into the grid. As a result, each cell in the grid contained the number of unique user visits. The results were normalized and represented as a 50x50 matrix. Then, we performed a matrix subtraction with other origin countries.

We observed a significant difference between different origin countries, a difference that remained about the same regardless of the grid size, indicating its significance.

These results were then compared to the average difference between each origin country's users to confirm that they were indeed larger.

4.3. Evaluation Method

For LSTM evaluation F1-score was used. For SVM a one-versus-all model was created for each country, and therefore we evaluated each model while accounting for the collective result.

We used each of the models' prediction probability to rank them in descending order. As a result, the model with the best probability for a prediction candidate was positioned first in the list. Because tourist data is extremely noisy, we allowed up to 3% of measurement error and therefore checked the next candidates in the range of 3% probability difference.

4.4. Results

Multi-class classification using LSTM showed poor results. Despite numerous optimization attempts varying from changing the number of feature vectors, to adjusting the clustering method and number of clusters, F1-score never crossed 0.3, which is considered low.

Results from SVM one-versus-all model are shown in Table 2. Out of 1,106 users in the test data, 168 users were classified with the best score by their home country's model, and 764 users were correctly classified with a measurement error of up to 3% compared to other country models. The rest 174 users were incorrectly classified with probabilities ranging from 10% to 100% error.

Table 2. Experiment Results Using SVM Classifier.

		Top Candidate		
		Yes	No	*Total*
Model Prediction	Correct	168	764	***932***
	Incorrect	14	160	*174*
	Total	*182*	*924*	*1106*

5. Conclusions and Future Work

In this paper, we proposed a method to infer the home location using a tourist's trajectory. Our approach used the tendency of tourists from the same country to travel to the same places. Experimental results showed that our classifier could estimate the candidates of tourist's home location.

In future work, we plan to expand the features to include not only the origin country but also other parameters such as gender, traveling season, and traveling form.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number 16K00157, 16K16158, and Tokyo Metropolitan University Grant-in-Aid for Research on Priority Areas Research on social big data.

References

1. Shiori Hironaka, Mitsuo Yoshida, Kyoji Umemura, *User's Centrality Analysis for Home Location Estimation*. Proceedings of IEEE/WIC/ACM International Conference on Web Intelligence, Thessaloniki, Greece, Oct. 14–17, 2019.
2. Tianran Hu, Jiebo Luo, Henry Kautz and Adam Sadilek, *Home Location Inference from Sparse and Noisy Data: Models and Applications.* Proceedings of IEEE 15th International Conference on Data Mining Workshops, Atlantic City, NJ, USA, Nov.14-17, 2015.
3. David Jurgens, Tyler Finethy, James McCorriston, Yi Tian Xu, Derek Ruths, *Geolocation Prediction in Twitter Using Social Networks: A Critical Analysis and Review of Current Practice*. Proceedings of the Ninth International AAAI Conference on Web and Social Media, Oxford, England, May. 26-29, 2015.
4. Hochreiter, Sepp & Schmidhuber, Jürgen, *Long Short-term Memory*. Neural computation. 9. 1735-80. 10.1162/neco, 1997.
5. C. Cortes and V. Vapnik, *Support-vector networks*, Machine Learning, vol. 20, no. 3, 1995, pp. 273–297.

The IoT Solution to Archive and Play the Digital Library of Kamishibai

Motohide Yoshimura, Ayumi Eikawa
Department of Information systems, Faculty of Information Systems
University of Nagasaki, Siebold Campus 1-1-1 Manabino, Nagayo-cho,
Nishi-Sonogi-gun, Nagasaki 851-2195, JAPAN

E-mail: yxsimura@sun.ac.jp, bs116004@sun.ac.jp
http://sun.ac.jp/e/

Abstract

We innovate an IoT solution to archive and play the digital library of Kamishibai. Kamishibai is a form of Japanese picture story show. The solution has two aspects. The one is a regional activation event to know history of the residential area, and the other is an educational platform to design the programming education by using NFC device. In this paper, we report the experimental results to build a display and sound system to play the digital library of Kamishibai.

Keywords: Kamishibai, Japanese picture story show, regional activation, programming education, IoT solution, near field communication

1. Introduction

Advances in information and communication technology have certainly improved the convenience of our lives. The household ownership ratio for smartphones such as iPhone and Android is 79.2%, and the grand total of mobile device ownership ratio is 95.7%. Overall, the percentage of Internet users from 13 to 59 years old exceeds 93%[1]. The role of Information Communication Technology (ICT) in tourism has become very important in the countryside vitalization. Here are some examples of it by ICT. At Tateyama-cho in Toyama Prefecture, various type of promotion are moved ahead for the global branding of the Tateyama Kurobe Alpine Route, in which mountains at altitude of 3,000m are high up[2]. In the Murodo Area located at the altitude 2,450m, an optical fiber cable has been laid and Wi-Fi Internet access become more convenient for the tourist. Live feed at the gazebo of Kurobe Dam by using web camera is delivered, and the tourist information such as a time table of the cable car and manners of national park usage, *etc.* is shown at the stations by using digital signage[3]. At Hiroshima City in Hiroshima Prefecture, the peace tourism around the Atomic Bomb Dome and Peace Memorial Park is promoted, so that the number of foreign tourists has been increasing year by year, since when Barack Obama has visited at the Peace Memorial Park. In that case, the fact that almost half of them return to cities such as Kobe City, Osaka City, and Kyoto City on a day trip is came out according to an analysis of tourist information of Hiroshima Free Wi-Fi users. Therefore, a round trip to various areas in Hiroshima Prefecture such as Miyajima island, Kure City, and Edajima City, *etc.* is planned to encourage tourists to stay in the prefecture[4]. Especially, through the stamp rally such as “Hiroshima Red Passport” is held from August 31th to December 25th, 2019. Stamp rally means going to different locations to collect stamps on a card. If you take part in the stamp rally, you can win prizes according to the number of stamps you collect![5] In Nagasaki City at Nagasaki Prefecture, the “Miracle Nagasaki Project” is being promoted to explore further appeal points of Nagasaki City and consider a new style of tourism, and the Wi-Fi Internet service has been provided to improve the convenience of tourists[6]. As an approach to barrier-free access to sightseeing facilities and spots, barrier-free information such as sidewalks (road steps, obstacles, *etc.*) is automatically collected based on the various sensors attached to a general wheelchair. The collected information is visualized on images and maps as the barrier-free street view web application[7,8].

Though these tourism services based on ICT will be an urgent issue in Japan, where the Tokyo Olympics are close at hand, the perspective after the Tokyo Olympics is also important. Making the further leap forward as the tourism-oriented country, it is important to create new tourism services based on flexible thinking. Considering tourism services as “things”, users are “tourists”, and tourism that users enjoy is “experiences”. Tourism functions in case that “things”, “tourists”, and “experiences” engage together, and tourism cannot uphold without the interaction between “things”, “tourists”, and “experiences”. Even in view of the current state of community-based new tourism, it is obvious that new kinds of tourism results from the regional activity. Tourism cannot be uphold without vitalizing the regional community.

Now, let's get into the main subject in this paper. What is the most important element among “things”, “tourists”, and “experiences”? It is perfectly obvious. The answer is “tourists”. Tourists use “things” and obtain “experiences”, and share “experiences” with another. Tourism must be centered on tourists. There are three essence for creating a new community-based tourism. The first is the innovation in tourism. The second is the sustainability in tourism. The last is the educational enhancement in regional community. Here, we make an emphasis on the last, and plan to organize a workshop which enhances the awareness of regional culture. How can we enhance the awareness of regional culture? Isn't it "accumulation of regional culture", "learning of regional culture", and "creation of new values"? In this paper, we apply Kamishibai to accumulating and learning of regional culture, and employ ICT to create new values. Kamishibai is a form of Japanese traditional picture story show and is a kind of play that a performer tells story while switching several sheets of pictures. In Omihachiman City at Shiga Prefecture, The guides belongs to tourism association produce and play Kamishibai. They conducts demonstrations of Kamishibai at elementary schools, so that the students can acquire knowledge of Omihachiman's history, culture, planning town, Omi merchants[9]. Kamishibai is utilized in programming workshops for junior high school students. In the workshops, the students create Kamishibai by using Scratch, a programming language[10]. In this study, we design a framework of traditional culture and ICT by blending Kamishibai and IoT, and build a display and sound system of Kamishibai experimentally. We report the experimental results to build a display and sound system to play the digital library of Kamishibai.

In Section 2, we describe the outline of the display and sound system of Kamishibai, and in Section 3, report the experimental results to build the system. In Section 4, we describe the conclusion.

2. Display and Sound System of Kamishibai

2.1. *What is Kamishibai?*

Kamishibai is a form of Japanese traditional picture story show. The beginning of Kamishibai is said to be a performance by the unemployed people in parks and at street corners where the children gathers during World Depression in 1929. Unemployed people set a miniature stage-like device on a carrier of bicycle and tells the story by switching the sheets of pictures. They sell the cheap sweets in order to gather children. The performer enhances the realism of picture story by communicating with the children. The impeccable interval of switching pictures deepens the concentration in the story. Kamishibai is getting noticed as an educational tools.

2.2. *Development Environment of Display and Sound System*

In our system, Kamishibai consists of several sheets of picture made by using drawing tools, and they are placed on a Web server. Each picture is coded as a web site by using HTML5 and is displayed on a monitor by switching each NFC tag in which URL information of each picture is written. In this study, we made each sheet of picture by using free illustrations. The story of Kamishibai is “Momotaro”, famous Japanese fairy tales. “Momotaro” consists of 12 sheets of pictures, so that the numbers of NFC tags to switch each sheet of pictures is

12. We also construct sound system by using raspberry pi 3 Model B and the "PaSoRi" RC-S380/S. NFC tag has an unique ID for every tag, we read and identify ID, make sounds according to the picture story The code to read and identify ID is written in Python.

2.3. *Architecture of Display and Sound System*

We explain the architecture of display and sound system to play the digital library of Kamishibai in Fig.1 Kamishibai consists of several sheets of picture on a Web server and is displayed on a device by switching the NFC tags in which URL information of each picture is written. Here we show the observe side of display system in Fig.2 and the reverse side of display system in Fig.3. As shown in Fig.2, we use android device to display Kamishibai and the device is affixed to cardboard with double-sided tape.

As shown in Fig.3, we set runners at upside and downside to slide the drawing paper to the right. Kamishibai performer is on the reverse side and are right-handed. Each NFC tag in which URL information of each picture on a Web server is affixed to the each drawing paper. In case of "Momotaro", 12 sheets of drawing paper to which each NFC tag is affixed are set at the runner.

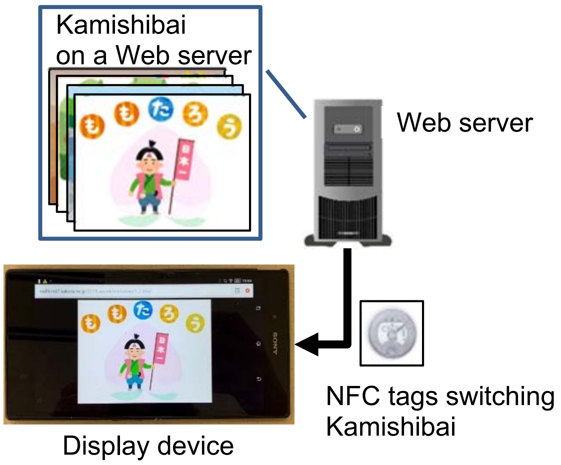


Fig. 1: Architecture of Display System.

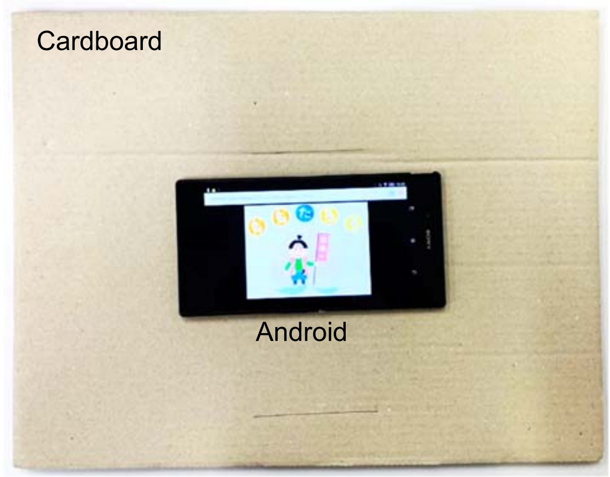


Fig. 2: Observe Side of Display System.

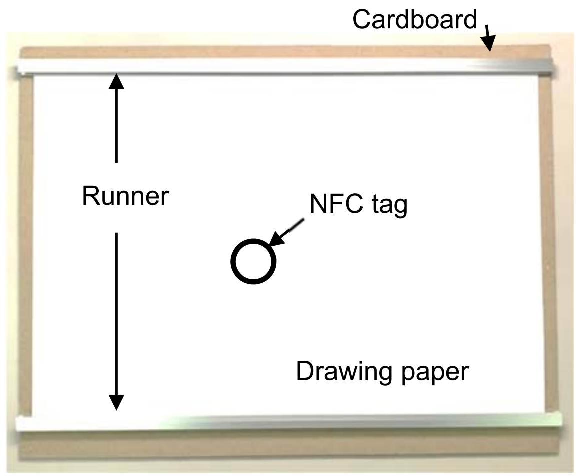


Fig. 3: Reverse Side of Display System.

3. Experimental Results to Build Display and Sound System

3.1. *Positional Relationship between NFC tag and NFC Reader*

3.1.1. *Thickness of Cardboard*

Between NFC tag and NFC reader, there is a cardboard and its thickness is 412g/m^2. We check whether read error is occurred or not according to the cardboard. We repeat the three sets of sliding 12 sheets of drawing paper, there is no read error and we can perform the full story of "Momotaro" by three times.

3.1.1. Position of NFC tag against NFC Reader

Fig.4 shows a cross-sectional view from left side.

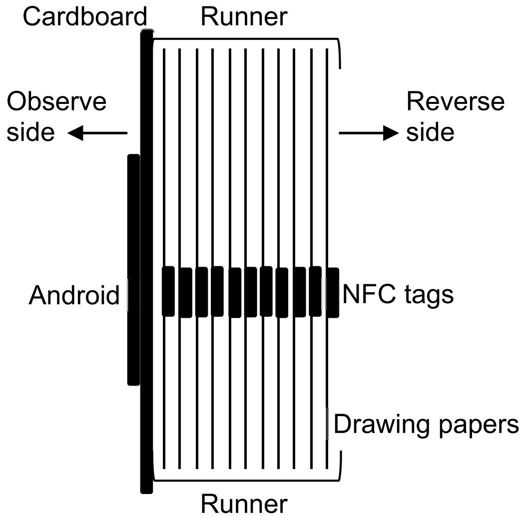


Fig.4: A Cross-sectional View from Left Side

In case that the position of the NFC reader of the android device and the position of the NFC tags is fully overlapped in line as to the performer's eye, error ratio is at 42% through sliding 12 sheets of drawing papers. In case that NFC tags are affixed at the position horizontally 2.3cm to the left, error ratio is at 0% through sliding 12 sheets of drawing papers.

3.2. *Making Sound by Unique ID of NFC tag*

We also construct sound system by using raspberry pi 3 Model B and the "PaSoRi" RC-S380/S. NFC tag has an unique ID for each tag, we read and identify the unique ID of each tag, make 12 kinds of sounds according to the story of "Momotaro". Fig.5 shows our sound system . At the Kamishibai performance, we can slide each sheet of drawing paper by right hand, and can make sound according to the story by left hand.

4. Conclusion

We innovate an IoT solution to archive and show the digital library of Kamishibai. In this study, we experimentally build the display system and sound system to show the digital library of Kamishibai. By using this system, we can perform Kamishibai without printing the sheets of pictures.

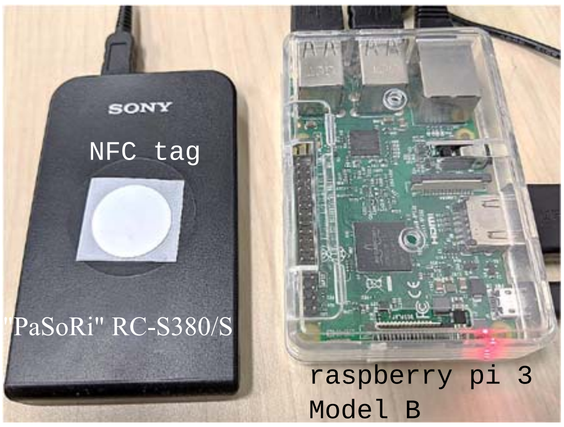


Fig.5: Sound System

References

1. Ministry of Internal Affairs and Communications, WHITE PAPER Information and Communications in Japan, 2019.
2. Nippon Telegraph and Telephone West Corporation, Case Study in ICT Applications, Retrieved December 19th, 2019, from https://www.ntt-west.co.jp/ict/casestudy/tateyama.html (in Japanese).
3. Susunaka Tsushin Kogyo Inc., Case Study in ICT Applications, Retrieved December 19th, 2019, from http://www.sasunaka.co.jp/records/index.cgi?pg=0040 (in Japanese).
4. Hiroshima Convention Bureau, Hirotabi, Retrieved December 19th, 2019, from https://www.hiroshima-navi.or.jp/ (in Japanese).
5. RCC Bunka Center Inc. 2019 Hiroshima Red Passport, Retrieved December 19th, 2019, from https://www.rccbc.co.jp/event/news/20190726.html (in Japanese).
6. Nihon Unisys, Ltd. MIRACLE NAGASAKI PROJECT, Retrieved December 19th, 2019, from https://www.unisys.co.jp/solution/biz/regional-revitalization/platform/nagasaki/ (in Japanese).
7. Chisato Miura, *et.al.*, Proposal of Slope Information Provision Method for Barrier Free Street View System, *The IPSJ Transactions on Consumer Devices and Systems*, vol.9, no.1, pp.11-21, 2019 (in Japanese).
8. Kenichi Arai, *et.al.*, Social Barrier-free Street View System, *The Journal of Information Processing*, vol.60, no.3, pp.821-829, 2019 (in Japanese).
9. Association of tourism volunteer guide of Omihachiman, Hidetsugu San to Omihachiman, Retrieved December 19th, 2019, from https://www.omi8guide.com/kanko-kamishibai/ (in Japanese).
10. Motoyuki Ozeki, Practice and Trends of Computer Programming Workshop for Junior High School Students, The Bulletin of Mukogawa Women's University. Humanities and Social Sciences, vol.65, pp.37-44, 2017 (in Japanese).

A Research on Prediction of Inter-firm Relationships

Takao Ito
Graduate School of Engineering, Hiroshima University
1-4-1 Kagamiyama, Higashi Hiroshima, 739-7527
E-mail: itotakao@hiroshima-u.ac.jp
Tsutomu Ito
Hamura Plant, Hino Motors, Ltd.,
5-2-1 Midorigaoka, Hamura City, Tokyo, 205-8601, Japan
E-mail: fw.eldorado.500cuin@gmail.com
Rajiv Mehta
Martin Tuchman School of Management, New Jersey Institute of Technology, University Heights,
Newark, New Jersey, 07102-1982, U.S.A.
E-mail: kumano@cc.miyazaki-u.ac.jp
Seigo Matsuno
Department of Business and Administration, NIT, Ube College,
2-14-1, Tokiwadai, Ube, Yamaguchi, 755-0096, Japan
E-mail: matsuno@ube-k.qac.jp
Makoto Sakamoto
Faculty of Engineering, University of Miyazaki,
1-1 Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan
E-mail: sakamoto@cc.miyazaki-u.ac.jp

Abstract

The formation of successful corporate strategy among vertically integrated organizations is predicated upon a plethora of dynamic relationships among buyer-seller firms. Albeit growing in interest, there is a paucity of research that focuses on predicting the nature of inter-firm relationships among Japanese keiretsu members. Discerning this knowledge can help formulate successful strategic decisions, such as whether to continue procuring parts or increasing future investments in partner firms. As studies on prediction of inter-firm relationships have been sparse, this research offers insights by showing a newer approach for making decisions regarding corporate strategy. Using data collected on transaction from Yokokai—parts suppliers of Mazda—in this study a new model is developed to forecast inter-firm relationship using degree index and DEA model. The main contribution of this research shows that prediction of inter-firm relationship is not only depends on future trends, but also determined by the interactive relationships among member firms of a keiretsu.

Keywords: out-degree, rate of change, Inter-firm relationship, cross-sectional interaction, Yokokai.

1. Introduction

In vertically integrated arrangements, formulation of successful corporate strategy is predicated upon a plethora of dynamic interrelationships among member organizations, such as frequently engaging in inter-firm transactions, cross-firm investments, and jointly coalesce towards attaining collective, supra-organizational goals. Although there is growing interest, research on these vertically integrated organizational forms has been fragmentary. Indeed, research that focuses on predicting the nature of inter-firm relationships among Japanese keiretsu members is scant. Discerning this knowledge can direct determine strategic decisions, such as whether to continue procuring parts or increasing future investments in partner firms. Accordingly, this study makes a contribution to the extant literature by illustrating a novel integrative forecasting approach for predicting the state of inter-firm relationships, which is not only dependent on future economic conditions, but also determined by the interactive relationships among keiretsu member organizations. This manuscript is structured as follows. Section 2 reviews the relevant literature on forecasting methods and inter-firm relationships. In Section 3, the model building procedure and process for predicting inter-firm relationships using our new proposed method is explicated. In Section 4, the application of the integrative forecasting approach is exemplified. In Section 5, the results of this empirical study using the data drawn from Yokokai, Mazda's Keiretsu, is compared and contrasted to validate our new proposed integrative forecasting method vis-à-vis traditional prediction methods and approaches. Section 6 identifies limitations and offers directions for future research.

2. Background

As well-known organizational structures, Japanese keiretsu typically comprise a family of inter-firm cooperative arrangements among member automotive manufacturers and their component parts suppliers. It is widely recognized that the success of corporate management of these inter-family structured business alliances depends to a large extent on inter-firm relationships among partner firms. A plethora of studies on inter-firm relationships have been published. For instance, Bensaou illustrated a framework for managing a portfolio of relationships among partner firms. More specifically, he (a) proposed the type of relation design a firm should choose under different external contingencies, and (b) what is the appropriate way to manage different types of relationships [1]. In another investigation, Mathijs De Vaan clarified how the composition and stability of network ties affect firm performance in industries characterized by rapid technological change [2]. All of these research studies focused on the inter-firm relationships using past data sets. While offering explanations to past occurrences is important, forecasting the future developmental trends of inter-firm relationships would be even more salient. Indeed, businesses place tremendous importance on the prediction in different fields, such as weather forecasting, and stock price movements [3]. In the context of keiretsu, as studies on prediction of interrelationships among inter-firm partners have been sparse, this research offers insights by showing a newer approach for making decisions regarding corporate strategy. Using data collected on transaction from Yokokai—parts suppliers of Mazda—in this study a new model is developed to forecast inter-firm relationship using graph theory.

3. Model Building

Generally, all forecasting methods, which include moving average and ARMA models, have focus on future trends and rate of change for time series data depending on past values of the variable being forecasted and on past prediction errors. Moreover, the input data sets are basically independent with each other. However, in case of network organizations, all indexes are connected with each other; thus, the criterion variable will vary if one of the indexes, such as degree and influence change [4-5]. Consequently, the prediction of inter-firm relationships basically comprises two aspects. The first is future trends, which may exhibit increasing or decreasing values. And, the second is the rate of change for time series data. Thus, future trend will be determined by not only time series change, but also cross-sectional interaction of inter-firm relationship. In this paper, time series change can be measured by least square method, and future trend will be identified by using comparative study between different term's efficiency. That is indicative of a strong incentive to develop next transaction if the t+1 term' efficiency is larger than its previous term. Therefore, to calculate relative efficiency another forecasting method—the DEA procedure—is proposed. Cross-sectional interaction can be calculated using graph theory, such degree, and influence of inter-firm relationships. Accordingly, the procedure can be shown as below.

1. Collect adjacent matrix data from Keiretsu.
2. Calculate network indexes such as degree using graph theory. In asymmetric network, degree will be divided into out-degree and in-degree. They will used as input data sets in DEA model. Degree will be calculated as below.

$$\mathrm{Deg} = \sum_{i}^{n} a(p_i, p_k) \tag{1}$$

3. Put the network indexes as input data set into DEA model with output data set. The generalized model of the CCR model is formulated as follows:
   Objective function

$$\max \ \theta = \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \tag{2}$$

Subject to

$$\frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1 \tag{3}$$

$$u_r \geq 0 \ \ (\mathrm{r} = 1, 2, \cdots, \mathrm{s}) \tag{4}$$

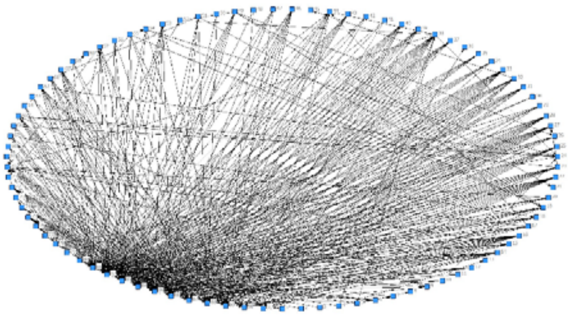


Fig. 2. An Example of Transaction of Yokokai in 2004.

$$v_i \geq 0 \quad (\mathrm{i} = 1,2,\cdots,\mathrm{m}) \qquad (5)$$

4. Select specific firm from the keiretsu and forecast the value of inter-firm relation using regression model.
5. Calculate the rate of change of the specific firm. The rate of change *r* will be calculated as below.

$$\mathrm{r} = \frac{Deg_{t+1} - Deg_t}{Deg_t} \leq 1 \qquad (6)$$

   where
   Deg_{t+1}……predicted value of degree at t+1 term
   Deg_t……..existed value of degree at t term
6. Input the degree index as input data into DEA model, and calculate the relative efficiency (θ_0).
7. Determine the transaction firms of the specific firm, and predict the value of inter-firm relation using the rate of change for each firms.
8. Calculate degree index again, and input the degree as input data sets of DEA model; calculate the efficiencies of each firms: $\theta_1, \theta_2, \ldots \theta_n$.
9. Determine the direction of the transaction firms, to keep going up if $\theta_i \geq \theta$; Otherwise keep going down.
10. Predict the values of inter-firm relationship.
11. Check the difference between prediction and observed value, find the reason if the big gap exists.

4. An Example

Suppose a five node network shown in Fig. 1 is given [4].

Based on degree equation, in-degree and out-degree will be calculated as below.

Table 1. This is the caption.

	Out-degree	In-degree
A	20	90
B	135	35
C	115	55
D	80	45
E	55	180

The relative efficiency of the five firms are 0.714, 1, 0.322, 0.543, 1 respectively when the sales revenue of the five firms are 100, 90, 250, 300, and 400 respectively. Suppose the regression model of out-

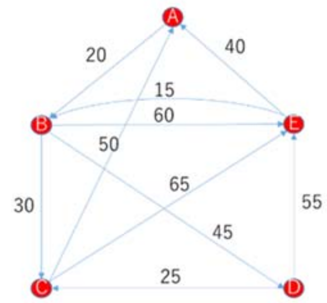


Fig. 1. An Example of Five-node Network.

degree of node D is y=20x-100 based on data sets of 9 historical term. Thus, the out-degree of B in next term will be 100. The rate of change is 0.25. Consequently, the relationships with C, and E will be 31.25, and 68.75, respectively. The new value of inter-firm relation of D-C and D-E are 0.505, and 0.482, respectively. They are less than θ_0=0.543; therefore, the direction of the future trend will be negative. The prediction value of D-C and D-E will be 18.75, and 41.25. It means that the relationship between D and C, and D and E will be weaker than before.

5. Empirical Studies

In order to predict the relationship of the keiretsu, data for the duration 2004 to 2008 was obtained from Yokokai, a network organization of the automotive parts suppliers. The network of Yokokai in 2004 is illustrated as in Fig. 2.

To predict the inter-firm relation in 2008, the out-degree of Hiroshima Aluminum are calculated.

Table 2. Out-degree of Hiroshima Aluminum.

	Out-degree
2004	76.5
2005	76
2006	79
2007	78.8

The relative efficiency of Hiroshima Aluminum is 0.04. The prediction value of out-degree in 2008 will be calculated as 80.05 using regression model. The observed value in 2008 is 81.2. The rate of change will be calculated as 1.58%. There are four companies have transactions with Hiroshima Aluminum in 2007. They are Mazda, Honda, Jatco, and Aishin AW. This rate is used to calculate the relationship and efficiency. The new efficiencies of Mazda, Honda, Jatco, and Aishin AW are 0.039, 0.092, 0.088, and 0.101, respectively. Thus, only the efficiency with Mazda is less than the former. It indicates that the future trend of the relationship with Mazda will be negative. The results of prediction and efficiency can be shown in Table 3 below.

Table 3. Prediction Results and Observed Data.

	Prediction	Observed Data Set
Mazda	53.04	48
Honda	8.84	6.9
Jatco	10.46	11.5
Aishin AW	5.99	14.8

Basically, the traditional approach of prediction is only to calculate the trend based on time series change. To compare the traditional approach with our newer proposed method, the prediction of inter-firm relationships is conducted. The value of inter-firm of Hiroshima Aluminum with Mazda, Honda, and Jatco are 54.7, 7.2, and 11.25 respectively using the regression model without considering the interaction among firms to predict. The relationship with Aishin AW cannot be calculated because of lack of data. Using the observed data, prediction using our proposed method, versus prediction without considering the interaction is illustrated in Fig. 3.

Prediction 1 indicates the proposed method, which is compared with the results of Prediction 2 (or the traditional approach) in Fig. 3. All other firms exhibit similar results. Similarly, we selected KYB and compared the results using the same method, which is reported in Fig. 4.

This study corroborates that the result of our newer proposed method is much more accurate than the traditional approach. Accordingly, the results obtained in this empirical investigation find validation and support of our integrative forecasting method.

6. Conclusion

In this paper, a new method using two integrative forecasting techniques was proposed. Using transaction data drawn from Mazda's Yokokai, the value of inter-firm relationship of Hiroshima Aluminum and KYB were selected and predicted. The results reported show validation that our newer method is empirically tested and supported.

A limitation of this paper, however, is that only one network index—degree—is used. Additional network indices, such as influence and closeness, should also be investigated to test the validity of the newer forecasting model. Furthermore, in future research investigations not only transactional data, but also cross-shareholdings among member firms, should be employed.

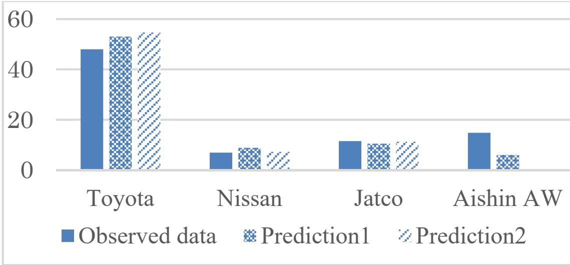


Fig. 3. Result of Prediction 1 and Prediction 2.

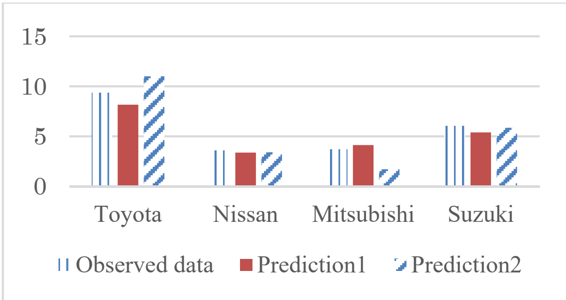


Fig. 4. An Example of Five-node Network.

References

1. Bensaou, Portfolio of buyer-supplier relationships, Sloan management Review, summer 1999.
2. Mathijs de Vaan, Interfirm networks and firm performance in the face of technological discontinuities, the DIME-DRUID Academy winter conference 2011, Comwell Rebild Bakker, Aalborg, Denmark, January 20-22, 2011.
3. Forecasting, https://en.wikipedia.org/wiki/Forecasting, retrieved November 29, 2019.
4. Tsutomu ito, The influence of inter-firm relationships on corporate performance, Scholars' Press, 2018.
5. T. Ito, M. Sakamoto, R. Mehta, and S. Ikeda: Examining Inter-Firm Relationships Using Network-Based DEA: A Case Study of Mazda Yokokai, International Journal of Mathematical Models and Methods in Applied Sciences, Volume 8, 2014, pp. 516-521, November 2014.

Generation of Arbitrarily-Oriented Ripple Images Using Circular-Sector-Type Smoothing Filter and Inverse Filter

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

Abstract

A non-photorealistic rendering (NPR) method for generating ripple images from photographic images has been proposed using intensity gradient. Ripple patterns imitate a wave on the water surface and are composed of continuous lines with fluctuations. Ripple images are expressed by superimposing ripple patterns on photographic images. However, the conventional method can only generate vertical and horizontal ripple patterns. Therefore, in this paper, we develop a method that can generate ripple patterns in any orientation. Ripple images generated by the proposed method are called arbitrarily-oriented ripple (AOR) images. The proposed method is executed by an iterative process using circular-sector-type smoothing filter and inverse filter. To verify the effectiveness of our method, we investigate the changes in AOR images by changing the values of the parameters.

Keywords: non-photorealistic rendering, ripple pattern, arbitrarily orientation, circular-sector-type smoothing filter, inverse filter

1. Introduction

NPR is a technology of computer graphics that can generate effective illustrations and artistic images. Many NPR methods have been proposed to simulate effective illustrations and artistic images[1,2,3,4,5]. In the past NPR methods, an NPR method for generating ripple images has been proposed[6]. Ripple patterns imitate a wave on the water surface and are composed of continuous lines with fluctuations. Ripple images are expressed by superimposing ripple patterns on photographic images. The conventional method was conducted by an iteration process with intensity gradients. Although the conventional method can generate ripple patterns in horizontal and vertical orientations, it is not possible to generate ripple patterns with arbitrary slope.

In this paper, we develop a method that can arbitrarily change the orientation of ripple patterns, then expand the range of expression of the conventional method. Ripple images generated by the proposed method is called AOR images. The proposed method is executed by an iterative process using inverse filter[7,8] and circular-sector-type smoothing filter. By adjusting the value of the parameter in the proposed method, ripple patterns of the proposed method is more linear or wavy than ripple patterns of the conventional method. In addition, by changing the value of the parameter, the proposed method can also change the interval between ripple patterns as the conventional method. To verify the effectiveness of the proposed method, experiments are conducted using Lenna image and other photographic images. As a result of the experiments, it is clarified that the proposed method can automatically generate AOR images. In addition, through experiments that change the values of parameters in the proposed method, it is also revealed the changes in appearance of AOR images.

The rest of this paper is organized as follows. Section 2 describes the proposed method for generating AOR images. Section 3 shows experimental results, and reveals the effectiveness of the proposed method. Finally, Section 4 concludes this paper.

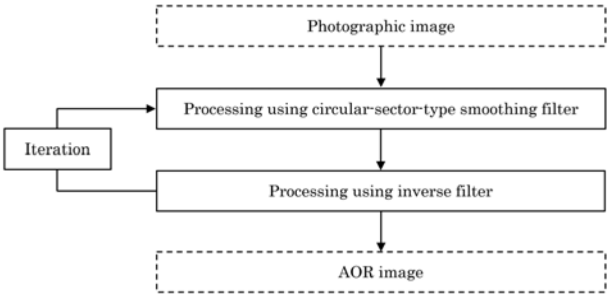


Fig. 1. Flow chart of the proposed method.

2. Proposed Method

The proposed method is executed in two processes. In the first process, images are smoothed using circular-sector-type smoothing filter. In the second process, the smoothed images are restored using inverse filter. By repeating the two processes, AOR images are generated. Accumulation of errors due to the iterative process is a factor that can generate ripple patterns. A flow chart of the proposed method is shown in Fig. 1. The detailed procedure in Fig. 1 is shown as follows.

The input pixel values for spatial coordinates (i, j) of a photographic image are defined as $f_{i,j}$. Then, the pixel values of the image at the t-th iteration number are defined as $f_{i,j}^{(t)}$, where $f_{i,j}^{(1)} = f_{i,j}$. The pixel values $f_{i,j}^{(t)}$ have value of M gradation from 0 to $M-1$.

In the first process, the image with the pixel values $f_{i,j}^{(t)}$ is smoothed using circular-sector-type smoothing filter that calculates an average of pixel values in circular sector as shown in Fig. 2. The circular sector with a range of θ_1 radian and radius D rotates θ_2 radian, and the pixel values that the center of the pixel is included in circular sector at the target pixel (i, j) are $g_{i,j,k}^{(t)}(k = 1,2,\cdots,K_{i,j}^{(t)})$. The pixel values $s_{i,j}^{(t)}$ of the smoothed image are calculated as the following equation.

$$s_{i,j}^{(t)} = \frac{\sum_{k=1}^{K_{i,j}^{(t)}} g_{i,j,k}^{(t)}}{K_{i,j}^{(t)}} \quad (1)$$

In the second process, the pixel values $f_{i,j}^{(t+1)}$ using inverse filtering are calculated as the following equation.

$$f_{i,j}^{(t+1)} = f_{i,j}^{(t)} - s_{i,j}^{(t)} + f_{i,j} \quad (2)$$

In case $f_{i,j}^{(t+1)}$ is less than 0, then $f_{i,j}^{(t+1)}$ must be set to 0. In case $f_{i,j}^{(t+1)}$ is greater than $M-1$, then $f_{i,j}^{(t+1)}$ must be set to $M-1$. An AOR image is obtained after the first and second processes of T times iteration.

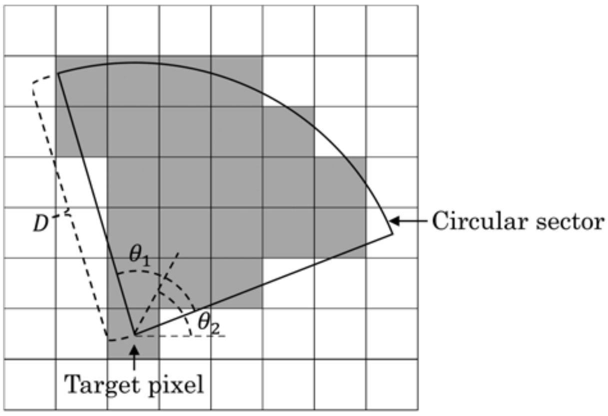
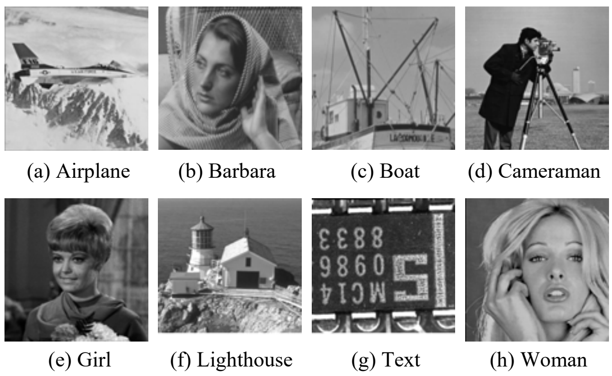


Fig. 2. Conceptual diagram of circular-sector-type smoothing filter.

Fig. 3. Lenna image.



(a) Airplane (b) Barbara (c) Boat (d) Cameraman

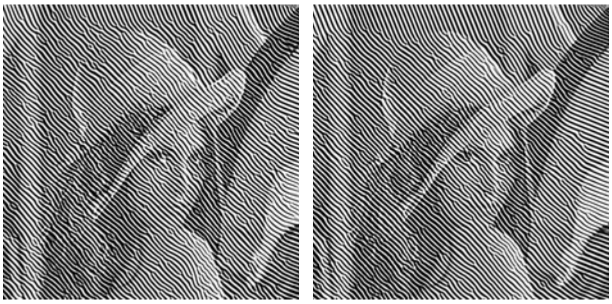
(e) Girl (f) Lighthouse (g) Text (h) Woman

Fig. 4. Various photographic images.

3. Experiments

We mainly conducts two experiments. First, the experiment with changing the values of the parameters in the proposed method is conducted using Lenna image shown in Fig. 3. Second, the experiment is conducted to verify that patterns can be generated using various photographic images shown in Fig. 4. All photographic images used in the experiments are 512 * 512 pixels and 256 gradation.

(a) $T = 10$ (b) $T = 20$



(c) $T = 50$ (d) $T = 100$

Fig. 5. AOR images in the case of the iteration number $T = 10{,}20{,}50$, and 100.

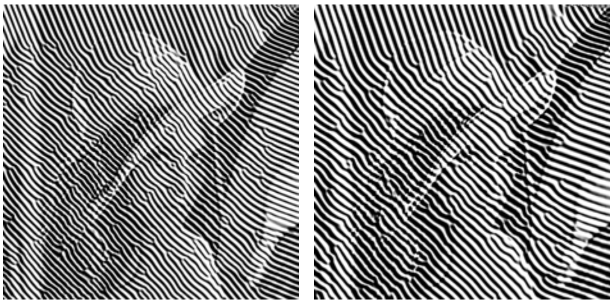
3.1. *Experiment with changing parameters*

AOR images by changing the iteration number T are visually confirmed using Lenna image. The iteration number T is set to 10, 20, 50, and 100. Other parameters D, θ_1, and θ_2 are set to $5, \pi/2$, and $\pi/4$, respectively. The results of the experiment are shown in Fig. 5. As the value of the iteration number T is larger, ripple patterns become clear and converge.

AOR images by changing the radius D are visually confirmed using Lenna image. The radius D is set to 3, 5, 7, and 9. Other parameters T, θ_1, and θ_2 are set to 100, $\pi/2$, and $\pi/4$, respectively. The results of the experiment are shown in Fig. 6. As the value of the radius D is larger, the intervals between ripple patterns become wider.

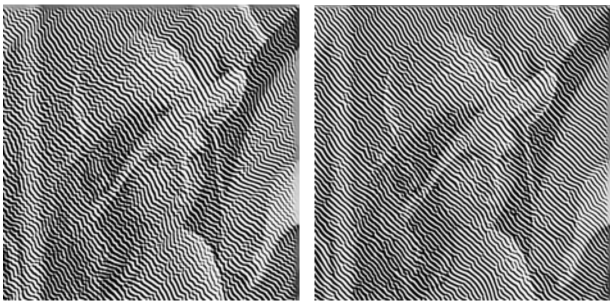
AOR images by changing the angle θ_1 are visually confirmed using Lenna image. The angle θ_1 is set to $\pi/6$, $\pi/3$, $\pi/2$, and $(2/3)\pi$. Other parameters T, D, and θ_2 are set to 100, 5, and $\pi/4$, respectively. The results of the experiment are shown in Fig. 7. As the value of the angle θ_1 is larger, ripple patterns become more linear. On the other hand, as the value of the angle θ_1 is larger, ripple patterns become wavier.

(a) $D = 3$ (b) $D = 5$

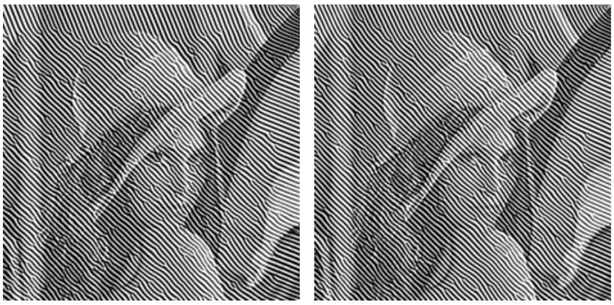


(c) $D = 7$ (d) $D = 9$

Fig. 6. AOR images in the case of the radius $D = 3{,}5{,}7$, and 9.

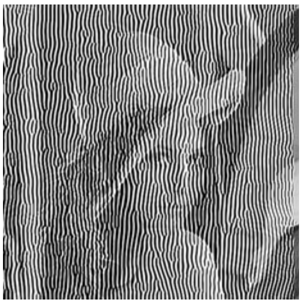


(a) $\theta_1 = \pi/6$ (b) $\theta_1 = \pi/3$

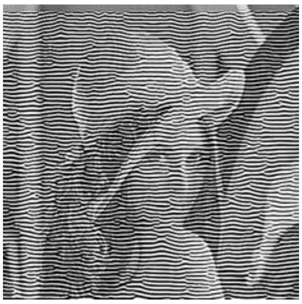


(c) $\theta_1 = \pi/2$ (d) $\theta_1 = (2/3)\pi$

Fig. 7. AOR images in the case of the angle $\theta_1 = \pi/6\,, \pi/3\,, \pi/2$, and $(2/3)\pi$.




(a) $\theta_2 = 0$ (b) $\theta_2 = \pi/4$



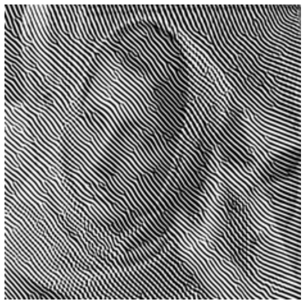

(c) $\theta_2 = \pi/2$ (d) $\theta_2 = (3/4)\pi$

Fig. 8. AOR images in the case of the angle $\theta_1 = 0, \pi/4, \pi/2$, and $(3/4)\pi$.

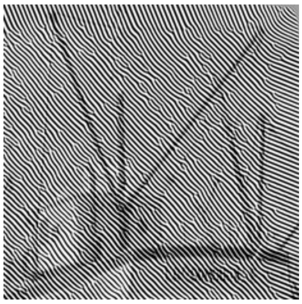
AOR images by changing the angle θ_2 are visually confirmed using Lenna image. The angle θ_2 is set to 0, $\pi/4$, $\pi/2$, and $(3/4)\pi$. Other parameters T, D, and θ_1 are set to 100, 5, and $\pi/2$, respectively. The results of the experiment are shown in Fig. 8. Depending on the value of the angle θ_2, the orientation of ripple patterns is changing.

3.2. *Experiment using various photographic images*

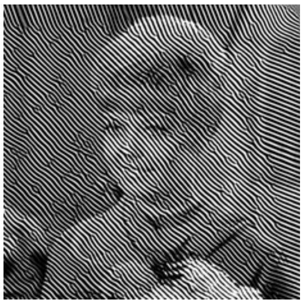
The proposed method is applied to eight photographic images shown in Fig. 4. The parameters T, D, θ_1, and θ_2 are set to 100, 5, $\pi/2$, and $\pi/4$, respectively. The results of the experiment are shown in Fig. 9. In all cases, ripple patterns can be automatically generated on the whole image. However, no ripple patterns are generated in the white areas shown in Fig. 9 (f) and (g). It is conceivable to generate a ripple pattern by adding noise to the white area.

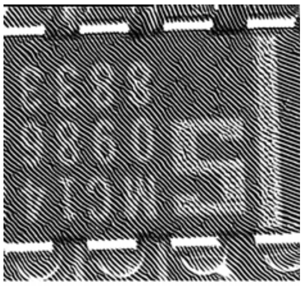
(a) Airplane (b) Barbara




(c) Boat (d) Cameraman




(e) Girl (f) Lighthouse




(g) Text (h) Woman

Fig. 9. AOR images.

4. Conclusion

We proposed an NPR method for generating AOR images from photographic images. The proposed method was executed by an iterative process using inverse filter and circular-sector-type smoothing filter. To verify the effectiveness of the proposed method, the changes in AOR images by changing the values of the parameters were investigated. As a result of the experiments, by changing the value of the angle θ_2 in the proposed method, ripple patterns could be generated in any orientation. And, by changing the value of the angel θ_1, ripple patterns could be generated more linear or wavy. And, by changing the value of the radius D, the proposed method could also change the interval between ripple patterns.

A subject for future study is to expand the proposed method for application to color photographic images and videos.

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.
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. W. Qian, D. Xu, K. Yue, Z. Guan, Y. Pu and Y. Shi, Gourd pyrography art simulating based on non-photorealistic rendering, *Multimedia Tools and Applications.* **76**(13), 2017, pp. 14559-14579.
4. D. Martin, G. Arroyo, A. Rodriguez and T. Isenberg, A survey of digital stippling, *Computers & Graphics.* **67**, 2017, pp. 24-44.
5. W. Qian, D. Xu, J. Cao, Z. Guan and Y. Pu, Aesthetic art simulation for embroidery style, *Multimedia Tools and Applications.* **78**(1), 2019, pp. 995-1016.
6. T. Hiraoka, Generation of ripple images using intensity gradient, *ICIC Express Letters.* **13**(1), 2019, pp. 1097-1101.
7. J. M. Ortega and W. C. Rheinboldt, Iterative solutions of nonlinear equations in several variables, *Society for Industrial Mathematic.* 1987.
8. Yu and K. Urahama, Iterative method for inverse nonlinear image processing, *IEICE Transactions on Fundamentals.* **E97-A**(2), 2014, pp. 719-72.

Modelling Autonomous Parallel Parking Procedure for Car-like Robot Avrora Unior in Gazebo Simulator

Dinir Imameev[1], Ksenia Shabalina[1], Artur Sagitov[1], Kuo-Lan Su[2], Evgeni Magid[1]
[1]Laboratory of Intelligent Robotic Systems (LIRS), Intelligent Robotics Department, Higher Institute for Information Technology and Intelligent Systems, Kazan Federal University, 35 Kremlyovskaya street, Kazan, Russia Federation
[2]Department of Electrical Engineering, National Yunlin University of Science and Technology, Taiwan
E-mail: dinir.imameev@gmail.com, ks.shabalina@it.kfu.ru, sagitov@it.kfu.ru, sukl@yuntech.edu.tw, magid@it.kfu.ru
http://kpfu.ru/robolab.html

Abstract

This paper focuses on implementation of path planning and control for Avrora Unior robot car that enable autonomous parallel parking. Path planning is based on existing geometrical approach, which was modified to fit specific kinematics of the robot shape and control geometry. Geometry and parking space size determine path key points: steering and counter-steering points. We implemented and tested the algorithm using Avrora Unior robot model in Gazebo simulator.

Keywords: autonomous parallel parking, Avrora Unior robot car, car-like robot, Gazebo, algorithm, modelling.

1. Introduction

Parallel parking is a hard problem for amateur drivers to perform and in some cases is challenging even for the experts. The parking process involves search for an empty spot and executing series of maneuvers in such a way that places a car in this spot without accidents, which are highly probable at narrow parking spots. Automobile manufacturers are independently developing specialized solutions (known as parking assistants[1]) to help driver in those situations, e.g., Toyota's Intelligent Parking Assist, Ford's Active Park Assist, BMW's Parking Assistant, and Mercedes's Active Parking Assist.

We are working on a similar problem for a smaller fully autonomous vehicle by adapting existing algorithm to suit car-like robot constraints. Our work is based on Avrora Unior robot. The paper discusses parking path planning[2] and control tested for a simulated Avrora Unior robot, but our solution is easily transferable to a real Avrora Unior and other car-like robots[3].

The paper is organized as follows: Section II describes car-like Avrora Unior robot and its relevant characteristics, Section III outlines generalized parking problem for non-holonomic vehicle; Section IV covers necessary modifications that were made to the existing robot model[4]; Section V covers updates of the robot's control module. In Section VI we discuss experimental setup and results. Finally, we conclude in Section VII.

2. Avrora Unior robot

Car-like Avrora Unior mobile robot (Fig. 1) was developed by Russian company Avrora Robotics. Robot's main purpose was to become a main teaching platform for mobile robotics students, which affected the composition of its sensors. Relevant linear measurements for Avrora Unior robot are listed in the Table I.

Avrora Unior robot is driven by two rear wheels (each featuring individual DC brushless motor coupled with a fixed ratio gearbox) and steered by separate DC motor that turns the two front wheel linkage. Design of steering linkage is based on a Ackermann steering geometry[5]. Microsoft Kinect camera is installed in the front section. We also requested manufacturer to install Hokuyo laser rangefinder mounting plate in the front upper part of the robot. Avrora Unior is equipped with off-the-shelf parking sensor utilizing 8 ultrasonic transducers to determine the distance to obstacles; 4 are mounted on the front and 4 on the rear bumper.

Table 1. Avrora Unior's Linear Dimensions.

Dimension	Dimension (mm)
Rim diameter	205
Wheel diameter	260
Wheel width	104
Wheelbase	700
Axle track	540
Body length	1120
Body height	570
Body width	650

Fig. 1. Avrora Unior robot.

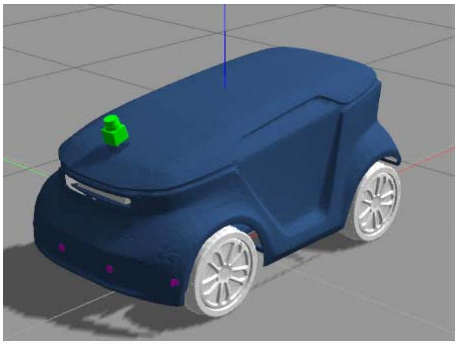


Fig. 2. Avrora Unior model in Gazebo.

To prevent unnecessary collisions and other accidents at this stage of research we are using robot model in a Gazebo[6] simulation that was developed previously[4] (Fig. 2). The model features and all sensors entirely replicate physical characteristics of the real Avrora Unior robot.

3. Problem specification

Adaptation of parallel parking algorithm for Avrora Unior robot involves solving two major issues. The first issue is implementation of non-holonomic control for Avrora Unior robot.

A. Non-holonomic vehicle control

To implement control for four-wheel mobile robot with rear driven wheels and front steering wheels we introduce two control variables (θ is a steering angle, v is a forward velocity); thus configuration space has 3 dimensions (x, y, θ). The vehicle motion is described through the following set of equations:

$$\begin{cases} \dot{x} = v\cos\phi\cos\theta \\ \dot{y} = v\cos\phi\sin\theta \\ \dot{\theta} = \frac{v}{L}\sin\phi \end{cases} \tag{1}$$

Equations (1) represents a system of non-holonomic constraints because they include derivative of vehicle's coordinates variable and are not integrable[7].

B. Parking Path Planning

For parallel parking path planning we modified path planning method based on geometrical approach[8]. This planning method can be used to precalculate all necessary trajectories for a given environmental parameters (distance to parked cars, length of available space), so we can generate trajectories for most common situations.

Before parking, the robot calculates a set of key points where steering and counter-steering trajectories should be executed. Calculation of those key points is based on the following vehicle parameters: maximum steering angle (β), wheelbase (e), front and rear overhang (p), body width (w) and vehicle length (L)[8]. A trajectory for a single try parallel parking could be calculated only if available parking spot length is larger than Lmin[8]; otherwise more than one trial (series of steering-countersteering trajectories) are calculated to successful park a vehicle.

4. Algorithm implementation

This work is based upon implementation of geometry-based algorithm that have been made available by Rohith Krishnan[9]. However, due to a critical difference in selection of a target coordinate frame used in calculations of key points (centered not at the vehicle center, but at the rear axle center) the original package did not generate appropriate trajectories for Avrora Unior parking. Nevertheless, the original paper formulas were successfully adjusted in order to resolve this issue.

4.1. Robot Control

The algorithm was adapted to comply with the control system of Avrora Unior model. Implemented parallel parking module tracking vehicle's position in relation to calculated key points. Changes in key points calculation were only subtractions of displacement about OX from a rear axle location. More complex transformation were made for calculation of counter-steering point because it includes both rotation and translation.

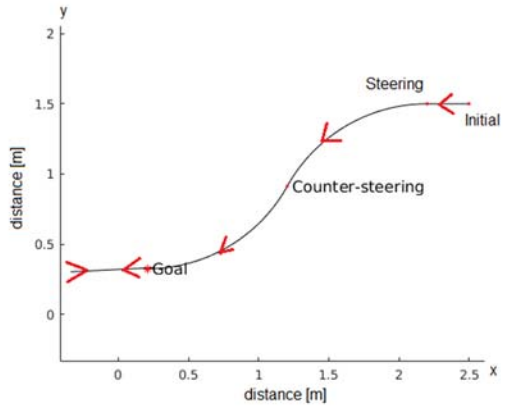


Fig. 3. Parking trajectory.

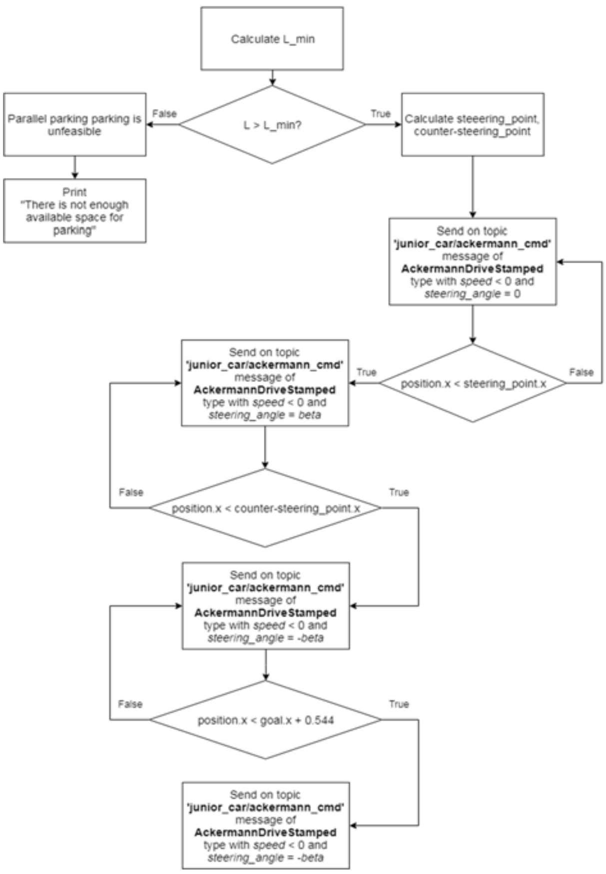


Fig. 4. Control algorithm diagram.

As Avrora Unior robot (and subsequently it's model) does not feature active braking system, we stop it in advance to the target point using only engine braking. We made measurements that determine the time required for a full stop and time it takes to turn steering for a full interval[10], and used those measurements to implement the final control algorithm, which is presented in Fig. 4.

4.2. Key Points

The first step is calculating a counter-steering point for the center of the vehicle. For calculation of a new counter-steering point, the coordinate frame {1} of the old one is used. The coordinate frame in the center of counter-steering point 1 is rotated by angle β relative to global coordinate frame {0} (a steering angle)[8] and further displaced by e/2 (Fig. 5). A separate ROS service was implemented for calculation of angle β from an input steering angle. The diagram of vehicle control architecture in Gazebo simulation is presented in Fig. 6.

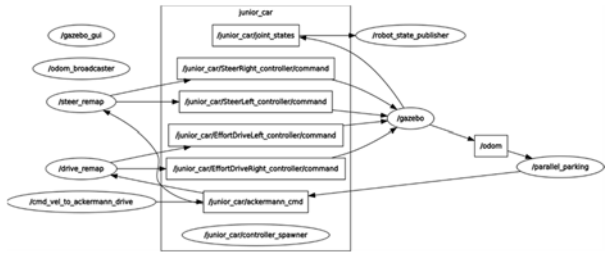


Fig. 5. Parking control system.

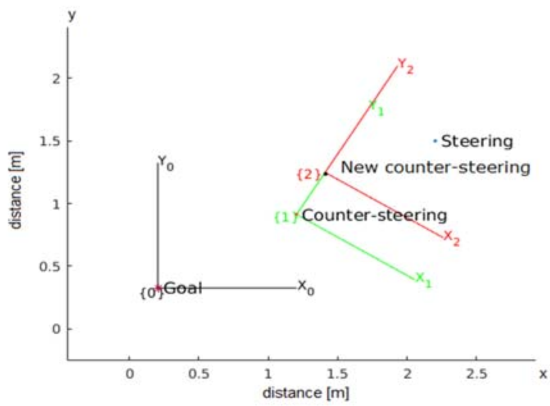


Fig. 6. Coordinate frames.

5. Avrora Unior model dynamic characteristics

We measured time required for a full stop by performing a series of experiments. After robot was accelerated to a predefined velocity (parking velocity) we engage a passive breaking (by sending ROS[10] AckermannDriveStamped message with parameters speed = 0 and steering angle = 0) to 'junior car/cmd' velocity control topic (Fig. 6). Time that takes the vehicle to stop starting from the moment of message being sent and ending when the vehicle reaches velocity 0, is taken as a full stop time. In average it takes about 40 seconds for a full stop. Measuring required time for full steering was performed while the vehicle was suspended. In this experiment AckermannDriveStamped message with speed = 0 and steering angle = β is sending onto topic 'junior car/cmd'. Three seconds are required for a full steering.

6. Experimental validation

Parking experiments were performed in a completely known environment. A virtual world[11] experiment consisted in defining an initial position of the vehicle and selection of a parking spot with a goal position.

For example, we describe sample experiment where a point (0.206, 0.325) is considered as a goal position and (2.5, 1.5) — as an initial one. Area of a parking spot is easily calculated for a goal position (for the rear axle) using p = 0.206 m, w = 0.65 m and Lmin = 1.7243 m.

Gazebo simulation with experiment configuration was created (Fig. 7). Validation of our implementation was made by checking trajectories followed when the robot passes through key points to a goal point. Avrora Unior robot model successfully went through all points although its trajectories were not perfect (Fig. 3). Errors were related to an imperfect stopping process. Average displacement with regard to OX axle was about 0.06 m and with regard to OY axle was about 0.013 m relatively to the goal point; the displacements were estimated in series of experiments.

Fig. 7. Parking space in Gazebo

7. Conclusions and future work

We have implemented an algorithm of parallel parking that is based on a geometrical approach. Current implementation of the algorithm has an obvious disadvantage, which is its dependency on the initial robot orientation with regard to a goal position.

We currently keep working on a solution for this problem by using dynamic coordinate frame. Moreover, PID controller has a direct influence on accuracy of turns in key points. After addressing this issue the algorithm will be validated on a real Avrora Unior robot. Our future work is focused on development of fully autonomous parking system of level 4 according to SAE J3016 standard.

Acknowledgements

This work was supported by the Russian Foundation for Basic Research (RFBR), project ID 19-58-70002.

References

1. Jung, H. G., et al. Parking slot markings recognition for automatic parking assist system. *IEEE Intelligent Vehicles Symposium* (2006), pp. 106-113.
2. Lavrenov, R. Smart Spline-Based Robot Navigation on Several Homotopies: Guaranteed Avoidance of Potential Function Local Minima. *International Conference on Artificial Life and Robotics* (2018), pp.407-410.
3. Shabalina, K., Sagitov, A., Magid, E. Comparative Analysis of Mobile Robot Wheels Design. In 11th *International Conference on Developments in eSystems Engineering* (IEEE, 2018), pp. 175-179.
4. Shabalina, K., Sagitov, A., Su, K.L., Hsia, K.H., Magid, E. Avrora Unior car-like robot in Gazebo environment, International Conference on Artificial Life and Robotics (2019), pp. 116–119,.
5. Mitchell W. C., et al. Analysis of Ackermann steering geometry. *SAE Technical Paper*, №. 2006-01-3638.
6. Lavrenov, R., Zakiev, A. Tool for 3D Gazebo map construction from arbitrary images and laser scans. In 2017 10th *International Conference on Developments in eSystems Engineering* (IEEE, 2017), pp. 256-261.
7. Latombe J.-C., *Robot motion planning.* **124** (Springer Science & Business Media, 2012).
8. Sungwoo, C., Boussard, C., & d'Andrea Novel, B., Easy path planning and robust control for automatic parallel parking, *IFAC Proceedings Volumes*, **44(1)** (2011), pp. 656–661.
9. "Rohith-k/autonomous-parallel-parking-car-like-robot-gazebo-ros," https://github.com/Rohith-K/Autonomous-Parallel-Parking-Car-likeRobot-Gazebo-ROS, accessed: 2019-04-11.
10. Saveliev, A., Malov, D., Edemskii, A., & Pavliuk, N. Proactive Localization System Concept for Users of Cyber-Physical Space. In *International Conference on Interactive Collaborative Robotics* (Springer, Cham, 2018), pp. 233-238.
11. Simakov, N., Lavrenov, R., Zakiev, A., Safin, R., & Martinez-Garcia, E.A., Modeling USAR maps for the collection of information on the state of the environment. *International Conference on Developments in eSystems Engineering* (IEEE, 2019).

Traffic Sign Recognition Algorithm for Car-like Robot Avrora Unior

Nikita Nikiforov[1], Ksenia Shabalina[1], Artur Sagitov[1], Kuo-Hsien Hsia[2], Evgeni Magid[1]
[1]Laboratory of Intelligent Robotic Systems (LIRS), Intelligent Robotics Department, Higher Institute of Information Technologies and Intelligent Systems, Kazan Federal University, Russia
[2]Department of Electrical Engineering, National Yunlin University of Science and Technology, Taiwan
E-mail: NANikiforov1337@gmail.com, ks.shabalina@it.kfu.ru, sagitov@it.kfu.ru, khhsia@yuntech.edu.tw, magid@it.kfu.ru
http://kpfu.ru/robolab.html

Abstract

Achieving high accuracy of traffic signs detection and recognition is difficult in real-time and is heavily influenced by non-ideal environment conditions. In this paper, we propose to combine a set of Haar cascades that had been trained on a large number of samples and could recognize different types of road signs in different positions and orientations. We use feature detection and feature matching in the process of traffic sign type identification. Our algorithm was validated on Avrora Unior robot model in a simulated environment within Gazebo.

Keywords: traffic sign recognition algorithm, Avrora Unior, car-like robot, Gazebo, simulation.

1. Introduction

Traffic sign recognition is an important task for unmanned cars and car-like robots[1]. Since road traffic is almost entirely formed by human drivers on a road, it is required for mobile unmanned ground vehicles (UGV) to strictly follow state traffic regulations and to guarantee road safety for other road users. Due to this, many computer vision tasks have become classic tasks of intelligent road agents development: traffic signs[2], traffic lights[3], and road markings recognition[4], determining speed and direction of movement of other road users[5].

In this paper we focus on traffic signs recognition task for car-like UGV Avrora Unior[6]. This is necessary in order to enable autonomous path planning[7] and locomotion possibilities of the robot within public roads[8] as well as applying parking[9] and overtaking[10] algorithms.

The rest of the paper is organized as follows: Section 2 describes the task and cascade training. Section 3 explains operation of the algorithm and its components. In Section 4 we discuss experimental results. Finally, we conclude in Section 5.

2. Traffic Sign Recognition Problem

To tackle the problem of traffic sign recognition, we have constructed shape detection cascading classifier. As cascading classifiers need to be trained with several hundred positive detection examples and several negative examples (for every particular sign type) we needed a decent training set. We have selected GTRSB[11] database as it contains convenient image format, images have small sizes and additional support information of objects' coordinates. Moreover, most of the dataset's sign design is similar to the design used in a Russian Federation, which allows direct transfer of the learned behavior from our Avrora Unior robot's control system to the system of a real full-size autonomous car. We use OpenCV library that has built-in routines for cascade training. Training process consists of two stages – creating a training dataset and subsequent model training.

2.1. *Creating a Training Set*

A cascade can find desired objects faster and more accurately if these objects have same proportions and shapes. Therefore, all traffic signs were divided into two groups: signs of a triangular shape and signs of a round shape. We prepared two datasets with information about the signs for each group. Figures 1 and 2 show examples of each traffic signs group that were used in a training.

Fig. 1. Examples of round shape signs

2.2. *Cascade Training*

After a training set was constructed we are constructed a cascade of classifiers. Two training functions were considered: HAAR[12] (in honor of Alfred Haar) and LBP[13] (Local Binary Patterns). HAAR has shown to be in some cases more accurate (showing accuracy advantage of 10-20%), but in some cases it took several days to complete its training. LBP on the other hand requires significantly less training time – the training procedure took just several minutes. Yet, in a case of road signs, our tests did not detect any significant difference in accuracy between HAAR and LBP functions. It was decided to use LBP, since in case of the database changes, it would be possible to quickly re-train the cascade.

Kuranov[14] have shown that a Haar feature cascade with 20x20 sample size achieved the highest detection rate, while cascade of four 18x18 nodes was computationally more efficient with slightly inferior result. Figure 3 shows cascade detection results that was used with training images.

Fig. 2. Examples of triangle shape road signs

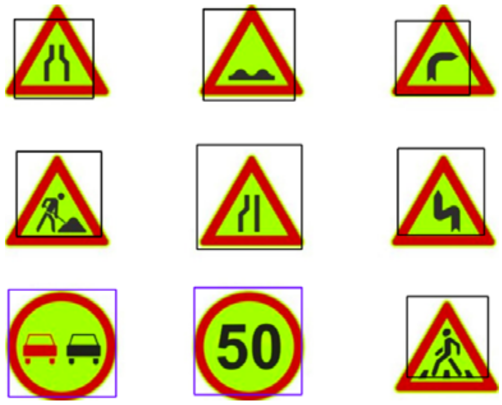


Fig. 3. Object detection by two cascades. The black contour is a cascade for triangular signs, the blue contour is a cascade for circular signs.

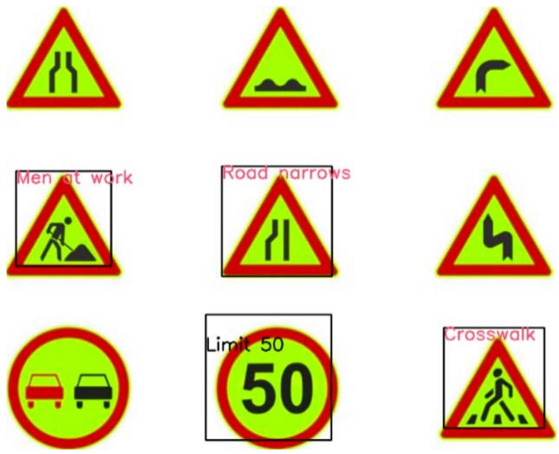


Fig. 4. Recognition of signs on a training image.

3. Algorithm Implementation

Avrora Unior is equipped with Microsoft Kinect, and since this camera's range is rather short, approximate dimensions of objects that the cascade could search for limits possible sign sizes. Our algorithm works in two stages. At the first stage ("LEARNING"), the algorithm uses one of available feature detector (ORB[15], SURF[16], SIFT[17]) to search for key features on training samples, and then stores results in a sign database. The second stage ("RECOGNITION") starts with a search of potential signs from a camera using our feature cascade, with results being passed to the classifier to determine an individual sign type. The comparison is performed using key feature comparators BFMatcher [18] and FLANN[19]. If a detected sign in a database has a relevant textual information it its displayed as an overlay on the frame (Fig. 4). Detected signs are saved into a separate database for further processing. Each of the cascades features an individual database, while each object from those databases is compared against samples for the training database.

Fig. 5. Sign models added into Gazebo simulator.

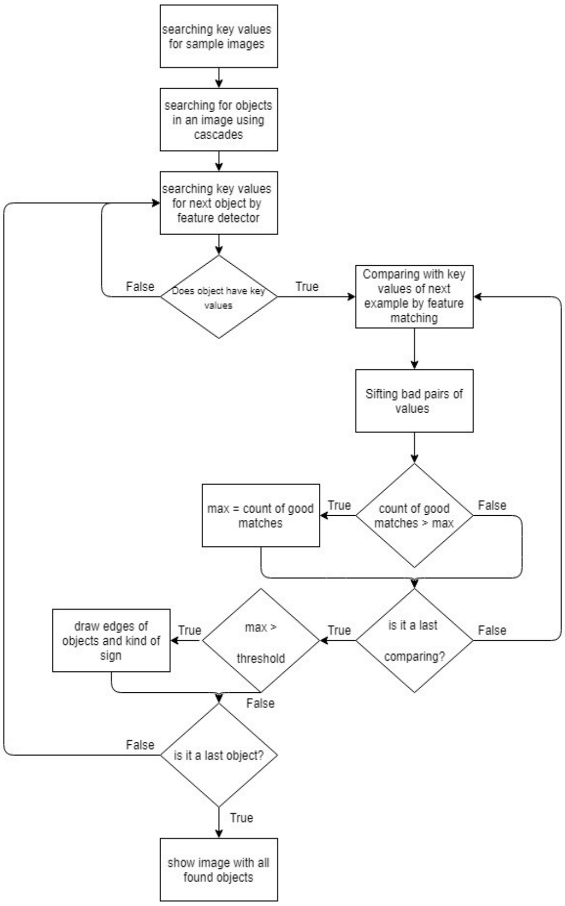


Fig. 6. Flowchart of a traffic sign recognition algorithm.

4. Algorithm comparison in Gazebo simulation

We performed validations in a simulated environment. For Gazebo simulator we created several models of signs of different type and stored them in SDF format. A set of traffic sign models was organized as a Gazebo model database[20] and a world was populated with the traffic signs (Fig. 5). that is used for Avrora Unior robot modeling to evaluate other parts of the autonomous control[12]. Our experimental design involved the following scenario:

- Avrora Unior robot drives in a straight line;
- Kinect camera captures frames to be used in traffic sign detection;
- ROS logging subsystem saves traffic sign recognition algorithm output of image processing, i.e., a detection status.

Table 1: Detection in different environments

Trial No.	Recognition time		
SIFT + BFMatcher			
	A	S	N
1	0.538	0.205	0.256
2	0.538	0.129	0.333
3	0.462	0.23	0.308
Average	**0.513**	**0.282**	**0.299**
ORB + BFMatcher			
	A	S	N
1	0.307	0.18	0.513
2	0.103	0.359	0.538
3	0.18	0.41	0.41
Average	**0.197**	**0.316**	**0.487**
SURF + FLANN			
	A	S	N
1	0.719	0.18	0.103
2	0.59	0.256	0.155
3	0.667	0.23	0.103
Average	**0.659**	**0.222**	**0.12**

The algorithm is presented in Fig. 6. A combinations of ORB with BFMatcher detectors and SIFT with BFMatcher, were tested in virtual experiments in Gazebo. Their accuracy in sign type recognition was lower than a combination of SURF with FLANN. The results are shown in Table 1: "A" means that it took less than 1 second to recognize a sign; "S" - it took more than 1 second to recognize a sign; "N" - a sign was not recognized. The first experiment took place in a regular environment illumination; the second experiment featured twilight-like environment with insufficient illumination; the third has featured excessive illumination. Road sign recognition algorithm failed to correctly recognize 4 out of 39 signs among the ones placed in the simulation world, i.e. had a 10.3% miss rate.

Algorithm	Detection rate
SIFT + BFMatcher	0.513
ORB + BFMatcher	0.197
SURF + FLANN	**0.659**

5. Conclusions and Future Work

In this paper we proposed a traffic sign recognition solution that uses several classifier cascades, one cascade for a triangular shape and one for a round shape. At the current stage the algorithm works poorly with narrow signs. Best accuracy in sign type recognition was demonstrated by a combination of SURF with FLANN.

In the future, we plan to add narrow sample images to training dataset. As a part of future work, we also plan to perform the same experiments in a laboratory environment with real car-like robot Avrora Unior.

Acknowledgements

This work was supported by the Russian Foundation for Basic Research (RFBR), project ID 18-48-160037.

References

1. Shabalina, K., Sagitov, A., Magid, E., Comparative Analysis of Mobile Robot Wheels Design. *International Conference on Developments in eSystems Engineering* (IEEE, 2018), pp. 175-179.
2. Greenhalgh, J., & Mirmehdi, M., Real-time detection and recognition of road traffic signs. *IEEE Trans. on Intelligent Transportation Systems* **13(4)** (2012), 1498-1506.
3. De Charette, R., & Nashashibi, F., Real time visual traffic lights recognition based on spot light detection and adaptive traffic lights templates. *Intelligent Vehicles Symposium* (IEEE, 2009), pp. 358-363.
4. Aufrere, R., Chapuis, R., & Chausse, F., A model-driven approach for real-time road recognition. *Machine Vision and Applications* **13(2)** (2001), 95-107.
5. Horn, B. K., & Schunck, B. G., Determining optical flow. *Artificial intelligence* **17(1-3)** (1981), pp. 185-203.
6. Shabalina K., Sagitov A.,Su K.L., Hsia K.H., Magid E., Avrora Unior Car-like Robot in Gazebo Environment. *International Conference on Artificial Life and Robotics* (2019), pp. 116-119
7. Magid, E., Lavrenov, R., Khasianov, A. Modified spline-based path planning for autonomous ground vehicle. *International Conference on Informatics in Control, Automation and Robotics* **2** (2017), pp. 132-141.
8. Klomp, M., et al., Trends in vehicle motion control for automated driving on public roads. *Vehicle System Dynamics* **57(7)** (2009), pp. 1028-1061.
9. Li, B., et al., Collaborative mapping and autonomous parking for multi-story parking garage. *IEEE Transactions on Intelligent Transportation Systems* **19(5)** (2018), pp. 1629-1639.
10. Dixit, S., et al., Trajectory planning and tracking for autonomous overtaking: State-of-the-art and future prospects. *Annual Reviews in Control* **45** (2018), pp.76-86.
11. Stallkamp, J., Schlipsing, M., Salmen, J., & Igel, C., The German Traffic Sign Recognition Benchmark: A multi-class classification competition. IJCNN **6** (2011), p. 7.
12. Soo, S., Object detection using Haar-cascade Classifier. Institute of Computer Science, University of Tartu, (2014) pp. 1-12.
13. Mu, Y., Yan, S., Liu, Y., Huang, T., & Zhou, B., Discriminative local binary patterns for human detection in personal album. *IEEE Conference on Computer Vision and Pattern Recognition* (2008), pp. 1-8.
14. Kuranov, A., An empirical analysis of boosting algorithms for rapid objects with an extended set of haar-like features. Intel Tech. Rep. (2002)
15. Rublee, E., Rabaud, V., Konolige, K., & Bradski, G. R., ORB: An efficient alternative to SIFT or SURF. *ICCV* **11(1)** (2011), p. 2.
16. Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L., Speeded-up robust features (SURF). *Computer vision and image understanding*, **110(3)** (2008), 346-359.
17. Vedaldi, A. An open implementation of the SIFT detector and descriptor. UCLA CSD (2007).
18. Jakubovic, A. & Velagic, J., Image Feature Matching and Object Detection Using Brute-Force Matchers. *Int. Symposium ELMAR* (IEEE, 2018), pp. 83-86.
19. Goel, A., Saxena, S. & Bhanot, S., Modified functional link artificial neural network. *International Journal of Electrical and Computer Engineering* **1** (2006), pp. 22-30.
20. Lavrenov, R., Zakiev, A. Tool for 3D Gazebo map construction from arbitrary images and laser scans. In *10th International Conference on Developments in eSystems Engineering* (IEEE, 2017), pp. 256-261.

Robotic Palpation Modeling for KUKA LBR IIWA Using Gazebo Simulator

Artur Shafikov[1], Artur Sagitov[1], Hongbing Li[2], Natalia Schiefermeier-Mach[3], Evgeni Magid[1]
[1]Laboratory of Intelligent Robotic Systems (LIRS), Intelligent Robotics Department, Higher Institute for Information Technology and Intelligent Systems, Kazan Federal University, 35 Kremlyovskaya street, Kazan, Russia Federation
[2]Department of Instrument Science and Engineering, Shanghai Jiao Tong University, Shanghai, China
[3]Health University of Applied Sciences Tyrol, Innsbruck, Austria
E-mail: artur.shafikov96@gmail.com, sagitov@it.kfu.ru, lihongbing@sjtu.edu.cn, natalia.schiefermeier-mach@fhg-tirol.ac.at, magid@it.kfu.ru
http://kpfu.ru/robolab.html

Abstract

Palpation is one of diagnostic methods being extensively used in medical practice. It is often used for tumor detection in cancer screening but its efficacy is highly dependent on examining physician's skill. Therefore, using a robotic tool could make this procedure more objective. In this paper, we present our control and perception modules of autonomous palpation robotic system. We have modeled KUKA LBR IIWA manipulator control using MoveIt motion planning in Robot Operating System and validated it in Gazebo simulator.

Keywords: robotic palpation, autonomous palpation simulation, surface traversing, medical robotics.

1. Introduction

Surgical palpation is an examination of a patient's body technique that is performed manually by a practitioner with his/her hands, which helps to decide on a diagnosis by analyzing obtained tactile information[1]. Palpation is commonly used in diagnosing of atherosclerosis, hypersplenism, thrombophlebitis and many other diseases and illnesses. One of the most significant areas where palpation is being used is an early cancer diagnosis that requires deciding on whether there is a tumor in examined body part according to tissue's relative stiffness; and, if possible, identifying its type[1].

The main reason why automation of a palpation procedure is an important and necessary task is its subjectivity, since it is strongly dependent on practitioner's capability of sensing slightest changes in stiffness that is followed by interpretations of these tactile sensations. Those skills could not be easily taught as tactile perception is very individual. Main goals of our research is to allow compensation of the aforementioned subjectivity by quantitative estimation of palpated tissues' and detected tumors' properties, which could increase a percentage of correct abnormalities detection and enhance a spatial measurement's accuracy.

In this paper we present a simulation of a palpation procedure being performed with a help of an autonomous robotic manipulator. Simulation makes possible to analyze the robot's behavior within the palpation procedure before performing real world experiments and help identifying weaknesses of the developed algorithm, its software implementation and the robot hardware[2]. Simulation is performed in Gazebo simulator[3] that provides necessary modules to verify algorithms, estimate sensors' performance and control multiple physical properties of different simulated objects[4,5].

2. Related Work

Though robotic palpation researches lack works that are directly related to simulation, still there is a variety of papers that are focused on automated and robotized methods of surface exploration and tissue properties acquisition in medical purposes.

There exist different ways of robot end-effector interaction with a surface being examined. Garg et al.[6] have applied indentation technique, which is one of the most simple and frequently used methods in robotic palpation research. Guo et al.[7] have introduced indentation with a static measurement, i.e. when a robot keeps an end-effector position at a certain indentation depth until sensor's data becomes stable, thus leading to uncertainties reduction. On the other hand, Liu et al.[8] implemented continuous rolling across an examined surface that was compared with the indentation technique results and demonstrated faster surface exploration with a higher density of extracted surface information. Goldman et al.[9] used continuous cycloid motions to compensate a bias that was influenced by a direction of movements across an examined surface. Chalasani et al.[10] proposed using a continuous motion across an examined surface with sinusoidal depth profiles, referencing technique being used in a traditional surgery. Konstantinova et. al.[11] introduced an indentation method that applied a technique of coupled normal and lateral forces modulation that was based on mathematical model that had been derived from manual palpation experiments.

Surface exploration strategy is a complementary problem being addressed by a robotic palpation research. Goldman et al.[9] applied a kernel search algorithm that generates a projection on a surface grid with an adaptive resolution. In Refs. 12 and 13 authors used a force feedback from force/torque sensors in order to compute normal and tangential forces to generate desirable end-effector position and orientation in a way that follows the normal of a palpated surface. Hess et al.[14] used representation of a surface exploration task as a traveling salesman problem and have used Kinect camera for extraction of a surface information in a graph generation. In Refs. 6 and 10 the authors applied machine learning techniques to help in a generation of trajectories for a surface exploration.

Some of the aforementioned researches have elements that use a finite element simulation to estimate a surface behavior; however, once an algorithm is being integrated into a robotic system, the simulation becomes computationally expensive.

This paper aims to provide a simulation setup for robotic palpation to be performed by KUKA LBR IIWA manipulator. Similar model has proven to be an appropriate choice for a surface exploration by Virga et. al.[15] who have used it for an autonomous ultrasound screening robotic system.

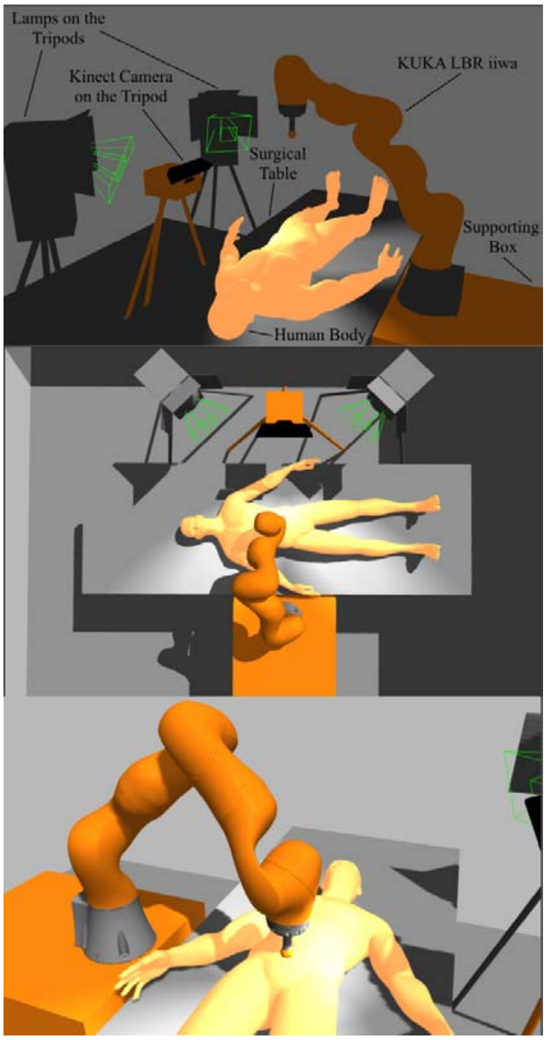


Fig. 1. Simulation scene.

3. Setup

Simulation setup is depicted in Fig. 1 and consists of a robotic manipulator that performs palpation task, clinical setting that includes auxiliary systems and an object of palpation procedure, which is a virtual patient.

3.1. *Simulation setup*

To implement a palpation procedure in a simulation we used KUKA LBR IIWA manipulator model[16]. Packages provide the model that fully supports Gazebo simulator with a full MoveIt! integration, which makes it possible to implement the control for a robot via Robot Operating System (ROS).

KUKA LBR IIWA robot is a 7 degrees of freedom lightweight redundant manipulator that was developed for a human-robot collaboration (HRC). Manipulator's redundancy as it have been pointed out by a Torabi et. al.[17] results in decreased inertia and friction during palpation procedures. HRC-compatibility is also an important feature of KUKA LBR IIWA manipulator as such robotic systems are developed to assist physician in the first place. The model was extended by adding a spherical indenter as an end-effector for a palpation, which is connected with a force sensor in the seventh link of the original manipulator. Similar types of end-effectors were used in Refs 6, 7, 9 and 11.

3.2. *Clinical setting*

Models for clinical setting were either developed specifically for this task (the surgical table, the tripod, the lamp models), or taken from default Gazebo simulator repositories (Kinect camera and the box).

Clinical setting consists of a surgical table on which a palpated object is to be placed, a rectangular box on which a robotic manipulator is mounted in a such way that its workspace covers most of the surgical table area, Kinect camera with ROS depth camera plugin[18, 19] to estimate surface geometry before palpation procedure starts and two lamps for better illumination. In order to be directed towards an palpated object, lamps and Kinect camera are mounted on tripods. Kinect camera and lamps are directed towards an abdominal area of a human body model - this area will be examined during the autonomous palpation process.

3.3. *Virtual patient*

A 3D model of virtual patient was obtained from Ref. 20 as an OBJ 3D model and then converted into STL format in order to be applied in Gazebo simulator. However at the current stage of development it serves only as a visual part of a virtual patient - in terms of collision virtual patient is a rectangular box, which surface is palpated.

4. Autonomous Palpation

There are two aspects of autonomous palpation: control, i.e., making a manipulator's end-effector traverse the target surface, and perception, i.e., getting data from manipulator's sensors. These aspects are further described in this section.

4.1. *Control*

Sliding movement across the palpated surface in a snake scan manner was implemented for examination of a surface with predefined position and dimensions. Prior to surface examination the manipulator moves slowly in a free space towards the palpation surface until the end-effector detects a collision with the surface.

4.2. *Perception*

Once manipulator's end-effector is in contact with the surface, z-component of force measurements from 3-axis force/torque sensor, which is embedded into 7-th robot's link, increases tenfold. This allows detecting a contact during movement towards the surface. To ensure consistent sensing, the end-effector is kept at the orientation that is orthogonal to the examined surface while running the palpation procedure.

During the palpation procedure data from the sensor is being encoded into a marker message and is published as a topic. From there these data may be fetched and visualized via RViz, which is a 3D visualizer for ROS framework. Thus it is possible to obtain a color-coded map of normal force values for each examined point of the palpated surface. This feature provides the ability to detect simulated abnormalities of an examined tissues.

5. Discussion and Future Work

This paper presented our first version of an open source software that contains KUKA LBR IIWA extended model and a simulation scene. Our software allows testing robotic palpation algorithms. Additionally one of the basic palpation methods, which consists of a surface traversing and data acquisition, was implemented and successfully tested in the proposed simulated environment.

Our future work is aimed at making simulated palpation procedure more general-purposed. That includes implementation of a surface geometry information acquisition by Kinect camera and subsequent support of a curved surface palpation without predefining the position and dimensions of a target object. Another goal is set at expanding the available palpation methods by implementing new ones. Finally, we plan to introduce a tissue model that possesses elastic properties so that a force sensor could react differently to surfaces with various elasticity characteristics.

Acknowledgements

This work was supported by the Russian Foundation for Basic Research (RFBR), project ID 18-48-160037.

References

1. J. L. Longe and D. S. Blanchfield, *Gale encyclopedia of medicine*, 2nd edn. (Detroit, MI: Gale Group, Detroit, MI, 2002).
2. Sagitov, A., Li, H., Magid, E. Toward automated open wound suturing using haptic feedback: detecting wounds and planning the suture. In *Int. Conf. on Artificial Life and Robotics* (2018), pp. 208-211.
3. Sagitov, A., Li, H., Schiefermeier-Mach, N., Magid, E. Extending Gazebo simulator for surgical robotics: tissue and suture modeling. *International Conference on Artificial Life and Robotics* (2019), pp. 124-127.
4. Lavrenov, R., Zakiev, A. Tool for 3D Gazebo map construction from arbitrary images and laser scans. *International Conference on Developments in eSystems Engineering* (IEEE, 2017), pp. 256-261.
5. Sagitov, A., Tsoy, T., Li, H., Magid, E. Automated open wound suturing: detection and planning algorithm. *Journal of Robotics, Networking and Artificial Life*, **5(2)** (2018), pp.144-148.
6. A. Garg, S. Sen R. Kapadia, Y. Jen, S. McKinley, L. Miller and K. Goldberg, Tumor localization using automated palpation with Gaussian process adaptive sampling, *IEEE International Conference on Automation Science and Engineering* (2016), pp. 194–200.
7. J. Guo, B. Xiao and H. Ren, Compensating Uncertainties in Force Sensing for Robotic-Assisted Palpation, *Applied Sciences*, **9(12)**: 2573 (2019).
8. H. Liu, D. P. Noonan, B. J. Challacombe, P. Dasgupta, L. D. Seneviratne and K. Althoefer, Rolling mechanical imaging for tissue abnormality localization during minimally invasive surgery, *IEEE Transactions on Biomedical Engineering*, **57(2)** (2010), pp. 404–414.
9. R. E. Goldman, A. Bajo and N. Simaan, Algorithms for autonomous exploration and estimation in compliant environments, *Robotica*, **31(01)** (2013), pp. 71–87.
10. P. Chalasani, L. Wang, R. Roy, N. Simaan, R. H. Taylor and M. Kobilarov, Concurrent nonparametric estimation of organ geometry and tissue stiffness using continuous adaptive palpation, *IEEE International Conference on Robotics and Automation* (2016), pp. 4164–4171.
11. J. Konstantinova, G. Cotugno, P. Dasgupta, K. Althoefer and T. Nanayakkara, Correction: Palpation force modulation strategies to identify hard regions in soft tissue organs, *PloS One*, **13(1)** (2018), p.e0192259
12. E. C. Li and Z. M. Li, Surface Tracking with Robot Force Control in Unknown Environment, *Advanced Materials Research*, **328-330** (2011), pp. 2140-2143.
13. J. Back, J. Bimbo, Y. Noh, L. Seneviratne, K. Althoefer and H. Liu, Control a Contact Sensing Finger for Surface Haptic Exploration, *IEEE International Conference on Robotics & Automation* (2014), pp. 2736–2741.
14. J. Hess, G. D. Tipaldi and W. Burgard, Null Space Optimization for Effective Coverage of 3D Surfaces using Redundant Manipulators, *IEEE/RSJ International Conference on Intelligent Robots and Systems* (2012), pp. 1923-1928.
15. S. Virga, O. Zettinig, M. Esposito, K. Pfister, B. Feisch, T. Neff, N. Navab and C. Hennersperger, Automatic force-compliant robotic ultrasound screening of abdominal aortic aneurysms, *IEEE/RSJ International Conference on Intelligent Robots and Systems* (2016), pp. 508-513.
16. C. Hennersperger, B. Fuerst, S. Virga, O. Zettinig, B. Frisch, T. Neff and N. Navab, Towards MRI-based autonomous robotic US acquisitions: a first feasibility study, *IEEE transactions on medical imaging*, **36(2)** (2017), pp. 538-548.
17. A. Torabi, M. Khadem, K. Zareinia, G. R. Sutherland and M. Tavakoli, Application of a Redundant Haptic Interface in Enhancing Soft-Tissue Stiffness Discrimination, *IEEE Robotics and Automation Letters* **4(2)** (2019), pp. 1037–1044.
18. Safin, R., Lavrenov, R. Implementation of ROS package for simultaneous video streaming from several different cameras, *International Conference on Artificial Life and Robotics* (2018), pp. 220-223.
19. Safin, R., Lavrenov, R., Martínez-García, E. A., & Magid, E. ROS-based Multiple Cameras Video Streaming for a Teleoperation Interface of a Crawler Robot. *Journal of Robotics, Networking and Artificial Life* **5(3)** (2018), pp.184-189.
20. free3d.com, (2015). *Male Base Mesh 3D Model*. [online] Available at: free3d.com https://free3d.com/3d-model/male-base-mesh-6682.html [Accessed 13 Dec. 2019].

Remote Control Application for “Servosila Engineer” on Android Mobile Devices

Daniel Kiryanov, Roman Lavrenov
Laboratory of Intelligent Robotic Systems (LIRS), Intelligent Robotics Department,
Higher Institute for Information Technology and Intelligent Systems,
Kazan Federal University, 35 Kremlyovskaya street, Kazan, Russia Federation
E-mail: danielkiryanov@gmail.com, lavrenov@it.kfu.ru
http://kpfu.ru/robolab.html

Abstract

Even though modern mobile robots’ autonomous navigation capabilities rapidly increase, teleoperation mode is still an important tool, especially in critical domains like rescue or military robotics. This paper presents Android OS based teleoperator control tool for Russian crawler robot Servosila Engineer. We changed the way of data exchange between a robot and its operator, which allows using Wi-Fi data standards in order to simply data transfer from OCU (Operator Control Unit) process to vehicle process. Our application provides robot remote control and video data transfer from robot onboard cameras.

Keywords: crawler robot, OS Android, remote-control tool, UDP, GUI, user interface.

1. Introduction

Today mobile devices are equipped with hardware that is not inferior in performance to computers. For example, smartphones have a wide range of built-in sensors such as an accelerometer, gyroscope, GPS, etc. High performance and multi-threading of mobile processors provide high-speed programs. Increased RAM and physical memory allow processing a large amount of incoming data, while growing capacity of batteries increases devices’ autonomy. All together these allowed using mobile devices in many areas of modern information technologies, including robotics.

The common ground between mobile development and robotics can be, for example, augmented reality, robot interaction with a mobile device, robot control via a mobile application, including a remote-control tool via Wi-Fi. In the later case, it’s important to consider interaction between the systems.

Nowadays a large number of developments in this area exist. One of them is ABR (Android Based Robotics) project[1]. It has both client and server software. Android app used to connect to a TCP server (ABR_server) and allow streaming of data and remote control. Connection between Android client and the server is via UDP (User Datagram Protocol)[2] protocol. The project demonstrates processes of connecting and receiving data from a TCP server, sending data over UDP sockets, capturing frames from a phone’s camera for streaming etc. Server is a QT (C++) project that runs on a PC and can receive connections from multiple phones (ABR_client), display video feedback and sensory information and serve to control robots remotely. Another example is “ROS Control” tool[3]. It is a universal tool using ROS Android. “ROS Control” has multiple teleoperation control options including joystick and tilt control, waypoint planning and navigation, remote camera view, laser scan visualization, GPS data and map view.

In our research, we developed Android OS based remote-control tool for Servosila Engineer robot[4] (Fig.1) using Android SDK only. This was possible due to the software on a vehicle process, which allows communicating between the robot and a smartphone using standard data transmissions[5].

2. Servosila Engineer Robot

The mobile robot Engineer (Fig.1) is developed by Russian company "Servosila" for search and rescue operations in natural and man-made disasters[6]. It is used when receiving remotely a video data from inside a dangerous room is required[7]. Special solutions embedded in the design of the robot chassis allow the robot to enter objects that are located inside industrial and residential buildings. Servosila Engineer is equipped with a powerful manipulator arm, which allows the robot to remove samples of dangerous objects from a disaster area for a detailed analysis or immediately perform a dangerous engineering operation on-site to eliminate or prevent further growth of a man-made disaster. Table 1 briefly describes robot equipment.

Fig. 1. Servosila Engineer servo drive scheme.

3. Android OS

Android operating system is Linux-based and has Google's own Java virtual machine implementation. Applications for Android OS are programs in non-standard byte code for Dalvik virtual machine, for which format of installation packages APK was developed.

Table 1. Robot equipment.

Sensors	Types
Cameras	24x optical camera
	Rear view camera
	Two front cameras for stereo vision
Thermal imager	External
Auto Navigation	Laser scanner
	GPS/GLONASS receiver
	Stereo cameras

Many libraries are available for working with applications: Bionic (library of standard functions), OpenGL (3D graphics engine), SQLite (a lightweight DBMS available for all applications), SSL (protocol for secure data transmission over the network), etc.

4. Existing tools

Programmers can develop software for Android in Java using the SDK or in a native language (C/C++) using the native development kit (NDK). It is also possible for developers to modify the Linux kernel if needed. Implementation of an Android application can be achieved using Android Studio IDE or Eclipse IDE with Android Development Tools (ADT) plug-in. Using this SDK, a developer obtains an easy access to different functionalities of Android phone such as graphical interfaces, multi-threading, networking, data storage, multimedia, sensors, location provider, speech-to-text, text-to-speech, and more. Since Android phones can connect to the Internet, cloud-based applications can also be used when high-performance computing is needed. In robotics, this feature can allow cloud-based robotics applications development. When developing an application that is CPU-intensive but doesn't allocate much memory, an alternative programming option is to use Android NDK. With the NDK, a programmer can create an Android Java application that interacts with native code (C/C++) using Java Native Interface (JNI)[8].

Programming in C/C++ on an Android platform can result in an increase of performance, but also increases complexity. The NDK also enables usage of existing C/C++ libraries. This maked possible exporting popular libraries (e.g., computer vision library OpenCV) to Java so that they can be incorporated in Android applications (OpenCV Android). The robot operating system (ROS) is also available for Android in Java (ROS Java, ROS Android)[9]. An example of SDK is "Pepper SDK"[10], which provides an Android Studio plug-in to develop human-robot interaction tools for Pepper Humanoid Robot. There are many Android applications for robots on Arduino hardware platforms[11]. Android OS mobile devices can also act as a hardware platform due to series of open-source PIC microcontroller-based boards "IOIO"[12]. IOIO allows Android mobile applications to interact with external electronics. Its hardware and software are entirely open source and enabled the creation of hundreds of DIY robotic projects around the world.

Frame Type ID	Axis #0	Axis #1	...	Axis #15	Button #0	Button #1	...	Button #15	Video Bit Rate Telemetry
1byte	2bytes	2bytes	...	2bytes	1byte	1byte	...	1byte	8bytes
uint8	int16	int16	...	int16	uint8	uint8	...	uint8	double

Fig. 2. Remote control packet structure.

5. Remote control

The above technologies are the result of porting software to Android. In order to interact with the robot in another environment without using third-party libraries, it is necessary to make an interactor on an on-board control computer of the robot. Therefore, we developed a communication system between the vehicle process and OCU process using data transfer over UDP, which was selected because of ease of use. For example, the OCU process sends a request to the vehicle process and hopes to get a response. The client after a certain time may try again if the request or response is lost. It allows developing easier code and reducing a required number of messages in comparison with protocols that request initial configuration. In order for the parties to understand each other, it is also necessary to determine a format of a transmitted message. Figure 2 shows a remote-control packet structure that consists of the following parts[2]:

- frame type id – value indicating a type of message being sent (e. g. remote-control packet)
- axis – fields are used to indicate a speed of servo drives' rotation
- buttons – fields are used for giving an amperage to servo drives
- video bit rate telemetry – section using to set a number of bits used for video data transmission/processing

Due to this structure, the OCU process can control all Engineer's servo drives (Fig. 1) movements. Therefore, any environment can act as a remote-control tool. We used Android OS as one of them.

6. Mobile Application

We designed a simple and intuitive interface when creating this tool. Robot servo drives were divided into groups "Movement" and "Joints" and placed in the separate windows (Fig. 3). There are options to set their

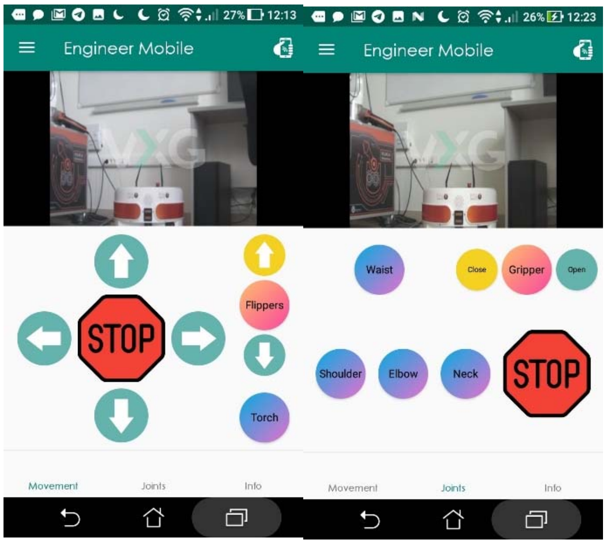


Fig. 3. "Movement" application window (left) and "Joints" application window (right).

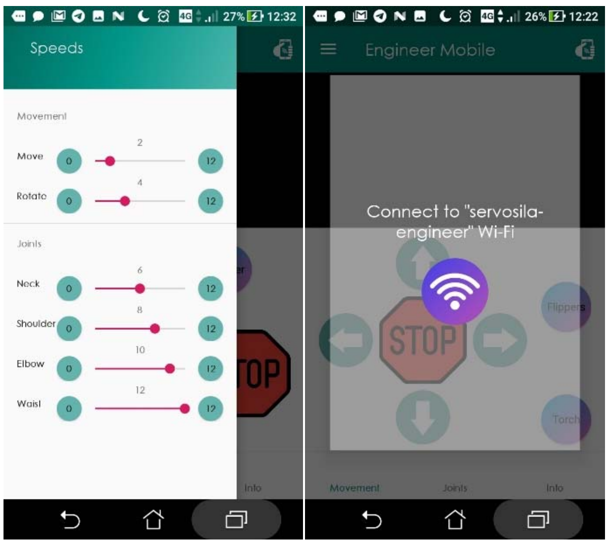


Fig. 4. "Speed" sidebar (left) and "Connection" dialog (right).

speed in the sidebar (Fig. 4). It's useful when it's important to change the speed while the robot is moving. Also, the stop button was provided (Fig. 3). We added video processing from the front camera via RSTP (Rapid Spanning Tree Protocol). The mechanism associated with transmitting the remote-control packet over the network is initialized in a separate process called "Service" (Fig. 5). It allows sending the remote-control packet to the server several times per second and does not affect the UI. The packet is formed inside the PackageManager and transmitted to the UDP client which uses the standard Java mechanism called "DatagramSocket". UI consists of MainActivity and Fragments that can be interchanged. Robot Servosila Engineer has DHCP server for connection, so the app can get its IP, form it in Utils and transfer it to the UDP Client.

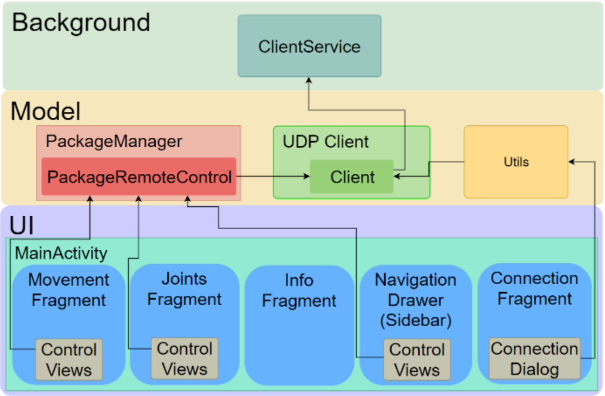


Fig. 5. Application data sharing scheme.

7. Conclusions and Future Work

In this paper, we presented Android OS based teleoperator control tool for Russian crawler robot Servosila Engineer. We developed software for the onboard control computer of the robot. It receives structured remote-control packets by UDP, so we designed a remote-control application using this standard. This tool allows controlling the robot's movement and servo drives. The application receives a video stream from the front camera via RSTP. As a part of our future work, we will supply the application with telemetry data and video streams from other cameras and use Google maps to represent the robot geolocation.

Acknowledgements

This work was supported by the Russian Foundation for Basic Research (RFBR), project ID 19-58-70002.

References

1. Nicolas Oros, Jeffrey L. Krichmar. Android™ Based Robotics: Powerful, Flexible and Inexpensive Robots for Hobbyists, Educators, Students and Researchers. [Web Link].URL:https://www.socsci.uci.edu/~jkrichma/ABR/index.html.
2. Mavrin, I., Lavrenov, R., Magid E. Development of a Graphical User Interface for a Crawler Mobile Robot Servosila Engineer, 11th International Conference on Developments in eSystems Engineering (IEEE, 2018), p. 192-197.
3. I. Rekleitis. Android app for remote controlling a ROS robot. [Web Link]. URL: https://mtbii.github.io/RobotCA
4. Sokolov, M., et al. Modelling a crawler-type UGV for urban search and rescue in Gazebo environment. In International Conference on Artificial Life and Robotics (2017), pp. 360-362.
5. Li, H., et al. Achieving Position Synchronization in Passive Bilateral Teleoperation, IEEE International Conference on Robotics and Biomimetics (IEEE, 2018), pp. 2208-2213.
6. Magid, E., Tsubouchi, T. Static Balance for Rescue Robot Navigation-Translation Motion Discretization Issue within Random Step Environment. In ICINCO v.2, 2010, pp. 415-422.
7. Safin, R., Lavrenov, R., Martínez-García, E. A., Magid, E. ROS-based Multiple Cameras Video Streaming for a Teleoperation Interface of a Crawler Robot. Journal of Robotics, Networking and Artificial Life, **5(3)** (2018), p.184-189.
8. S. Goebel, et al., Using the Android Platform to control Robots, in Proceedings of 2nd International Conference on Robotics in Education, 2011.
9. J. Kerr and K. Nickels, Robot operating systems: Bridging the gap between human and robot, in Southeastern Symposium on System Theory (IEEE, 2012), pp. 99–104.
10. N. Fung, Light weight portable operator control unit using an Android enabled mobile phone, in Proc. SPIE Unmanned Systems Technology XIII, Orlando, FL, 2011.
11. Y. Ji-Dong, et al., Development of Communication Model for Social Robots Based on Mobile Service, in IEEE Second International Conference on Social Computing (2010), pp. 57-64
12. S. Wu, Y. Chen, Remote robot control using intelligent hand-held devices, in Fourth Int. Conf. on Computer and Information Technology, Wuhan, China, 2004.

Network Failure Detection and Autonomous Return for PMB-2 Mobile Robot

Dmitry Bereznikov, Aufar Zakiev
Laboratory of Intelligent Robotic Systems (LIRS), Intelligent Robotics Department,
Higher Institute for Information Technology and Intelligent Systems,
Kazan Federal University, 35 Kremlyovskaya street, Kazan, Russia Federation
E-mail: bereznikovdmitry@gmail.com, zaufar@it.kfu.ru
http://kpfu.ru/robolab.html

Abstract

In real world teleoperated tasks a robot connection with its operator is not always stable, so it is important to increase the robot autonomy. This paper focuses on increasing robot autonomy through autonomous return and charging station docking in a case of connection loss. We integrated the algorithm into real robot control system or PAL Robotics PMB-2 robot and experimentally demonstrated its good efficiency. The algorithm analyzes network failure through incoming TCP/IP packets, uses Simultaneous Localization and Mapping (SLAM) and path planning algorithms for autonomous return, and dock station plugin for the robot docking and recharging, which continues until the connection to teleoperator station is restored.

Keywords: mobile robot, algorithm, autonomous return, network failure detection, dock station, PMB-2.

1. Introduction

Mobile robotics is a very practically-oriented field of robotics with robots being widely used in broad range of different tasks[1], including cargo delivery[2], environment mapping and hazardous zones exploration[3,4]. Mobile robots often have operator[5,6], which controls robot actions and monitors robot state remotely from a safe place. Communication between mobile robot and the operator mostly is settled by a wired or wireless connection. Thus, teleoperated robots could lose control because of imperfections in communications technology, human factor or force majeure events[7]. Lost control may lead to entire robot loss or its partial damage[8], which requires mobile robots to demonstrate some level of autonomy in order to avoid such situations.

In this paper, we developed and integrated in PAL Robotics mobile robot PMB-2 control system[9] our network failure detection and autonomous return to dock station algorithm. Our goal was to implement operator-independent autonomous return system that does not require any software changes on operator's side. Therefore, the detection is based on incoming TCP/IP packets analysis. After connection failure is detected, robot returns to its initial position; during the movement, it searches for its dock station in order to start charging until a connection link with the operator restores.

Existing solutions for network failure detection include different approaches to detect connectivity loss. Algorithm[10] requires software configuration on both robot and operator's PC. Papers[11,12] are focused on multiple robots to detect network failure between them. Other methods[13,14] imply special device usage to measure received signal power. The method, which we present in this paper, is developed to overcome these limitations and is experimentally validated.

This paper is organized as follows: Section II describes PMB-2 robot used for experimental validation. Section III is dedicated to network fail detection algorithm. Next, autonomous return is described and experimental results are presented in Section IV. Last Section contains conclusions.

Fig. 1. PMB-2 robot (left) and its charging in the dock station (right); the dock is the white box in the top of the picture.

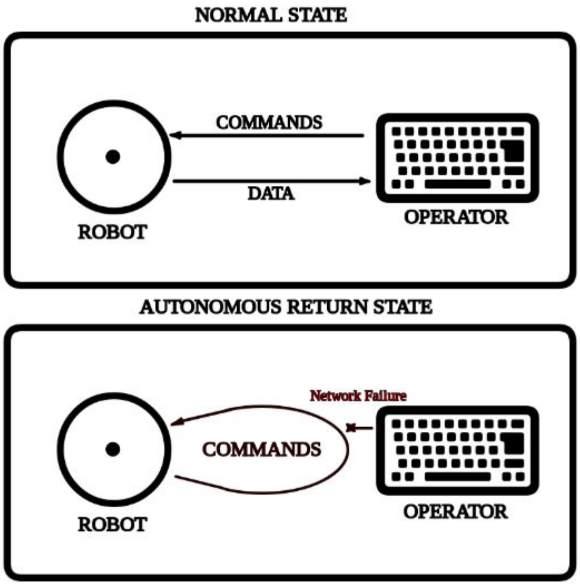


Fig. 2. Network failure detection diagram.

2. System Setup

PMB-2 (Fig. 1) is a PAL Robotics[15] company mobile robot that is equipped with Sick TiM 571 laser range finder (LRF) and 6 DoF inertial measurement units (IMU) sensors, which we use for SLAM and navigation. The robot has 54 cm diameter, 30 cm height, maximal speed of 1 m/s and a round shape, which makes it convenient for indoor use. Its built-in 20,000 mAh Li-Ion battery allows to work autonomously without recharging for a long time. The robot is equipped with 802.11b/g/n/ac Wi-Fi interface and could act as a wireless access point. PAL Robotics also provides charging dock station (Fig. 1, right) with the robot.

PMB-2 runs on Ubuntu 16.04 operating system and ROS Kinetic. Software has preconfigured SLAM and navigation packages. Mapping is performed using *slam_gmapping*[16] ROS package; AMCL[17] is responsible for localization and actual navigation is implemented in *global_planner*[18] package. Operator PC has Ubuntu 16.04 and ROS Kinetic environment installed. However, our system is software-agnostic, thus, it only requires to communicate with robot through TCP/IP connection.

3. Network Failure Detection and Recovery

PMB-2 is controlled through ROS interface and teleoperation mode is also implemented using ROS capabilities. The workflow is following:

(i) PMB-2 turns on internal Wi-Fi interface and creates access point.
(ii) Operator connects to the PMB-2's access point.
(iii) Operator configures ROS on his PC to connect to PMB-2's ROS.

The robot and operator's PC are united into a common network with a full software compatibility, i.e. ROS nodes launched on robot are visible and interactive for ROS nodes launched on operator's PC and vice versa. Such tough integration is very convenient for common software development workflow since ROS nodes are indifferent to source of data or commands. They could come through Wi-Fi, Ethernet, Bluetooth or even Zigbee networks[19] encapsulated in ROS topics. However, this encapsulation hides the data source information, therefore, there is no common way to determine whether topic is disconnected (data could not reach the destination point) or topic is simply empty (data is not sent at all). Therefore, we developed network failure detection node that takes teleoperation commands topic as an input and determines if the connection to operator is alive or not. Connection state is determined by *ping* utility[20], which is built-in into modern OSs as well as into Ubuntu.

At the system startup the node waits for initial teleoperator connection to prevent the robot from returning to a starting point instantly after turning on (Fig. 2). In normal state, ping is not used and no network load by failure detection algorithm is applied. However, if there are no incoming commands for 10 seconds, network failure possibility is recognized by the robot. It automatically turns on the ping utility and checks whether an operator is reachable in the network. Returning ping messages mean that the operator is in the network and does not send commands; if no ping returns, the operator is considered as disconnected and the autonomous return algorithm is launched.

After switching to autonomous return mode, the robot launches a built-in function of an automatic connection to a dock station. This feature allows the robot to safely wait for the operator reconnection if the starting position is successfully reached. Docking feature detects dock station surface laser pattern at up to 1-meter distance and performs a docking maneuver. Successfully docked robot starts to recharge (Fig. 1, right).

During the return movement, the robot continuously checks for operator's availability in the network. If the connection restores, it switches back to teleoperation mode. The autonomous navigation and movement stop; also, if the robot is already at the dock station, it undocks and waits for commands in fully operational state. Thus, no operator assistance is needed for the robot to recover from a network failure.

4. Autonomous Movement

Autonomous movement is fully implemented in ROS framework. When the algorithm detects network failure between robot control system and teleoperator, algorithm sends command to *move_base*[21] node which is part of the implemented at this robot navigation stack. The goal command is filled with position of starting point, orientation of robot and frame ID which is used for global navigation. After sending the goal, *global_planner* node builds a plan on the map from robot current position to goal position. This global plan goes as an input to *move_base*; this node turns the planned path into movement commands to robot's motor drives (Fig. 3). Algorithms performs searching for dock station during the movement to save time for docking.

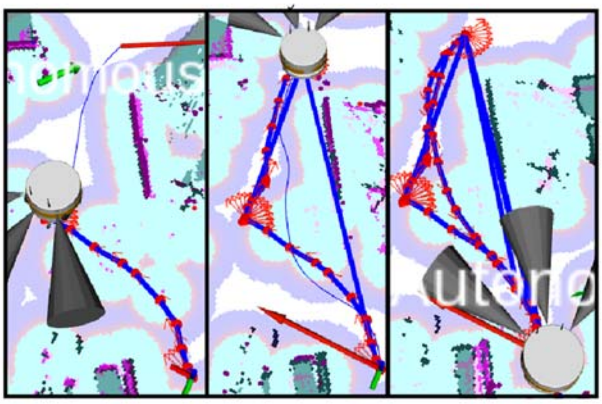


Fig. 3. Network failure detection and autonomous return. Robot is teleoperated (left); network failure detected (center); autonomous return performed (right).

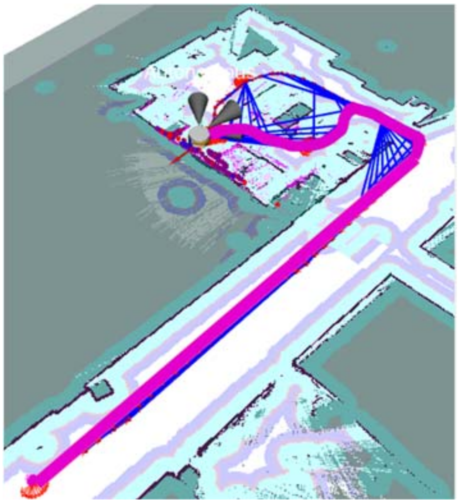


Fig. 4. Travel distance pointed purple line - 24 meters.

Autonomous return process relies on a map that was obtained during teleoperated movement and could be further used for localization and navigation[22]. PMB-2 uses laser-based SLAM *slam_gmapping* node which allows us to construct an occupancy grid map. During autonomous return, the robot is able to localize itself on this map using *amcl* package which uses Adaptive Monte Carlo Localization method.

5. Experiments

We conducted a set of seven experiments to determine the effectiveness and applicability of our algorithm. In each trial, we measured (Table 1):

- the distance traveled by the robot during the autonomous return (second column);
- time intervals from the moment of connection break, to the moment of switching robot to the autonomous return mode (third column);
- return time to the starting point (fourth column).

Experimental results in Table 1 show that delay of network failure detection is not stable and fluctuates around 12-16 seconds. However, it is not dangerous, because if there are no commands from a teleoperator, the robot stops. Figures 3 and 4 illustrate that the robot autonomous return time is more dependent on a path complexity than on a path length. The shorter path had multiple obstacles on the robot's way, which slowed it down. The longer path of the seventh trial consumed relatively little time because of a higher speed on straight segments of the path.

All tests passed successfully; the algorithm detected communication break with the teleoperator, switched to autonomous return and the robot reached starting point, found the dock station and connected for recharge.

Table 1. Algorithm testing results.

Trial number	Traveled distance in meters	Time to detect network failure after disconnect in seconds	Traveled time in seconds	Figures to illustrate
1	4	13	10	Fig. 3
2	4	16	12	
3.	4	14	18	
4	4	13	20	
5	4	12	11	
6	24	15	50	Fig. 4
7	40	15	60	---

6. Conclusions

The paper presents network failure detection algorithm and its usage in autonomous return task. The algorithm is fully compatible with ROS, does not need additional hardware or software installed on operator's PC and could be generalized for any TCP/IP connections. Both network failure detection method and autonomous return were experimentally validated on PMB-2 robot and showed their practical applicability for path planning[23].

Acknowledgements

This work was supported by the Russian Foundation for Basic Research (RFBR), project ID 19-58-70002.

References

1. K. Shabalina, et al., Comparative Analysis of Mobile Robot Wheels Design, in *Int. Conf. on Developments in eSystems Engineering* (IEEE, 2018), pp. 175-179.
2. I. Shimchik, et al., Golf cart prototype development and navigation simulation using ROS and Gazebo, in *MATEC Web of Conf.* **75** (EDP Sciences, 2016), p. 09005.
3. E. Magid, R. Lavrenov and I. Afanasyev, Voronoi-based trajectory optimization for UGV path planning, in *Int. Conf. on Mechanical, System and Control Engineering* (IEEE, 2017), pp. 383-387.
4. M. Sokolov, et al., Modelling a crawler-type UGV for urban search and rescue in Gazebo environment, in *Int. Conf. on Artificial Life and Robotics* (2017), pp. 360-362.
5. I. Farkhatdinov, JH. Ryu and J. Poduraev, A user study of command strategies for mobile robot teleoperation, in *Intel Serv Robotics* **2** (2009), pp. 95-104.
6. M. Wang and James N.K. Liu, Interactive control for Internet-based mobile robot teleoperation, in *Robotics and Autonomous Systems* **52**(2-3) (2005), pp. 160-179.
7. E. Magid, et al., Artificial Intelligence Based Framework for Robotic Search and Rescue Operations Conducted Jointly by International Teams, in *Int. Conf. on Electromechanics and Robotics "Zavalishin's Readings"* (Springer, 2019), pp. 15-26.
8. K. Nagatani, et al., Emergency response to the nuclear accident at the Fukushima Daiichi Nuclear Power Plants using mobile rescue robots, in *J. of Field Robotics* **30**(1) (2013), pp. 44-63.
9. P. Jordi, L. Marchionni and F. Ferro, Tiago: the modular robot that adapts to different research needs, in *Int. workshop on robot modularity IEEE IROS* (2016).
10. N. Alishev, et al., Network Failure Detection and Autonomous Return Algorithms for a Crawler Mobile Robot Navigation, in *Int. Conf. on Developments in eSystems Engineering* (IEEE, 2018), pp. 169-174.
11. A. Derbakova, N. Correll and D. Rus, Decentralized self-repair to maintain connectivity and coverage in networked multi-robot systems, in *Int. Conf. Robotics and Automation* (IEEE, 2011), pp. 3863-3868.
12. M. A. Hsieh, et al., Maintaining network connectivity and performance in robot teams, in *J. of field robotics* **25**(1-2) (2008), pp. 111-131.
13. D. Tardioli, et al., Enforcing network connectivity in robot team missions, in *Int. J. of Robotics Research* **29**(4) (2010) pp. 460-480.
14. A. H. Mong-ying, et al., Towards the deployment of a mobile robot network with end-to-end performance guarantees, in *Robotics and Automation* (IEEE, 2006), pp. 2085-2090.
15. PAL Robotics, official site: http://pal-robotics.com/
16. M. Quigley, et al., Sub-meter indoor localization in unmodified environments with inexpensive sensors, in *IROS* (IEEE, 2010), pp. 2039-2046.
17. D. Fox, Adapting the sample size in particle filters through KLD-sampling, in *Int. J. of robotics research* **22** (2003), pp. 985-1003.
18. global_planner, ROS package: http://wiki.ros.org/global_planner
19. A. Fernandes, Ad hoc communication in teams of mobile robots using zigbee technology, in *Computer Applications in Engineering Education* **23.5** (2015), pp. 733-745.
20. Ping utility, Ubuntu xenial package: https://packages.ubuntu.com/xenial/iputils-ping.
21. move_base, ROS package: http://wiki.ros.org/move_base
22. A. Zakiev, et al., Path planning for Indoor Partially Unknown Environment Exploration and Mapping, in *Int. Conf. on Artificial Life and Robotics* (2018), p.399-402.
23. E. Magid, et al., Combining Voronoi graph and spline-based approaches for a mobile robot path planning, in: *Informatics in Control, Automation and Robotics ICINCO 2017, Lecture Notes in Electrical Engineering,* ***495*** (Springer, Cham, 2020), pp. 475-496.

Design and Modeling of an Automatic Cartesian Farming Robot

Wisanu Jitviriya, Noppadol Pudchuen, Amornphun Phunopas
Faculty of Engineering, King Mongkut's University of Technology North Bangkok.
1518 Pracharat 1, Wongsawang Rd., Bangsue, Bangkok, Thailand 10800
wisanu.j@eng.kmutnb.ac.th, noppadol.pudchuen@gmail.com, amornphun.p@eng.kmutnb.ac.th

Eiji Hayashi
Mechanical Information Science and Technology, Kyushu Institute of Technology.
680-4, Kawazu, Iizuka, Fukuoka, 820-8502, Japan.
haya@mse.kyutech.ac.jp

Abstract

Currently, farming robots have become increasingly for households. The combination of modern technology and agriculture that is automatic system design, it is extremely accurate but also cost less than hiring human labor to work. All processes of the farming robot must be organic. The development of Cartesian Coordinate Robot (CCR) can operate in the multi-function such as tillage, applying fertilizer, sowing, and watering plants. In this research, which focuses on image processing to detect the plants and eliminate the weeds. The OpenCV library was used to detect the color green from the plant leaves, and from that, the program will do the rest. The elimination part of the system was constructed based on a brand-new idea. From all the experiments done, the conclusion looked promising, the blades were able to function at 80 percent efficiency.

Keywords: Smart farm, Cartesian robot, Weed detection.

1. Introduction

Supply and demand are the factors of an economic model that involves having enough supplies to meet the demands. Food is mainly grown in the agriculture industry. Agriculture utilizes the farming of land to produce crops, that will eventually become food for consumption. Fig. 1 shows a graph that breaks down all the usage of the entire land mass on this Earth. The figure shows that a huge portion of land is unusable due to its terrain. The land is made up of only 30% of the Earth's surface. That means that land is a limited resource on the Earth. To increase the area for growing crops, the world will have to sacrifice other portions of land. That is what makes the land a limited resource and therefore limited supply. The largest portion of demand for crops come from humans. Humans are the biggest consumers of food on Earth. Fig.2 represents the projected population growth as well as the population today and in the past. Assuming the graph is correct, the projected increase in the human population will cause more demands for the limited supply [1]. Currently, the robotics and automation systems are the most important and growing technology that can increase productivity in business and

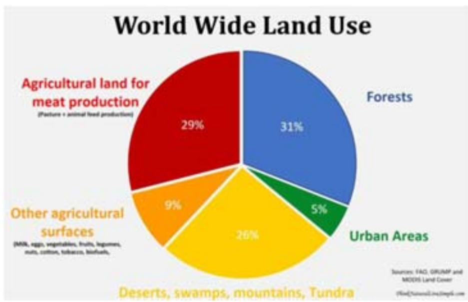


Fig.1 Land Usage in the World

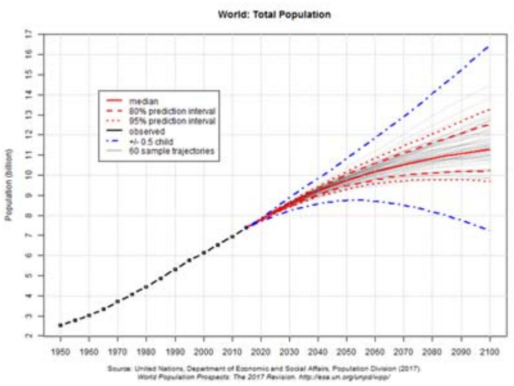


Fig.2 Population Growth

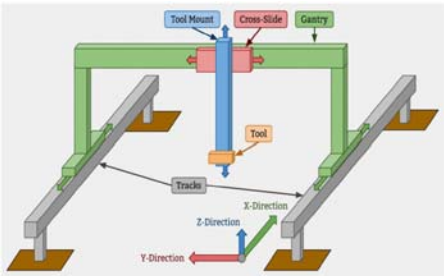


Fig.3 Mechanic diagram of FarmBot

reduce the working time of the workforce. It is fascinating if the researchers or inventors can apply their knowledge of robotics and automation technology to develop with together the agriculture industries. Therefore, that is the reasons, why the robots are imported to the agriculture industry is supply and demand. There are many researchers who proposed the new technologies combined with the agriculture industry, such as Wolfert et al. who propose the cyber-physical management cycle of the Smart Farming enhanced by cloud-based event and data management [2]. In the meanwhile, some researchers attempt to apply the image processing technic for recognizing different between weeds and crops for controlling weeds [3] – [5].

Recently, FarmBot [6] is designed to use in Smart Farming. The system is able to move in three axes; X, Y, and Z. This robot is done with a similar technology that has been a 3D printer and CNC milling machine and Fig.3 that shows FarmBot mechanic diagram. The main goal of FarmBot system is created to become a completely automated system from adding the seed in the soil, water the plant, detection the weeds and elimination them.

This research, which is divided into two parts. The first part is the hardware of the cartesian farming robot that is created similar to the FarmBot but, the tool head is redesigned for ease of use in farming. The second part is the software for controlling the cartesian farming robot. The robot can operate in the multi-function such as tillage, applying fertilizer, sowing, and watering plants. Moreover, the image processing is developed to detect the plants and eliminate the weeds. Our paper is organized as follows: the structure of the cartesian farming robot is explained in Chapter 2. In Chapter 3, the results are shown how the robot can operate and move to the target point accurately, and also the weed detection and elimination are verified. In the final section, we conclude the paper and describe our future works.

2. Structure of Cartesian Farming Robot

Motor Z-Axis
Control Box
Motor Y-Axis
Motor X-Axis
Web camera
Rubber drive belt
Tool head
Plastic Container

Fig.4 Cartesian Farming Robot

The main components of the hardware of cartesian farming robot are the motors, aluminum profiles, tool head, wheels, plastic containers, web camera, and control box. Fig. 4 shows how all components were assembled. The motors were a very important part because the mechanical parts are the ones receiving the heaviest load and control all the moving parts. There is a total of four motors that are used in this prototype. A rubber drive belts are used to line the track in the middle of the aluminum profile because rubber has high friction, and with the added gear like features of the drive belt, it makes it easier for the motor to turn the belt. The motors (X-axis and Y-Axis) are attached to a wheel and when that wheel is turned, it will not turn the belt, making only the mechanism to move along the axis of the aluminum profile. The container is used to put soil in grew some plants. The material for the container is plastic because it is light yet very strong. The multi-tools head of the robot has the seed injector, watering nozzle, soil sensor and Weeder which were designed in the compact tool head. The web-camera was used to detect the plants and weeds and all weeds will be eliminated by using Weeder.

3. Verification of the Cartesian Farming Robot

The proposed paper that presents the cartesian farming robot, which is designed and created to crop the plants and control the weeds automatically.

3.1. *Verification I*

- Accuracy and Repeatability testing

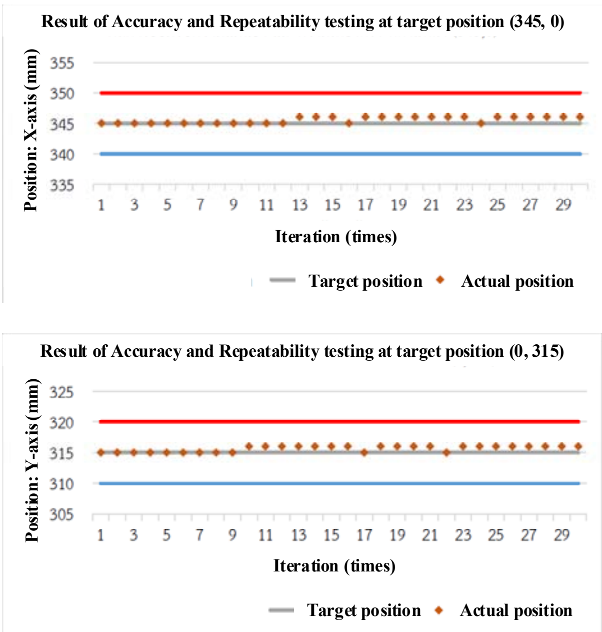


Fig.5 The results of robot movement in X-axis and Y-axis

As for the experimental method of accuracy and repeatability testing, the tool head is modified to attach the permanent pen. The robot is controlled to draw the point on the test plate (repeating 100 times for 1 target position). Fig.5 shows the results of the accuracy and repeatability testing that can confirm the effectiveness of the robot movement.

3.2. *Verification II*

- Detection plants and weeds

The detection software utilized the OpenCV library to help detect the pigments of the plants and weeds. The process was to have an ordinary web camera take a photo of the plant at the position that was predetermined and then has the OpenCV library program detected the pigments of the plants and the weeds in that frame. The diameters of the plants and the weeds were also found. After both of the diameters were found, a point of the center could be determined for each circle. It was important to locate the center point because it will be the point on the x-axis and y-axis that the program will be referencing to eliminate the weeds. After the program had determined all the diameters of the color green, the program will determine which circle was the plants and which circle was the weeds and will mark all the plants with blue circles and the weeds circles in red.

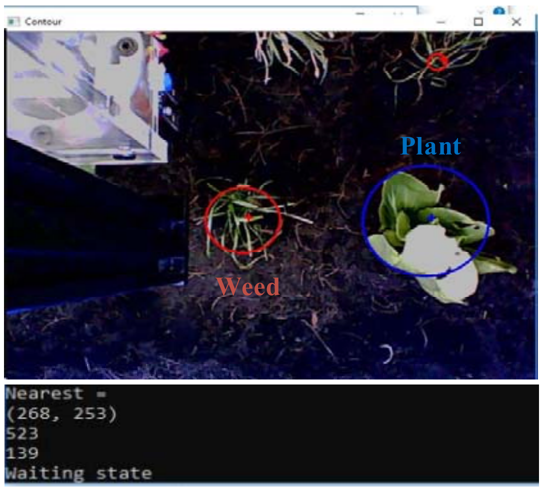


Fig.6 Detection program

TABLE I. PLANTS MEASUREMENT

No.	Hand Measure (x, y) cm	Program Measure (x, y) cm	% Error (%X, %Y)
1	(53, 15)	(54, 17)	(1.89, 13.33)
2	(64, 1)	(65, 0.5)	(1.56, 50)
3	(50, 15)	(52, 13)	(4, 13.33)
4	(66, 10)	(64, 12)	(3.03, 20)
5	(70, 68)	(73, 67)	(4.29, 1.47)

The experiment was about testing the accuracy of the program. The first step was measuring the actual position of where the plants were by hand and compared that to what the data showed. Table I and Fig. 6 is the chart that showed the results and an example of what the camera saw when it took a picture. The results were based on the origin of the tool head when it was at the home position.

- Elimination Weeds

Then, in order to completely terminated the weeds, it will not be advisable to pierce the tool head into the soil only once. Instead, it will better to create multiple points around an area where the program had detected the weeds. Fig.7 shows that there will be a total of nine points around the area where the tool head (4 blades) was going to pierce the soil. This experiment was to test the elimination of weeds in ten different spots. The definition of elimination was when all leaves were separated from their stems. Fig. 8 presents a picture of the weeds that were going to be destroyed by the four blades.

4. Conclusion

The research started by studying why weeds are such a destructive force for crops, and it was found that weeds were any unwanted plant species that were undesirable. This problem occurred because weeds would take nutrients away from the main crops, not allowing it to grow at its maximum capacity. There are two main components, the camera and the tool head for weed elimination. The camera was just an ordinary web camera because the software did most of the work and not the camera. The second component was the tool head for weeds elimination. The tool head had four thin blades attachment on the bottom in order to cut the weeds. The z-axis moved downward very slowly and with very little force, so the blades had to be extremely sharp in order to cut the weeds. From all the experiments done, that confirm the automatic cartesian farming robot can operate in the multi-function accurately and correctly.

5. Acknowledgment

This research was funded by King Mongkut's University of Technology North Bangkok. Contract no. KMUTNB-62-NEW-17.

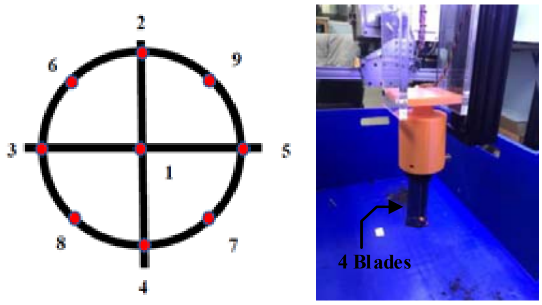


Fig.7 Piercing Position and tool head

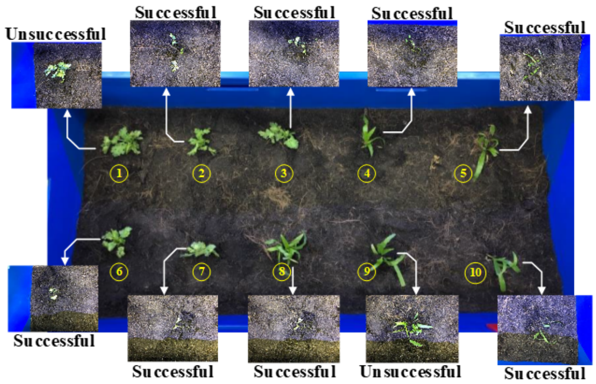


Fig.8 Results of 4 blades for elimination

6. References

1. "World Population Prospects - Population Division." United Nations, United Nations, population.un.org/wpp/Graphs/Probabilistic/POP/TOT/.
2. Wolfert, J., Sørensen, C.G., Goense, D., "A Future Internet Collaboration Platform for Safe and Healthy Food from Farm to Fork," Global Conference (SRII), 2014 Annual SRII. IEEE, San Jose, CA, USA, pp. 266–273. 2014.
3. Astrand, B., & Baerveldt, A. J. "An agricultural mobile robot with vision-based perception for mechanical weed control. Autonomous Robots, vol. 13, pp. 21-35, 2002
4. Søgaard, H. T., "Weed classification by active shape models Biosystems Engineering," vol. 91, no. 3, pp. 271-281, 2005.
5. Slaughter, D. C., Giles, D. K., Fennimore, S. A., & Smith, R. F., "Multispectral machine vision identification of lettuce and weed seedlings for automated weed control," Weed Technology, vol. 22, no. 2, pp 378-384, 2008.
6. Aronson, Rory L. Farmbot: Humanity's Open-Source Automated Precision Farming Machine. Tech. Farmbot, 19 Sept. 2013. Web. 21 Jan. 2014.

K-APF Algorithm to Avoid Obstacles in Path Planning

Dong-Kyo Jeong
Dep. Of Electronic Eng., Pusan National University, 2, Busandaehak-ro 63beon-gil Geumjeong-gu
Busan, 46241, Republic of Korea

Jang-Myung Lee
Dep. Of Electronic Eng., Pusan National University, 2,Busandaehak-ro 63beon-gil Geumjeong-gu
Busan, 46241, Republic of Korea

E-mail: dongkyo1696@pusan.ac.kr, jmlee@pusan.ac.kr
www.pusan.ac.kr

Abstract

In this paper, a K-APF algorithm has been proposed to resolve the local minima and unstable path problems that occur when the conventional APF (Artificial Potential Field) algorithm is used for path planning in a real environment. Which is the most popular path planning algorithm for robot manipulators and mobile robots, to avoid the local minima with the conventional APF, repulsive coefficients have been added to the potential field and the Kalman filter has smoothed the unstable path. In the main page, the construction and operation of existing APF algorithms and these Algorithms problems will be described. Also, in order to demonstrate the performance of the proposed algorithm, compared simulation with existing APF algorithm and compared to existing APF, A*, and Dijkstra algorithms through the real experiments in terms of accuracy and speed.

Keywords: Artificial Potential Field, Path Planning, Kalman Filter, A* Algorithm, Dijkstra Algorithm.

1. Introduction

To drive safely in autonomous driving, it is essential to acquire the external environment information using sensors and to create the map based on and then perform the optimal path planning. Algorithms used for path planning are typically A * algorithm and Dijkstra algorithm. The A * algorithm performs searching based on the experience of moving, which greatly reduces the size of the search space and thus shortens the computational speed. In a space where there is no frequent contact or experience with obstacles, the route must be search again from the beginning [1-3]. Dijkstra algorithm has the feature of accurately finding the shortest path in the path to be moved, but it takes a relatively long time to search.

In this paper, I used the APF (Artificial Potential Field) as an algorithm that is less affected by obstacle contact and capable of faster operation than the above two algorithms. APF is an algorithm that is used to move paths without collisions when obstacles exist in the work path, mainly mobile robots or robot manipulators. However, APF is an algorithm that moves relatively short range in a limited platform, robots, so when it is used as a path planning algorithm at a real distance, there existed a case where a planned route or an unnecessary route penetrated. To solve this problem, this paper applies Kalman Filter to APF and compares it with other algorithms while presenting a new algorithm called K--APF.

The paper is organized as follows. Chapter 2 presents the technical descriptions of the APF algorithm and Kalman Filter used in avoidance path planning, and Chapter 3 presents the K-APF algorithm. In the next Chapter 4, I have verified the performance of K-APF algorithm through experiments and compare it with A *, Dijkstra, and existing APF algorithm. In the last Chapter 5, the conclusion was composed.

2. Artificial Potential Field & Kalman Filter Algorithm

The APF algorithm generates a virtual force field based on the attractive force between the starting point and the arrival point, the repulsive force between the current position and the obstacles. The sum of these two forces acts to avoid collisions with obstacles. The APF algorithm can be divided into Attractive Potential Field, Repulsive Potential Field, and Universal Potential Field

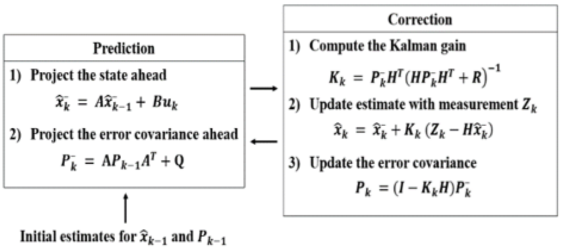


Fig. 1. Kalman Filter

[4-6].

2.1. *Attractive Potential Field*

In the case of Attractive Potential Field(F_{att}), it is a space that creates virtual manpower to reach the target point. It corresponds to Eq. (1).

$$F_{att} = G\sqrt{(x_{goal} - x)^2 + (y_{goal} - y)^2}. \quad (1)$$

x_{goal} and y_{goal} are the x and y axis coordinates of the destination and x and y are the x and y axis coordinates of the current position. G is the attraction coefficient, which is the amount of force pulling to the point of arrival.

2.2. *Repulsive Potential Field*

Repulsive Potential Field (F_{rep}) creates a virtual repulsive force between an obstacle and its current position so that it does not collide with the obstacle between the movement of the control target and the target. It corresponds to Eq. (2).

$$F_{rep} = \frac{O}{\sqrt{(x_o - x)^2 + (y_o - y)^2}}. \quad (2)$$

x_o, y_o is the x and y axis coordinates of the obstacle and O is the repulsive coefficient that is the degree of pushing force of the obstacle. In particular, the information in the field recognizes the surrounding environment through external sensors and collects the position of the obstacle and applies it to the coordinates.

2.3. *Universal Potential Field*

The Universal Potential Field (F_{Uni}) is a combination of the two forces, which creates an overall path plan in the field, with a path plan that avoids collisions from high potential to low potential. The equation for that corresponds to Eq. (3).

$$F_{Uni} = F_{att} + F_{rep}. \quad (3)$$

In the case of Eq. (4). below, the algorithm is used to terminate the algorithm without further calculation depending on how close the current position is to the arrival point. At this time, considering the repulsive force of the obstacle, set the Z coefficient to a range that is not affected.

$$D_{goal} = \sqrt{(x_{goal} - x)^2 + (y_{goal} - y)^2} \leq \text{Z}. \quad (4)$$

2.4. *Kalman Filter Algorithm*

Kalman Filter is shown in Fig. 1. it is divided into the prediction process and the estimation process. In detail, it can be divided into the initial value and the estimated value and the error covariance prediction, the Kalman gain calculation, the estimation value calculation, and the error covariance calculation. The characteristic of Kalman Filter algorithm is to repeat the process while changing the system model until the estimated value derived through this process is similar to the desired value [6-7].

Fig. 1. B_{u_k} and Q which are used in estimating selection, are system noise in the algorithm. In this paper, it was set to 0 and the rest of A, H, and R were used as the system model variables.

3. K-APF Algorithm

When planning the path based on the attraction and repulsion using the APF algorithm, the forces collide with each other, resulting in a point where the result is zero. This is called local minima. In local minima, when the path becomes narrower, and the distance between nearby obstacles becomes closer, the same repulsive force affects each other, and the difference between the two forces becomes zero. At this time, the planned algorithm is not affected by the force, it stops and trapped at that position. This problem was caused by using all the same repulsive force values of obstacles in the existing APF algorithm. In this paper, decided to use the repulsive force according to the size of obstacles. Eq. (2). F_{rep} was redefined in Eq. (5).by adding the coefficient of repulsion magnitude according to the building size.

$$F_{rep} = \frac{1}{2} O\left(\frac{1}{D_o} - \frac{1}{O_W}\right) \quad (5)$$

$$D_o = \sqrt{(x_o - x)^2 + (y_o - y)^2} \quad (6)$$

In the case of O_w is the range factor of the building size and D_o is the difference of the current position according to the obstacle position. In addition, there is one more problem that should be considered when path planning using APF algorithm. Fig. 2. When there is a curved part of an obstacle as shown, even if the same repulsive force is applied to both sides of the part, the force does not reach the path of the center. So when the path is drawn, it finds the range where the repulsive force acts and then writes the inside of the curved part.

In this paper, the Kalman Filter is combined with the APF algorithm to solve the problem and to plan the path stably. Fig. 2. When calculating the path of the measured value, the path biased to the curved part is selected to be formed by the measured noise. Therefore, Kalman Filter was used to reduce the measurement noise and to make the estimated value close to the existing path value so that the smooth path can be obtained.

$$Z_k = kx_k + v_k - w \quad (7)$$

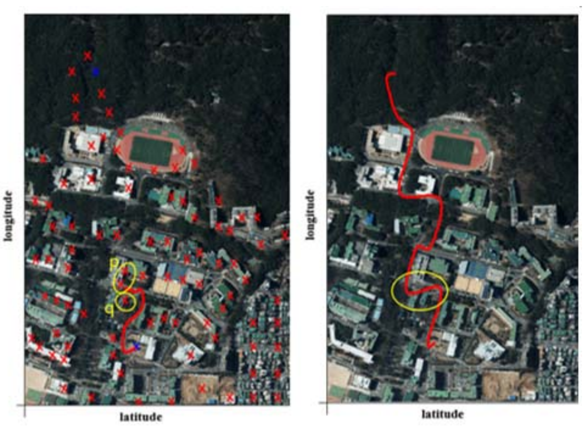


Fig. 3. Local minima of APF Algorithm and solved

Eq. (7). k is the inherent force constant of the algorithm applying the Kalman Filter and x_k is the final position of the path minus the current position of the path. The v_k value is the average value of the position where the path is drawn in the curved portion, and the w value is shown in Fig 2 as the noise value minus the a and b positions.

4. Experiment and Analysis

Points expected to be obstructions are marked with a red X and starting and destination points are marked with a blue X. In this case, the experiment was conducted with the blue mark on the bottom as the starting point and the point on the top as the arrival point. First, path planning was carried out through Eq. (3) to implement APF algorithm. As a result, Fig. 3. It is written as the left picture of but the algorithm is stopped by local minima while path is being formed. Accordingly, change F_{rep} in Eq. (3) to the content of Eq. (5), so that the p building located at the top of the local minima section is larger in size than the q building located at the bottom, so that the

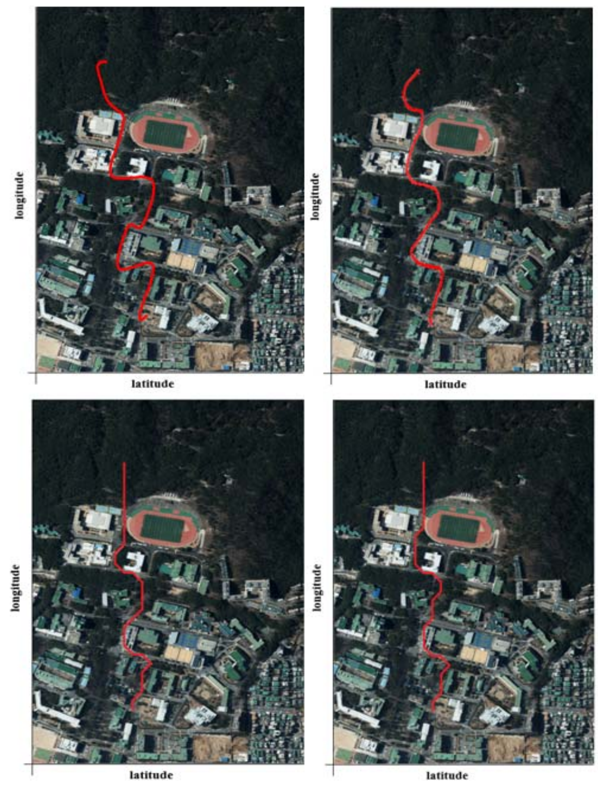


Fig. 4. Path planning results by Algorithm

repulsive coefficient O_w of the p building is doubled given. The problem was solved and the path to the arrival point was formed as shown in the right picture of Fig. 3.

However, even though the problem of the local minima was solved, in the Fig. 7. there was a section where the path was off the road like a yellow circle. The problem is solved by applying the aforementioned Kalman Filter to the APF algorithm. Eq. 8 was rewritten by combining Eq. 7 with the Kalman Filter to F_{att} of the existing Eq. 1.

$$F_{katt} = F_{att} + Kalman(Z_k) \quad (8)$$

$$F_{Uni} = F_{katt} + F_{rep} \quad (9)$$

After that, build Eq. (3) as if I was calculating Eq. (9) and K-APF algorithm was executed. The results of the K-APF algorithm are shown in Fig. 4. together with the A * algorithm and Dijkstra algorithm, which are frequently used for APF algorithm and path planning.

From the top left of Fig. 4. K-APF, APF at the top right, Dijkstra at the bottom left, and A * algorithm at the bottom right are shown. To compare the results, the distance, time and accuracy from the starting point to the arrival point is divided and shown in Table 1 below.

Table 1. Comparison by Model

Model	Distance(m)	Time(min)	Accuracy(%)
A*	915	13.72	84.21
Dijkstra	928	3.91	78.95
APF	935	14.03	89.47
K-APF	928	13.92	94.76

In Table 1., the distance measured for each algorithm is selected as the distance value for each distance and in terms of time, when a person normally walks for 1 hour, the speed takes 4 km/h, which can be converted to 66.67 m/min in terms of walking time per minute. The distance divided by the speed was chosen as the time value by using the speed and the preceding distance. Finally, the accuracy was measured as Eq. (10) depending on whether the obstacles were interfering based on the obstacles around the path when the path planning work was performed on the map.

$$\text{Accuracy} = \frac{Number\ of\ obstacles\ that\ interfere\ with\ the\ path}{Number\ of\ obstacles\ in\ the\ path} \quad (10)$$

5. Conclusion

In this paper, I solved adding a repulsive coefficient to solve the local minima problem that occurs when applying APF algorithm to real environment. In addition, Kalman Filter was added to the APF algorithm to alleviate the curved shape of the path. A new algorithm that improves these two characteristics is defined as K-APF algorithm and compared experimentally with conventional algorithms. The results confirmed that the probability of safe driving without obstacle contact increased by 5% ~ 16% when moving to the correct destination point in the same background. The algorithm is expected to be applied with high accuracy if it is used in connection with the task that reaches the destination by minimizing human interference in the area with many obstacles.

Acknowledgements

This research is based upon work supported by the Ministry of Trade, Industry & Energy (MOTIE, Korea) under Industrial Technology Innovation Program. No.10073147

References

1. Jun-Hyuck Im, Seung-Hwan You, Gyu-In Jee, Dal-Ho Lee, A Path Generation Algorithm for Obstacle Avoidance in Waypoint Navigation of Unmanned Ground Vehicle, *Journal of Institute of Control, Robotics and System*, Vol. 17, No. 8, pp.843-850, 2011.
2. Chi-Beom Noh, Min-Ho Kim, Min-Cheol Lee, Path planning for the shortest driving time considering UGV driving characteristic and driving time and its driving algorithm, *Journal of Korea Robotics Society*, Vol. 8, No. 1, pp.43-50, 2013.
3. Yong-Hu Lee, Sang-Woon Kim, A hybrid search method of a* and dijkstra algorithms to find minimal path lengths for navigation route planning, *Journal of the Institute of Electronics and Information Engineers*, Vol. 51, No. 10, pp.109-117, 2014.
4. Dong-Ju Park, Dong-Eon Kim, Jin-Hyun Park, Jang-Myung Lee, Singlarity and Collision Avoidance Path Planning based upon Artificial Potential Field and Manipulability Measure, *Journal of The Institute of Electronics and Information Engineers*, Vol. 55, No. 5, pp.89-96, 2018.
5. Seung-Gyu Ahn, Dong-Jin Lee, Collision Avoidance for UAV using Potential Field based on Relative Velocity of Obstacles, *Journal of the Korean Society for Aviation and Aeronautics*, Vol. 26, No. 2, pp.47-53, 2018.
6. Ho-Sun Kang, Jong-Woo An, Hyun-Soo Lim, Seul-Woo Hwang, Yu-Yeong Cheon, Eun-Han Kim, Jang-Myung Lee, Localization Algorithm for Lunar Rover using IMU Sensor and Vision System, *Journal of Korea Robotics Society*, Vol. 14, No. 1, pp.65-73, 2019.
7. Young-Sik Park, Jun-Woo Song, Dong-Hyuk Lee, Jang-Myung Lee, Path Estimation Method in Shadow Area Using Underwater Positing System and SVR, *Journal of Korea Robotics Society*, Vol. 12, No. 2, pp.173-183, 2017.

Selection System of Robot type for cell assembly production
(Production efficiency comparison of single arm robot and double arm robot)

Takahiro Watanabe
Department of Human Information Systems, Gifu University, Yanagido 1-1,
Gifu, Gifu 501-1194, JAPAN

Hidehiko Yamamoto
Department of Mechanical Engineering, Gifu University, Yanagido 1-1
Gifu, Gifu 501-1194, JAPAN

Takayoshi Yamada
Department of Mechanical Engineering, Gifu University, Yanagido 1-1
Gifu, Gifu 501-1194, JAPAN

E-mail: x4525098@edu.gifu-u.ac.jp, yam-h@gifu-u.ac.jp, yamat@gifu-u.ac.jp

Abstract

The purpose of this research is to compare the production efficiency of single arm robot and double arm robot. This research determines which robot is suitable for an assembly cell production, a single arm robot or a double arm robot. First, we develop the system to use a double arm robot to determine the best parts location by genetic algorithm (GA). The system consists of two modules, the conditions module and GA module. The conditions module reads work data for robot and sets up various parameters required for GA. The GA module decides the efficient arrangement place of parts, jigs and robot hands by GA and outputs the acquired arrangement visual images. Next, we use a single arm robot at the same parts location and compare the working time with a double arm robot. Finally, we conclude which robot is suitable for an assembly cell judging from the results of a total robot working time.

Keywords: genetic algorithm, Unit arrangement decision, Dual arm robot, Assembly machine

1. Introduction

In recent years, it has been demanded to produce a large quantity of products in a short time at the production site. Automations of the assembling works by robots has been also carried out. When developing an automatic assembly machine including robots, the production efficiency depends on the arrangement of each unit of the assembly machine. However, in many factories, the placement of units is based on the experienced engineers It remains questionable whether the determined placement by the experienced engineers is really efficient. In this research, the layout decision system that assists unit layout determinations of an efficient cell type assembly machine is developed by using genetic algorithm (GA). The system compares the production efficiency of a single arm robot and a double arm robot and determines which robot is suitable for the assembly job.

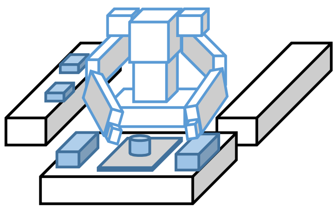


Fig. 1 Cell type assembly machine of double arm robot

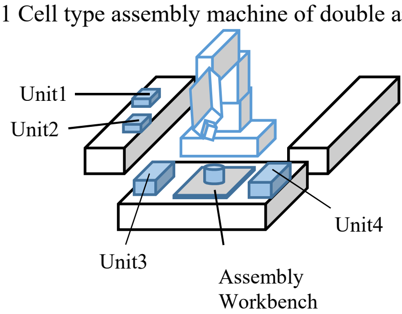


Fig. 2 Cell type assembly machine of single arm robot

2. Cell type assembly machine

As shown in Fig. 1 and Fig. 2, the cell type assembly machine to be used in this research has a configuration in which an assembly machine is arranged at the center, a work stand at the front and assembly units on the periphery. After each arm chucks the needed unit, the arm carries it to the workbench to assemble the parts. In the research, approach 1, approach 2 and the chuck position of the robot arm movement route are defined as the three points that exist on the arm moving route as follows

- Chuck position ... potision for chuking and unchuking parts
- Approach 1 ... position upward away from the chuck position
- Approach 2 ... position further upward than Approach 1

As shown in Fig. 3, the arm passes from approach 2 to approach 1 and arrives at the chuck position.

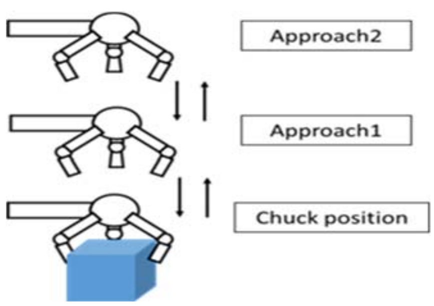


Fig. 3 Movement of arm

3. Layout determination system

The layout decision system we develop decides the arrangement of each unit with the best operation efficiency during designing a cell type part assembly machine by using GA. The system consists of two modules, the condition creation module and the GA module. The condition creation module performs to read the arrangement possible area of the unit and work contents, to divide the arrangement possible area and to decide various necessary GA parameters. The GA module determines the arrangeable place of the units, evaluates the working time and outputs coordinates of the arrangement place of the unit and the layout image of the unit.

3.1 Condition creation module

The procedure of the condition creation module is as follows.

Step Ⅰ: Read work data of parts assembly, placement possible area, parts arrangement data and create their database.

Step Ⅱ: Divide possible placement area into a lattice shape to create possible placement place.

Step Ⅲ: Carry out genes coding.

Step Ⅳ: Define the fitness function.

Step Ⅴ: Determine various parameters.

3.2 GA module

The process flowchart of the GA module is shown in Fig. 4. The process procedure is shown below.

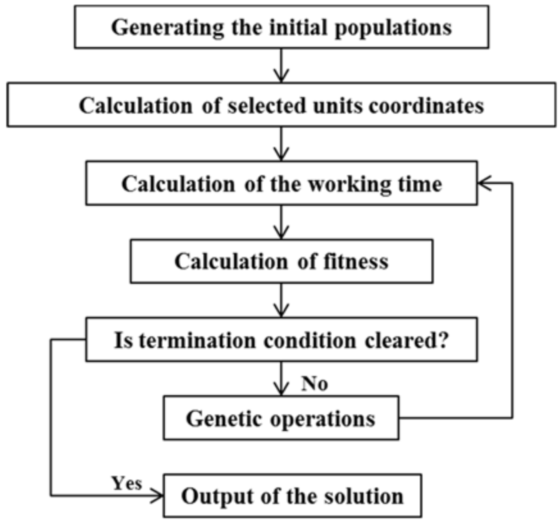


Fig. 4 Flowchart of GA module

Step 1: Generate an initial population.

Step 2: Calculate the placement coordinates of each unit.

Step 3: Calculate the time taken for each work.

Step 4: Calculate fitness.

Step 5: Apply genetic operations (selection, crossover, mutation).

Step 6: Judge whether the termination condition is satisfied. If it is, go to Step 8. If not, go to step 3.

Step 7: The unit arrangement of individuals with the smallest operation time is adopted as the optimum arrangement and the output as a solution.

A single arm robot comparison procedure.

Step 1: Calculate the unit optimal placement coordinates for 10 times and each work time with a double arm robot using GA.

Step 2: The same work is performed on the industrial robot with the optimal placement coordinates for 10 times obtained with the double-arm robot, and the work time for each 10 times is obtained.

Step 3: Compare work times obtained by both robots to find work efficiency.

Step 4: Determine which robot is suitable for cell-type automatic assembly production from the comparison of work time and work efficiency.

Using the results of the double arm robot above mentioned, the single arm robot performs the same work above. The single arm robot is set with the same conditions such as the arrangement coordinates and the work process and arm speeds. As shown in Fig.2, the robot arm approach speeds are changed. Table 1 shows the both robot speed.

Table 1. Each arm speed

	Don't have parts[mm/s]	Have parts[mm/s]
Approach2-2	770	380
Approach2-1	400	250
Approach1-chuck	25	50

4. Application examples and performance comparisons

The layout decision system was applied to the assembly machine and the production simulations were carried out. As a comparison, simulations of the following two systems were carried out.

・The system that determines unit layout by GA

・The same work with a double arm robot and a single arm robot

The units that needs to determine the placements have eight left tray, right part tray, left parallel hand, left three nail hand, right parallel hand, right four nail hand, jig A, jig B and also the assembly work. The layout map is shown in Fig. 5. Fig. 6 is the left and right part trays with 52 parts. Each number represents each part.

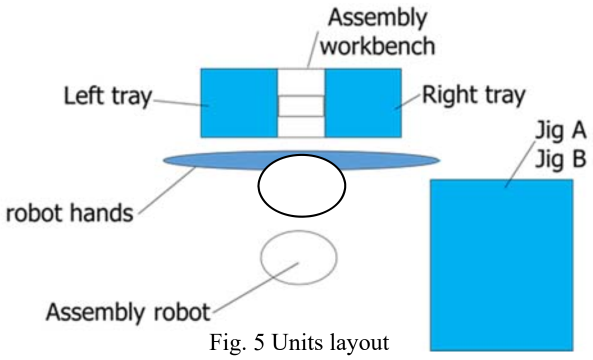


Fig. 5 Units layout

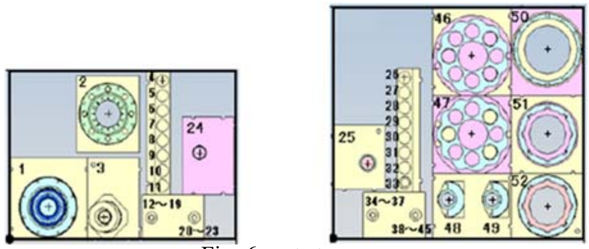


Fig. 6 parts trays

Next, Table 2 and Table 3 show the position data and placeable area data of both robot.

Table 2. Coordinates of double arm robot

	x[mm]	y[mm]	z[mm]
Robot	2,210	915	740

Table 3. Placement area of each unit

Unit	x1[mm]	y1[mm]	z1[mm]	x2[mm]	y2[mm]	z2[mm]
Left tray	1,811.5	1,760	740	1,918.5	1,804.5	740
Right tray	2,522	1,806.5	740	2,628	1,807.5	740
Left parallel hand	1,280	1,480	629	2,320	1,480	629
Left three claw hand	1,280	1,480	629	2,320	1,480	629
Right parallel hand	2,100	1,480	629	3,140	1,480	629
Right four claw hand	2,100	1,480	629	3,140	1,480	629
Jig A	3,034	83	740	3,656	1,117	740

Jig B	3,034	83	740	3,656	1,117	740

5. Simulation results

Simulations were carried out ten times in each of the three systems. The results of the simulations are described below.

・Minimum cycle time：
The smallest work time in each simulation

・Average cycle time：
Average work time of simulation for 10 times

・Average simulation time：
Average time taken for 10 simulations

Fig. 7 shows the output drawing of the unit arrangement with the highest efficiency. Table4 shows the simulation times and the average times.

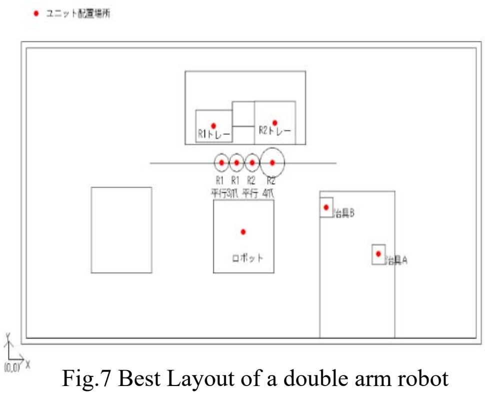


Fig.7 Best Layout of a double arm robot

Table 4 shows the best position of units. Table 5 shows the both simulation times and the average times.

Table 4. Best position of unit

Unit	x1[mm]	y1[mm]	z1[mm]
Left tray	1915	1775	740
Right tray	2527	1807	740
Left parallel hand	2055	1480	629
Left three claw hand	2205	1480	629
Right parallel hand	2365	1480	629
Right Four claw hand	2565	1480	629
Jig A	3379	708	740
Jig B	3039	1115	740

Table 5. Both robot simulation times

	A double arm robot	A single arm robot
One time	237.253	291.591
Two times	237.274	291.602
Three times	237.265	291.585
Four times	237.265	291.595
Five times	237.252	291.595
Six times	237.251	291.586
Seven times	237.267	291.590
Eight times	237.300	291.593
Nine times	237.256	291.590
Ten times	237.251	291.586
Average times	237.257	291.585

6. Testing

The average difference of the average work time of 10 times obtained by each robot was tested. In the alternative hypothesis, the difference between the average value of the working time of the double arm robot was tested and the average difference was tested. In both cases, the null hypothesis was rejected and the alternative hypothesis results that there was a difference in the working time of two robots.

7. Conclusions

The layout decision system developed in the research is the system to improve design and production efficiency by automating the unit arrangement determinations of a cell type assembly machine by GA. The system compares the production efficiency of a single arm robot and a double arm robot and determines which robot is suitable for assembly.

Comparing the working time of a double arm robot and a single arm robot, it was found that a double arm robot is more efficient in this assembly job. It is better to use a double arm robot because of its work efficiency.

It is ascertained that the developed system is more useful to use a double arm robot with this equipment.

Curvature Surface Magnetic Wheel Climbing Robot with Adaptive Electromagnetic Adhesive Force

Richit Palangwatanakul, Apisit Thungsang, Kaned Thungod*

Faculty of Engineering, Mahasarakham University, Mahasrakham, 44150, Thailand†

Arsit Boonyaprapasorn, Wanachart Borisut

Dept. of Mechanical Engineering, Chulachomklao Royal Military Academy, Nakhon-Nayok, 26001, Thailand

Thavida Maneewarn, Suriya Natsupakpong

Institute of Field Robotics, King Mongkut's University of Technology Thonburi, Bangkok, 101401, Thailand

Thunyaseth Sethaput

Dept. of Mechanical Engineering, Sirindhorn International Institute of Technology, Thammasat University, Pathum Thani, 12120, Thailand

E-mail: kaned.th@msu.ac.th

www.msu.ac.th

Abstract

Various industrial structures or machines mostly consist of different shapes of ferromagnetic curvature surfaces. The magnetic wheel climbing robot is the suitable approach for achieving both adhesion and locomotion of the inspection robot. However, the adjustable magnetic force for robot adhesion is necessary, especially when the thickness of the surface is not uniform or the variation of the air gap between the magnetic adhesion units caused by the curvature of the surface. This can lead to the insufficient adhesive force. Furthermore, unnecessary driving torque of the motor to actuate the climbing robot from the over design of the magnetic adhesive force from the magnetic wheels can be avoided. Due to the level of the adaptive adhesive force is necessary to be considered, we designed the adaptive electromagnetic adhesive force mechanism for the curvature surface climbing robot with magnetic wheels. The PID controller was employed to control the electromagnetic force, and the adhesive force was measured by a load cell. This measurement signal was used as a feedback signal. In the paper, we investigated the capability of this adjustable magnetic force system. Five aspects of experimentation were implemented. It was clear that the light weight electromagnetic force adjustment mechanism could provide the flexibility to regulate the adhesive force for the magnetic robot while traveling on the ferromagnetic curvature surface.

Keywords: Climbing robot, Magnetic wheel, Locomotion design, Adaptive electromagnetic adhesive system.

*242 Khamriang Sub-District, Kantarawichai District, Maha Sarakham Province, 44150 Thailand.

1. Introduction

Industrial storage tanks have been widely used as oil or chemical reservoirs which are often made up of ferromagnetic materials such as steel. To inspect the welding defect and internal crack to prevent accidental leak of the reservoir, the wall climbing robot has been introduced. The robots play an important role to replace human operation in risky tasks and limitations such as the complex structure and curvature surface, high altitude, narrow space, and hazardous environments. Besides, the purpose of inspection[1-3], the wall climbing robot also used in various applications: labeling[4], rust removal[5], grit blasting[6], painting[7], and welding[8].

The magnetic wheel climbing robot is the suitable approach for achieving both adhesion and locomotion of the inspection robot. In addition, magnetic-based adhesive climbing robot provides faster response to operate compared to suction-based mechanism due to magnetic-based mechanism requires a shorter time for sufficient adhesive force[9]. However, the adjustable magnetic force for robot adhesion should be considered, especially when the thickness of the surface is not uniform or the variation of the air gap between the magnetic adhesion units caused by the curvature of the surface. This can lead to the insufficient adhesion force. Furthermore, unnecessary driving torque of the motor to actuate the climbing robot from the over design of the magnetic adhesive force from the magnetic wheels can be avoided.

This study aims to design two-module wall climbing locomotive robot prototype with adaptive electromagnetic adhesive force based on PID controller. The robot must be enabled in climbing smooth vertical ferromagnetic surface and moving along ferromagnetic curvature surface.

2. Design and Control of a Climbing Robot

The magnetic wheel climbing robot presented in this paper was designed as two-module locomotive robot and consists of six major components as shown in Fig. 1. First, the robot structures were formed by 3D printer with 3 mm PLA filaments. The dimension of robot was 275 mm length x 166 mm width x 116.75 mm height. Total weight of the robot was 4.56 kg. Two separate modular design of robot allowed the robot to move along a curvature surface. Second, six magnetic wheels were used to increase the adhesive force between the robot and the steel surface. Each wheel was constructed from 8 pieces of 10 x 10 mm neodymium magnet and two 75-mm diameter outer wheel discs. The outer wheel discs were assembled to avoid direct contact between the neodymium magnet and the ferromagnetic surface. Third, three 12-Volts, 25-kg lifting force, electromagnet were used to provide an adhesive force and able to adjust the force to desired level. Each electromagnet was placed between each pair of wheels. Fourth, three load cells were placed between the structure and the electromagnet. The load cell was used as a force sensor to detect the adhesive force from the electromagnet. Fifth, four sets of 12V DC motor and worm gears were used to control the electromagnetic adhesive force and drive the robot wheels. Lastly, to control the robot, an Arduino-based PID controller was implemented which can be divided into two operating parts: 1) the electromagnetic adhesive force control and 2) the robot motion control. The electromagnetic adhesive force can be controlled using the force provided by the load cell, and then adjust the force of the electromagnet. The robot motion can be controlled by derived the input signals from a joystick wireless remote control. Then, the digital signals were sent to the Arduino board to process and sent the duty cycle of output signals in order to drive DC motors and the wheels through the worm gears.

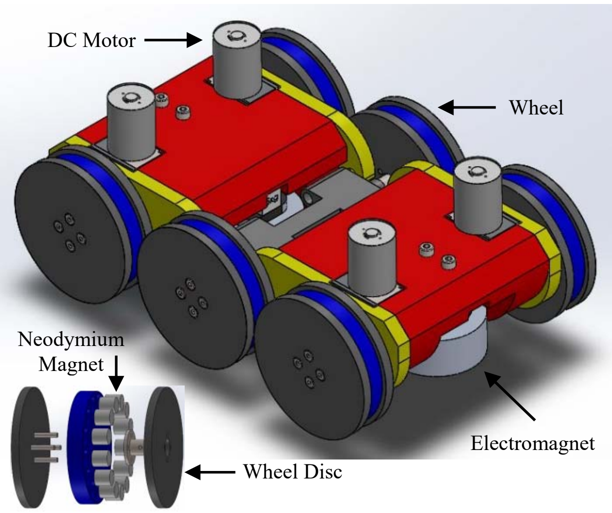


Fig. 1. Schematic of the magnetic wheel climbing robot.

3. Experimental Results and Discussion

In this paper, five experiments were conducted and presented as following:

3.1. *Performance of PID controller*

Tests of PID control performance were conducted to evaluate the ability of PID controller for adjusting electromagnetic adhesive force to reach the target value. In this test, the target force of 0.03 grams was assigned. Zero gram of adhesive force was applied as an initial condition. As a result shown in Fig. 2, the adhesive force powered by electromagnet can achieve the target force within 40 seconds. Also, when the target forces of 0.02 and 0.04 grams were assigned, the adhesive force was able to track the target within 40 seconds. The results showed the adaptive ability of PID controller.

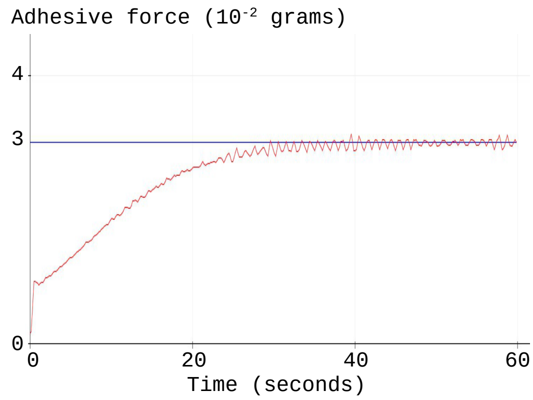


Fig. 2.PID control for electromagnetic adhesive force.

3.2. *Electromagnetic adhesive force*

Tests of adhesive force provided by electromagnet were performed by placing a 1.5 mm acrylic sheet between the electromagnet and a steel sheet as shown in Fig. 3. We found that the average electromagnetic adhesive force was 0.52 kg.

3.3. *Magnetic force of neodymium magnets*

For each wheel, eight pieces of neodymium magnet were plugged inside round drum magazine where placed between two outer wheel discs. A spring weighing scale was used to measure the adhesive force provided by neodymium magnetic wheel. The average adhesive force of 11.58 kg was measured. The measured force met the minimum required adhesive force from calculation. The minimum required adhesive force can be simply derived from the following equation:

$$F_m = mg\left(\frac{1}{\mu} + \frac{d}{b}\right). \qquad (1)$$

where F_m is the required adhesive force (N), m is mass of the robot (4.56 kg), g is gravitational acceleration (9.81 m/s^2), μ is friction coefficient (0.6), d is the perpendicular distance from the robot's center of mass to the surface (30.2 mm), and b is the robot's width (166 mm). After calculation, the minimum required adhesive force of the climbing robot is 82.71 N or 8.27 kg. We observed that eight pieces of neodymium magnet in each wheel can overcome the minimum required magnetic force and suitable for the robot.

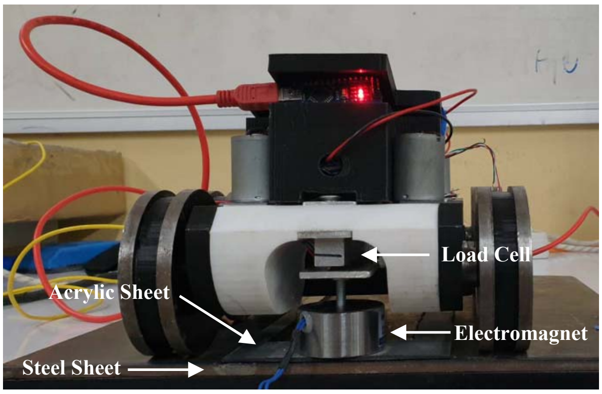


Fig. 3. Experimental setup for eletromagnetic adhesive force.

3.4. *Lift capacity for vertical surface climbing*

The lift capacity tests were conducted by adhering the robot on the steel wall. Then, the mass of 1, 2, 3, 4 kg were hinged on the robot as illustrated in Fig. 4. We observed that the robot was able to climb up for maximum of 3 kg weight. For 4 kg weight, the robot tended to slip and was unable to climb up.

3.5. *Vertical movement speed*

Vertical movement speed of robot along the steel wall was tested in 2 directions: upward and downward. Both tests were performed by tuning the DC motor to maximum speed. The travelled time of robot in 1 meter were recorded. Then, the average speed was calculated.

We found that average upward speed was 0.0245 m/s and downward speed was 0.0418 m/s.

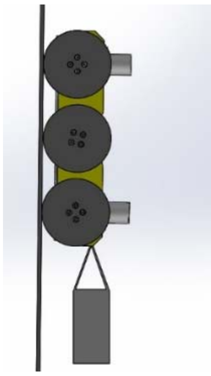


Fig. 4. Experimental setup for vertical lift capacity.

4. Conclusions

In the paper, the magnetic-based climbing robot was designed and constructed. Five aspects of experimentation were tested. Under different circumstances, the robot provides satisfactory results. PID controller has an ability to provide the desirable adhesive force and its adaptive ability. Due to the weight of the robot, eight pieces of neodymium magnet for each wheel are suitable for the robot which can overcome the minimum required magnetic force. The robot is able to climb up with the vertical load up to 3kg. It was clear that the light weight electromagnetic force adjustment mechanism could provide the flexibility to regulate the adhesive force for the magnetic robot while traveling on the ferromagnetic curvature surface.

References

1. L. P. Kalra and J. Gu, An Autonomous Self Contained Wall Climbing Robot for Non-Destructive Inspection of Above- Ground Storage Tanks, *Industrial Robot: An International Journal*, Vol. 34, No. 2, pp. 122–127, 2007.
2. J. Shang, B. Bridge, T. Sattar, S. Mondal and A. Brenner, Development of a Climbing Robot for Inspection of Long Weld Lines, *Industrial Robot: An International Journal*, Vol. 35, No. 3, pp. 217–223, 2008.
3. W. Fischer, F. Tache and R. Siegwart, Magnetic Wall Climbing Robot for Thin Surfaces with Specific Obstacles, *Field and Service Robotics*, Vol. 42, pp. 551–561, 2008.
4. Z. Xu and P. Ma, A Wall-Climbing Robot for Labelling Scale of Oil Tank's Volume, *Robotica*, Vol. 20, No. 2, pp. 209–212, 2002.
5. Y. Zhengyao, G. Yongju, W. Zuwen and W. Xingru Development of a wall climbing robot for ship rust removal, in *Proc. of the International Conference on Mechatronics and Automation (ICMA)* , 2009, pp. 4610-4615.
6. A. Faina, D. Souto, A. Deibe, F. Lopez-Pena, R. J. Duro and X. Fernandez, Development of a Climbing Robot for Grit Blasting Operations in Shipyards, in *Proc. of the 2009 IEEE International Conference on Robotics and Automation*, pp. 200–205, 2009.
7. W. Chuan-sheng, L. Hui-juan and W. Rui-qin, in *Proc. of the 2010 International Conference on E-Product E-Service and E-Entertainment*, 2010, pp. 1-4.
8. M. Wu, G. Pan, T. Zhang, S. Chen, F. Zhuang and Z. Yan-zheng, Design and optimal research of a non-contact adjustable magnetic adhesion mechanism for a wall-climbing welding robot, *International Journal of Advanced Robotic Systems*, vol. 10, 2013.
9. B. Chu, K. Jung, C. S. Han and D. Hong, A Survey of Climbing Robots: Locomotion and Adhesion, *International Journal of Precision Engineering and Manufacturing*, Vol. 11, No. 4, pp. 633-647, 2010.

Intention Classification of a User of a Walking Assist Cart by Using Support Vector Machine

Noritaka Sato, Tomoki Yokotani, Yoshifumi Morita
Department of Electrical and Mechanical Engineering, Nagoya Institute of Technology,
Gokiso-cho, Syowa-ku, Nagoya, Aichi 466-8555, Japan

Abstract

To develop better assist function for a walking assist cart, we focused on the prediction of the intention of a user. As the first step of the research, the forces and torques to the cart from the user's hands, and the rotational velocities of the wheels are sensing. And the Support Vector Machine is used for intention classification. As a result, we confirmed that our method was able to predict the intention of the user with enough accuracy.

Keywords: Walking assist cart, Support Vector Machine, Intention classification, Artificial Intelligence

1. Introduction

With the improvement of medical technology and living environment, the ratio of elderly people is increasing [1]. Elderly people have weak legs, so they take less walks and shopping, and tend to stay home. This makes walking more difficult. Moreover, the number of elderly people living alone is increasing, and many elderly people cannot get help from others. It is essential for the elderly to be able to live independently and comfortably.

One way to solve this problem is to use a walking assist cart. To improve the walking assist cart, some researchers and some companies add an electric assist function to the cart [2][3]. This research started as a joint research with SHINTEC HOZUMI Co., Ltd. Fig. 1 shows the current walking assist cart of SHINTEC HOZUMI Co., Ltd [4]. This cart can perform assist control such as deceleration on downhill.

However, there is a problem in the current cart that a sense of incongruity arises because the assist according to the intention of the user cannot be performed. In this research, the final goal is to develop an algorithm that estimates human intention based on sensor information on a cart and enables an assistant to feel comfortable.

In this paper, as the first step of the research, we aim to predict the state of the cart after 200 milliseconds using the wheel speed and the force on the handles of the cart. Specifically, we use a support vector machine (SVM) to create a learning model from data collected in advance and make predictions. We intend to achieve the final goal by increasing the number of motions that can be predicted. Note that the 200 milliseconds were determined based on discussions with the joint research partner.

Fig. 1. Walking assist cart of SHINTEC HOZUMI Co., Ltd

Fig. 2. Assist cart used in the data collection experiment.

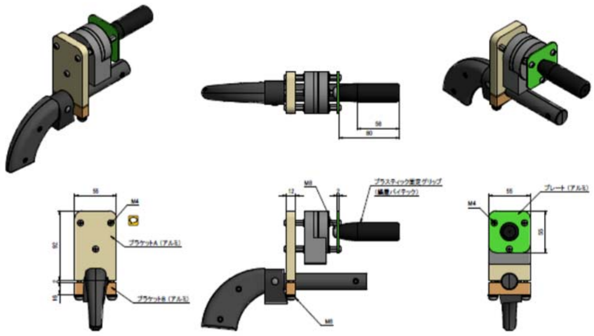


Fig. 3. Attaching method of the force sensor.

2. Data collection

In this study, we performed data collection experiments using real machines for learning SVM. Fig. 2 shows the walking vehicle used in this study. Although this cart has an assist function, the assist function was not used in the data collection experiment.

The motors on the left and right wheels of the cart are equipped with encoders. In this study, the inexpensive force sensor PFS055YA251U6 made by Leptrino was attached to the left and right handles as shown in Fig. 3.

The subject moved the cart forward, backward, and turn as shown in Fig. 4 and Fig. 5, and we recorded the forces and torques on the left and right handles, and the left and right wheel speeds. The turning radius was 0.5 m. Five subjects performed each operation three times.

Fig. 6 shows an example of the acquired data. When the cart goes straight from the stop state (between 0.2 s and

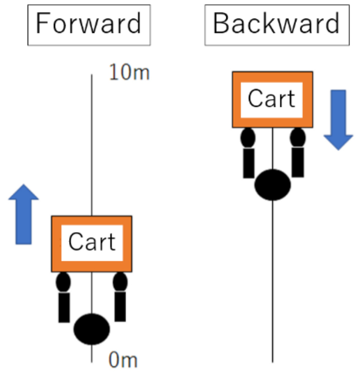


Fig. 4. Forward and backward motion in the data collection experiment.

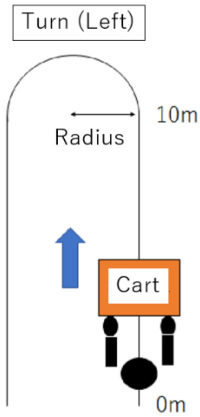


Fig. 5. Turn motion in the data collection experiment.

0.3 s), the value of the force sensor changes before the cart starts moving. Therefore, we think that the state of the cart can be estimated by using the information of the current wheel speed and the force sensor.

3. Proposed method

In this study, the state of the cart 200ms ahead (stop, forward, reverse, right turn, or left turn) will be estimated using SVM. Therefore, the state of the walking car at each time was determined from the acquired data and the video taken during the data collection experiment. After that, the state 200 milliseconds ahead of each time was determined as the correct state for estimation by using SVM.

Fig. 7 shows a scatter diagram of the force in the front-rear direction on the left and right handles after the noise is removed by the low-pass filter. Here, 0 and -1 mean

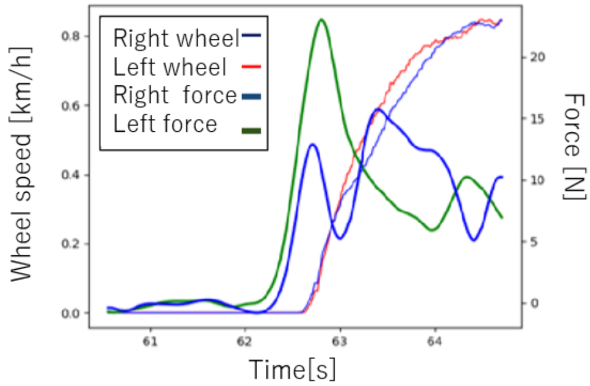


Fig. 6. Example of acquired data.

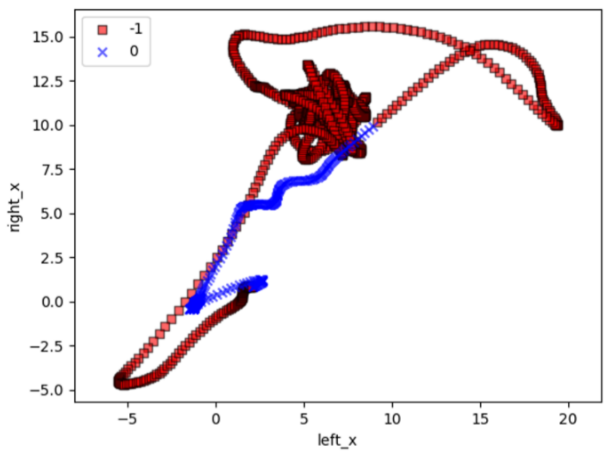


Fig. 7. scatter diagram of the force in the front-rear direction on the left and right handles.

stop and go straight after 200 milliseconds, respectively. It was determined that the point in a certain state was within a certain range, and that it could be estimated 200 ms ahead by clustering with SVM. In Fig. 7, a graph was drawn using two representative values, but in actuality, estimation by using SVM was performed using 14 data, including the values of the left and right 6-axis force sensors and the left and right wheel speeds, as inputs.

"scikit-learn", which is the Python library is used for the calculation of the SVM in this study.

4. Result

Cross-validation was performed on the data divided into 15 pieces. Fig. 8 shows a confusion matrix classified by SVM. The numbers on the label are -1 for forward, 0 for stop, 1 for left turn, 2 for right turn, and 3 for backward.

True value \ Estimated value	−1	0	1	2	3
−1	80323	771	72	34	0
0	284	71919	0	77	130
1	407	0	4453	0	0
2	827	175	0	9290	59
3	0	180	0	0	20544

Fig. 7. Confusion matrix classified by SVM.

The numbers in the matrix represent how many points were output as which labels.

Using this value, the correct answer rate for each cart condition is calculated as 98.9% for forward, 99.3% for stop, 91.6% for left turn, 89.7% for right turn, and 99.1% for backward. The overall correct answer rate is 98.4%. From the above, it is understood that the state of the cart 200 ms ahead can be accurately predicted by the proposed method.

5. Conclusion

The final goal of this study was to develop an algorithm that predicts human intentions and assists a cart that makes people feel more comfortable. The movements were classified by SVM based on wheel speed and force / torque on the handle. We created an SVM program using Python library named scikit-learn and applied it to the collected data. The results showed that the correct answer rate of 98.4% was obtained as a whole, and that the state of the cart could be accurately estimated.

In the future, the prediction accuracy at the time of state switching will be improved, and a predictor that can be used in practical use will be developed.

References

1. Cabinet Office. White Paper on Traffic Safety in Japan 2006. 2006, p.11.
2. T. Kurinobu, et al., The effects of electric walking assistance to a walking cart on safety feeling of older adults., Proceedings of the Japanese Society for Cognitive Psychology, 2015, pp.91. (in Japanese)
3. RT-Works Corporation, http://www.rtworks.co.jp/.
4. T. Yokotani et al., Technique of visualizing walking motion : Identification of a part enabling discrimination of motion of a walking assist device user, 2018, 263-5. (in Japanese)

Development of an Environmentally Adaptable Autonomous Mobile Robot

Iwata Naokazu
Graduate School of Engineering, Kyushu Institute of Technology
Kitakyushu, 804-8550, Fukuoka, Japan

Joo Kooi Tan
Faculty of Engineering, Kyushu Institute of Technology
Kitakyushu, 804-8550, Fukuoka, Japan
E-mail: etheltan@cntl.kyutech.ac.jp

Abstract

To realize a service-robot system for supporting a human life, we propose techniques for estimating self-position of a robot using local features of images, segmenting an image for finding a movable area, and also planning a route for finding the destination. In the route planning, junctions are labelled numbers for the robot to travel in order. The developed mobile robot travels to the destination employing the information on the estimated road region, its self-position and the planned route.

Keywords: Autonomous Mobile Robot, Oriented FAST and Rotated BRIEF, Bag of Features, K-means++

1. Introduction

In recent years, robot technology has been widely used not only in industrial robots, but also as service robots for supporting a human life. One of the main backgrounds is the technological innovation of both hardware and software such as sensor technology, artificial intelligence technology, and data processing technology which are necessary for robots. In particular, much attention has been focused on service robots for supporting a human life. These kinds of robots are expected to be used in a daily life, and necessary to adapt flexibly to various usage, scene, or even to environmental change. This paper proposes an autonomous mobile robot system that adapts to various environments and travel routes.

Robot self-localization methods of previous work include using sensors such as GPS, LiDAR[1,2], markers[3], and template matching[4]. However, these methods have the problem that they depend on the sensor characteristics, limit in setting the location of markers, and affected by the weather.

In this paper, we propose a self-localization method using local features based on a RGB-D camera. Since the developed robot needs to move autonomously in various environments, Oriented FAST and Rotated BRIEF (ORB)[5] are used for local features. The ORB procedure is high speed without affected by scaling, rotation and illumination. In order to reduce the local features calculation cost, we use Bag of Features (BoF)[6] that represents an image by a feature vector However, for developing an autonomous mobile robot, it is necessary to find an area where the robot can travel. In this study, we define that an initial area is a movable road region of the robot. We segment the image into areas and find a road region/area which is a similar area as the initial robot area. For travelling autonomously to a destination, the robot must plan a route from the current location to the destination. In this study, for the route planning, junctions, corners and a destination are labelled numbers for the robot to travel in order. The developed mobile robot travels to the destination employing the information on the estimated road region, its self-position and the planned route.

2. Self-Position Estimation of a Robot

In this section we describe a robot's self-position estimation method based on local features. Since many local features are extracted from a single image, comparison and verification of each local feature, furthermore estimating the position based on the results of similarity or coincidence of each local feature increases the calculation cost. Therefore, in the proposed method, local features are extracted from the image, and the obtained local features are converted to a histogram representing the frequency of appearance of local features using Bag of Features. The robot's self-position is estimated from the similarity of the histogram.

2.1. *Bag of Features*

Bag of Features (BoF) represents one image as one feature vector. The process of BoF is shown as follows.

(i) Extracting local features from images.
(ii) Creating Visual Words by clustering local features.
(iii) Creating a histogram of image based on Visual Words.

(i) *Extracting local features*

We extract local features using Oriented FAST and Rotated BRIEF (ORB). **Fig. 1** shows an example of feature points extracted by ORB. The points surrounded by red circles in Fig. 1 are extracted feature points. Here, let us suppose that $\boldsymbol{F}_j (j = 1,2, \ldots, m)$ local feature points are extracted from an image.

(ii) *Clustering the local features*

We apply K-means++[7] to cluster the local features $\boldsymbol{F}_j$. The K-means++ is an improved version of the K-means method[8] It can cluster the local features independently on the initial cluster allocation.

(iii) *Creating a histogram*

After having clustered the local features, the center of each cluster $i (i = 1, \ldots, n)$ is defined as Visual Words (VWs) $\boldsymbol{V}_i (i = 1,2, \ldots, n)$ The nearest VW is searched from each local feature $\boldsymbol{F}_j (j = 1,2, \ldots, m)$, and a histogram is created by voting to the VWs by considering a score . The score is obtained by the following equation to determine which VWs is closer to each local feature.

$$s = 1 - \frac{1}{2}\left\| \frac{\boldsymbol{V}_i}{\|\boldsymbol{V}_i\|} - \frac{\boldsymbol{F}_j}{\|\boldsymbol{F}_j\|} \right\| \quad (1)$$

Here, $\boldsymbol{V}_i (i = 1,2, \ldots, n)$ is the VWs obtained by clustering and $\boldsymbol{F}_j$ is the local feature of the image obtained by ORB. By voting $\boldsymbol{F}_j$s to the VWs with the smallest score s, a histogram representing the frequency of the appearance of local features is created. The histogram is a feature vector that represents an image.

Fig. 1 Example of extracted feature points from ORB

2.2. *Self-position Estimation of the Robot*

In order to estimate the robot's self-position, the histograms of local features are created using the BoF described in Section 2.1 from an input image and key frame images such as junctions. Next, the similarity is calculated between the input image and the key frame images using the histograms. The robot's self-position is estimated based on the similarity. The procedure of the robot's self-position estimation is shown below.

(i) The appearance frequency histograms H_{in} and H_{key} of local features are created by BoF from the input image and the key frame image, respectively.
The similarity (S_{SAD}) is calculated between H_{in} and H_{key}. It using Sum of Absolute Difference (SAD) defined by

$$S_{SAD} = \sum_{n=0}^{N-1} \left| H_{in} - H_{key} \right|_n \quad (2)$$

where N is the number of classes of the VWs.

(ii) To make the self-position estimation more accurate, the sum of the similarities of the past M input images is used. If it is less than or equal to a threshold T_s, it is judged that the robot is at the position where the key frame indicates.

3. Road Region Estimation

Road region is estimated using Graph Based Segmentation[10], which is one of the image segmentation methods.

3.1. *Graph Based Segmentation*

Graph Based Segmentation (GBS) is one of the renowned image segmentation methods, which combines pixels with similar characteristic pixel values into multiple regions. The algorithm of GBS is shown below.

(i) Smoothing the input image $I(x, y)$ using the following equation.

$$L(x, y, \sigma) = G(x, y, \sigma) \cdot I(x, y) \quad (3)$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma} \exp\left(-\frac{x^2 + y^2}{2}\right) \qquad (4)$$

Here σ is the standard deviation, $G(x, y, \sigma)$ is the Gaussian function, and $L(x, y, \sigma)$ is a smoothed image.

(ii) Creating a graph with the node v_i representing each pixel of the smoothed image and the edge e_q connecting adjacent pixels in the image. Using the illuminance difference between the pixels connected by an edge, the weight of the edge $w(e_q)$ is calculated by the following equation.

$$\omega(e_q) = \sqrt{(R_i - R_j)^2 + (G_i - G_j)^2 + (B_i - B_j)^2} \qquad (5)$$

where, $R_\#$, $G_\#$, and $B_\#$ are the red, green and blue values of the node.

(iii) Assignment of a separate region $S_q (q = 1, \ldots, m)$ to each node where m is the total number of pixels in the image.

(iv) When nodes v_i and v_j connected by edge $e_q (q = 1, \ldots m)$ satisfy the following equation, the region S_i and S_j are combined.

$$\omega(e_q) \leq \min\left(\max_{e \in E_i}\{\omega(e)\} + \frac{c}{|S_i|}, \max_{e \in E_j}\{\omega(e)\} + \frac{c}{|S_j|}\right) \qquad (6)$$

Here, c is a predetermined fixed value, $|S_\#|$ is the number of nodes constituting the region $S_\#$, and $E_\#$ is a set of edges connecting nodes in the region $S_\#$.

For all edges, the node is divided into several regions by determining the region connection by Equation (6). If, with a certain region, the number of nodes is less than or equal to a threshold S_{th}, the region is merged with an adjacent region with which an edge connecting the two regions has the smallest weight among adjacent regions.

3.2. *Road Region Estimation*

The input image is segmented by GBS described in Section 3.1. In the proposed method, the rectangle region surrounded by a red frame as shown in **Fig. 2** is defined as the robot's foot candidate region. The maximum region which contains the foot candidate region is regarded as the road region. If the area of the provisionally estimated road region A_R is smaller than a threshold A_{th}, the latest road region in the past frames is chosen as the current

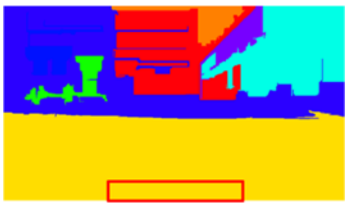


Fig. 3 Example of the robot's foot region

Fig. 2 Example of estimation of a road region

road region. **Fig. 3** shows an example of the road region. The gray area is the estimated road region.

4. Algorithm on Autonomous Movement

First, the robot estimates its own position using local features obtained from input images. Second, the road region is estimated from the input image by GBS. Third, it is examined whether the estimated robot's self-position is at a key frame position in the database or not. If yes, the key frame is examined whether it is the destination or not. If yes, the robot stops. If the key frame is not the destination, the robot's direction of travel is changed to the direction determined in the route setting. If the robot's self-position is not at a key frame position in the database, the estimated road region is divided into three sub-regions in the lower part of the image. We calculate the area of each sub-region. If the area is small, it is judged that there are few areas to advance ahead of the robot, which is dangerous, and the robot stops temporarily. If the areas of the sub-regions are larger than a certain threshold, the robot moves in the direction of the sub- region having the maximum area by finely adjusting its direction.

5. Experiment

5.1. *Experimental Environment*

The robot moves autonomously in an outdoor environment. The robot travels along two types of travel routes, route A and route B, 5 times each. The performance of the robot is evaluated by if it has reached the destination successfully. If an emergency stop is necessary due to the passage of a vehicle during the run, the experiment is interrupted and is resumed after the vehicle has passed. **Table 1** shows the parameters related to this experiment. The forward speed of the robot is 0.2 [m/s] and the rotational speed is 0.2 [rad/s].

5.2. *Experimental Results*

Table 2 shows the results of the experiment. In Table 2, '〇' indicates that the robot reached the destination successfully, and '×' indicates failure. The processing time was 65.4 [ms/frame]. The travel of the robot was successful 3 times

Table. 1 Parameters used in the experiment

No. dimensions of ORB	32
No. features	1000
No. clusters	500
No. learning images	1

Table. 2 Experimental results

	1	2	3	4	5
route A	○			○	○
route B		○	○	○	○

out of 5 with route A, whereas it was successful 4 times out of 5 with route B. As for the weather, the success cases were 2 times out of 4 when it was fine, whereas they were 5 times out of 6 when cloudy.

6. Discussion

In the performed experiment, as shown in Table 2, the destination could be reached in many cases regardless of the travel route and the weather. This is thought to be due to the high accuracy of self-position estimation and road region estimation. However, in the third experiment on route A, the destination could not be reached. This is because the estimation of the road region was incomplete. When the shadow is in the road region, the part that is originally the same region is divided into some regions like the situation in the white circles as shown in **Fig. 4**. Moreover, in this study, the road region was determined using the maximum area's region among the divided sub-regions, so the estimation of the road region was incomplete. More robust road region estimation may be possible by using an algorithm that does not depend on color and includes a process of removing shadows.

7. Conclusion

In this paper, we proposed an environmentally adaptable autonomous mobile robot system. In this system, the robot's self-position estimation based on the local feature of the image and the road region estimation using the image segmentation were performed. In the experiment on autonomous movement, it was possible to reach the destination 3 times out of 5 times with the travel route A and 4 times out of 5 times with the travel route B. By these results, the effectiveness of the proposed autonomous mobile robot system was confirmed. In addition, it was possible to reach the destination 2 times out of 4 times when it was fine, and 5 times out of 6 times when it was cloudy. This also confirms the effectiveness of the proposed system for various weather conditions.

Fig. 4 Example of a shaded road region

In the system, only local features of input images are used for robot's self-position estimation. However, it is considered that self-position estimation can be performed with high accuracy by using distance information in addition to local features. Moreover, because the road region was estimated by color, there were cases where the estimation of the road region was incomplete. Therefore, it is considered that the road region can be estimated in various environments by devising a method that does not depend on color.

8. References

1. Masatoshi Shinohara, S. A. Rahok, Kazumichi Inoue, Koichi Ozaki : "Development of Localization Method Using Magnetic Sensor and LIDAR", *Transactions of the Society of Instrument and Control Engineers*, Vol.49, No.8, 795-801, 2013.
2. Hisashi Date, Shinya Ohkawa, Yoshihiro Takita, Jun Kikuchi : "High Precision Localization of Mobile Robot Using KIDAR Intensity of Surface", *The Japan Society of Mechanical Engineers*, Vol.79, No.806, pp.100-109, 2013.
3. J. Wendlerr, S. Bruggert, H. Burkhard, H. Myritz, "Fault-tolerant self localization by case-based reasoning", *RoboCup 2000: Robot Soccer World Cup IV*, pp,259-268, 2001.
4. Kosei Demura, Yu Nakagawa : "A Monte Carlo Localization based on Template Matching using an Omni-directional Camera", Journal of the Robotics Society of Japan, Vol.27, No.2, pp.249-257, 2009.
5. E. Rublee, V. Rabaud, K. Konolige, G. R. Bradski : "ORB: An efficient alternative to SIFT or SURF", *Proceedings of IEEE International Conference on Computer Vision*, pp.2564-2571, 2011.
6. G. Csurka : "Visual categorization with bags of keypoints", *Proceedings of ECCV Learning in Computer Vision*, pp.1-22, 2004.
7. D. Arthur, S. Vassilivitskii : "k-means++: The advantage of careful seeding", *Proceedings of the 18 Annual ACM-SIAM Symposium on Discrete Algorithm*, pp.1027-1035, 2007.
8. P. Agarwak, N. Mustafa, "k-Means Projective Clustering", *PODS '04 Proceedings of the twenty-third ACM DIGMOD-SIGACT-SIGART symposium on Principles of the database systems*, pp.155-165, 2004
9. P. Felzenswalb, D. Huttenlocher : "Efficient graph-based image segmentation", *International Journal of Computer Vision*, Vol.59, pp.167-181, 2004.
10. M. Calonder, V. Lepetit, C. Strecha, P. Fua, "BRIEF: Binary Robust Independent Elementary Features", *11^{th} European Conference on Computer Vision*, pp.1-14, 2010.

Automated Task and Path Management for Industrial AGVs in Foam Manufacturing Plant

Amornphun Phunopas, Wisanu Jitviriya and, Noppadol Pudchuen
*Faculty of Engineering, King Mongkut's University of Technology North Bangkok, 1518 Pracharat 1 Road, Bangsue, Bangkok 10800, Thailand**

Songklod Tunsiri
Urban Community Development College, Navamindradhiraj University, Wachira Phayaban, Khao Road, Dusit, Bangkok 10300, Thailand

Eiji Hayashi
Computer Science and Systems Engineering, Kyushu Institute of Technology, 680-4 Kawazu Iizuka, Fukuoka 820-0067, Japan
E-mail: amornphun.p@eng.kmutnb.ac.th, wisanu.j@eng.kmutnb.ac.th
https://iraps.eng.kmutnb.ac.th

Abstract

AGVs are increasingly used in the automated warehouse with a high demand for changing traditional workflow management to industrial 4.0. The heart of the computerized system is the central software that can distribute work functions from the queues and manage the AGVs' traffic. On the 2D floor plant layout, the girds are initially from marked points or the place that AGVs have to transit to do an assigned task. This research proposes autonomously generating paths via four nearest grids and path switching scenarios. The results show the generated paths with sequential tasks concurrently in random conditions. The task management method can prevent the AGVs' crash and bottleneck from the operation of nine machines in the foam manufacturing plant.

Keywords: automated tasking, AGVs traffic and queue management

1. Introduction

Nowadays, automated warehouses are widely used in the industry. One example is the IPIN competition. It challenges the problem of position-tracking in the indoor environment, which is the initial scenario that industry needs. The industrial mobile robots or AGVs have to track the magnetic path in a closed route under the defined trajectory [1]. The system consists of many modules of software and hardware, for example, the central computing unit, the fixed automation unit, and the movable agent. These AGVs are one part of the system that complies with the assigned tasks sequentially. They can move to the target and do an assigned task; when it finishes a task, it then moves to the next position according to the production process. Moreover, this implemented technology is suitable for a different industry that has to compromise with the investor on factors of break-even point and payback period. The central software is the key to success in optimized productivity. This paper proposes a scenario that progresses previous work in using AGVs in foam a manufacturing plant [2]. The central software manages tasks in the queue and assigns them to two AGVs, then tests this in the real plant.

2. System Overview

The plant layout is created from part of the foam manufacturing process. The foam product must be transferred from the front machine to the hot temperature room to drain the humidity—the AGVs work in zone 1, zone 2, and zone 3. Completing the primary task takes a full cart from zone 2 to zone 3; then, the secondary task takes an empty cart from zone 1 to replace the taken full cart in the front machine. The AGVs can travel via all grids, which are connected as a path. It takes a series of grids from the current position to any destination in the plant layout in Fig. 1. The path is generated from the central software by traffic management to make the AGV pause or generate a new path that can avoid the collision to another AGV.

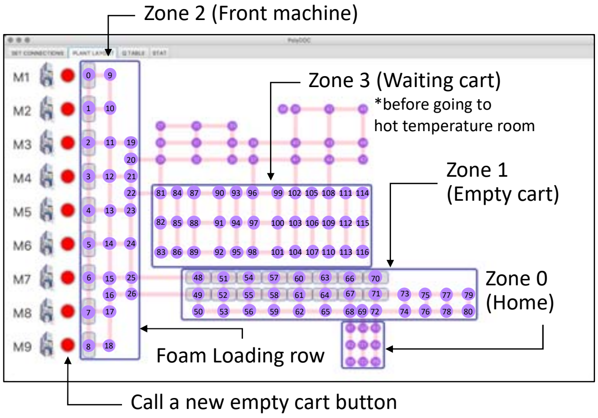


Fig. 1. Plant layout with four nearest grids and guided lines.

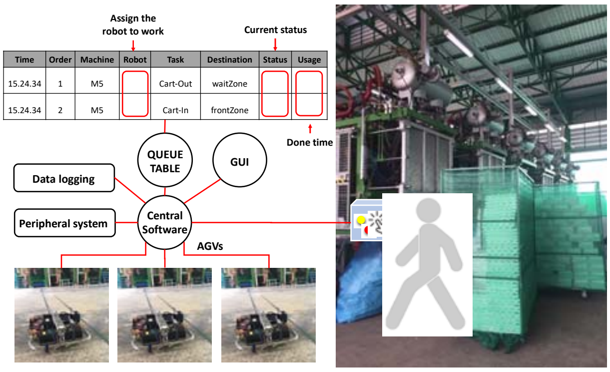


Fig. 2. The worker loads the foam product to the empty cart. When the cart is full, the worker will press the button. The central software pools the information and manages the tasks.

The system consists of the central software, pushbuttons, and AGVs. The operation starts from the worker pressing the button in front of nine machines, randomly depending on putting the foam to the cart finish in Fig. 2. Then, the central software will add the number of the front machine into the queue table. If there is an AGV available, it will be assigned to the queue and instantly start doing the assigned task. Thus, there are two sequential queues in the queue table meant to complete the cart-out and cart-in task.

3. Automated Task

The central software will provide the job that the AGVs have to do. The concept is that an AGV goes, arrives at a target, and performs a job until the last subtask. The AGV always communicates with the central software and is then assigned to the next target and do the next job. After finishing the task, the AGV is free and prompt to receive a command from the central software in Fig. 3.

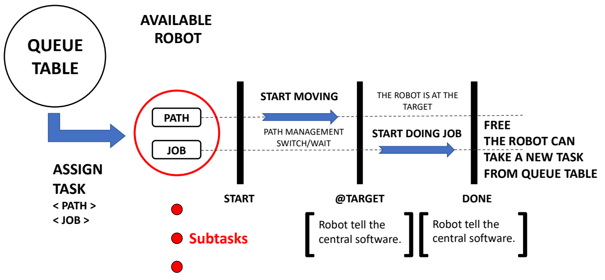
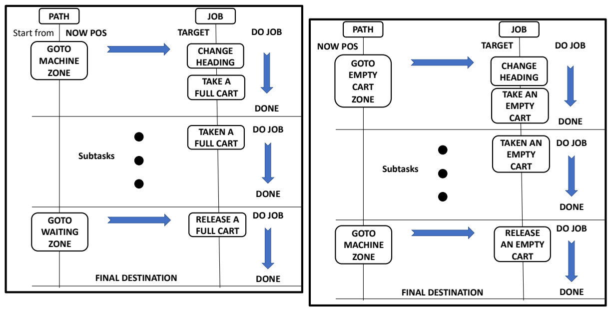


Fig. 3. To assign the available AGV to the queue and update the table queue.

The subtasks that the AGV can do are a cart's heading adjustment, an AGV's heading adjustment, cart hooking, and cart releasing in Fig. 4. The subtask concept is also



flexible for various applications in other industries.

Fig. 4. (Left) Task of taking the full cart to the waiting zone (Zone 3). (Right) Task of taking the empty cart to the machine zone (Zone 2).

4. Path Management

The grids are marked on the layout for significant tasks that the AGV can move past or stop to do a task. The grid positions are not symmetrical or balanced in rows and columns in this plant layout. The central software manages the whole system operation and traffic for the AGVs.

4.1. *Grids Connection to Guided Line*

The guided lines are designed to cover the work area in the plant layout that the grids can be connected to, creating a path for AGVs. The AGV has tasks assigned differently, but it has the same design and capability. The guided line uses magnetic tape that has to be constructed under the plant's floor because of durability and long-term operation. The AGV uses the magnetic sensor for line detecting and tracking. The grids of the main lines are mapped to the plant layout, as seen in Table 1. The grid can be detected by the cross junction of magnetic tape and RFID.

Table 1. Main line for AGVs

Line	Grid sets of main lines
0	{0,1,2,3,4,5,6,7,8}
1	{9,10,11,12,13,14,15,16,17,18}
2	{9,10,11,19,20,21,22,23,24,25,26,16,17,18}
3	{6,15,25,48,51,54,57,60,63,66,70,73,75,77,79}
4	{6,15,16,26,49,52,55,58,61,64,67,71,73,75,77,79}
5	{7,17,16,26,49,50,53,56,59,62,65,68,69,72,74,76,78,80}
6	{4,13,22,81,84,87,90,93,96,99,102,105,108,111,114}
7	{4,13,22,81,82,85,88,91,94,97,100,103,106,109,112,115}
8	{4,13,22,81,82,83,86,89,92,95,98,101,104,107,110,113,116}

4.2. *Path Switching*

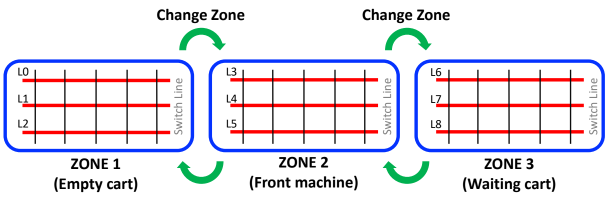
The AGV travels along the defined main lines. If the line is not available or is occupied by others, the AGV has to pause or switch being in the current mainline to prevent collision damage, as shown in Table 2. When the system has many AGVs, the central software has all of the information to analyze and assign a proper path for each AGV.

Table 2. Switching lines across the main lines

Zone	Grid sets of switching lines
1	{{48,49,50},{51,52,53},{54,55,56},{57,58,59}, {60,61,62},{63,64,65},{66,67,68,69},{70,71,72}, {73,74},{75,76},{77,78},{79,80}}
2	{{0,9},{1,10},{2,11,19,20},{3,12,21,22},{4,13,23}, {5,14,24},{6,15,25},{16,26},{7,17},{8,18}}
3	{{81,82,83},{84,85,86},{87,88,89},{90,91,92}, {93,94,95},{96,97,98},{99,100,101},{102,103,104}, {105,106,107},{108,109,110},{111,112,113}, {114,115,116}}

The path-switching method helps to manage the traffic of the multiple AGVs' operation under the limited guided lines. It is similar to railroad switching, but Fig. 5 indicates that it has many possibilities for switching the main lines via the cross lines.

Fig. 5. Main lines in zone 1, zone 2, and zone 3, including switching lines in each zone.



4.3. *Traffic Manager*

In Fig. 6, the central software has perceived the intersection points of the AVGs' paths. It monitors the critical distances by extending each AGV's path. However, it takes time to do the job at a target. The central software also has to calculate the time that corresponds to doing a job and traveling. Besides the collision problem, bottleneck conditions can occur when AGVs have to wait for too long for the current path to clear. This might result in a traffic problem that is inefficient for productivity, even if the AGVs can get to the destination.

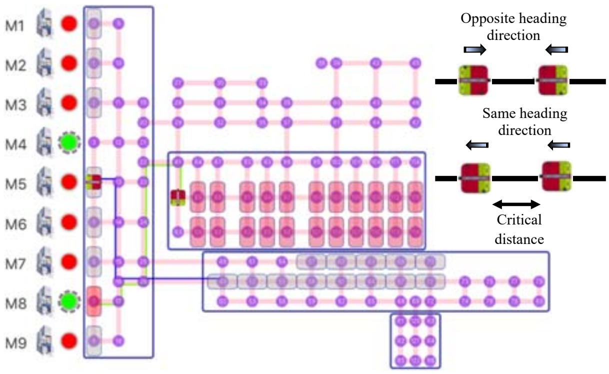
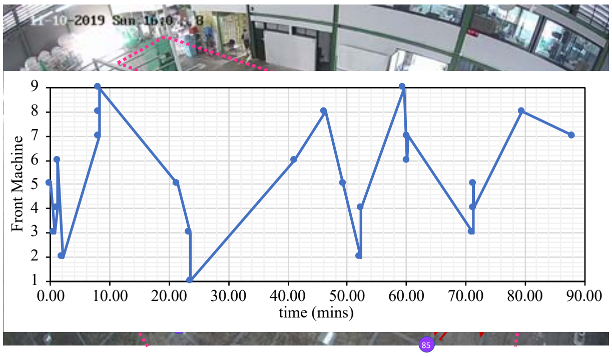


Fig. 6. Collision monitoring using all paths of AGVs in the central software

5. Results

The experiment took one hour and 32.54 minutes (from 3:24:34 p.m. to 4:57:28 p.m.) to complete 24 instances of transportation by cart. There are work areas 7 m^2, 57 m^2, and 72 m^2 for zone 1, zone 2, and zone 3, respectively. The priority task started from zone 2 to zone 3 in order from grid 116 to 82. The pattern to transfer the empty carts from zone 1 to zone 2 had six grids, from grid 48 to 55, which ran in a loop repeatedly, as in Fig. 7. The buttons were independently pressed by the workers and then appended to the queue, as in Fig. 8.

Fig. 7. The travel pattern via grids in the plant layout for carts' transportation



It took every 3.40 minutes to press the button once of totally one hour and 28.15 minutes (or 88.15 minutes).

Fig. 8. The nine buttons in front of the machine are pressed independently 24 times

There were two AGVs separately operating in the experiment. One AGV had the full-cart-out task, and another AGV had the empty-cart-in task. It took 3.52 minutes, on average, for one cart to finish taking out the full cart and taking in the empty cart to the front machine. The AGVs had the same design, traveled by the same speed (0.5 meters per second), and had the same setting and behaviors. They worked overlap in zone 2, which made it possible for the AGVs to crash and bottleneck. As long as there is a task in the queue, the AGVs do not stop working but continue running to the next queue. The completion time of the full-cart-out task and the empty-cart-in task were approximately in the linear. Meanwhile, the time to pause and do a job were the interferences, as in Fig. 9. The workers take time to load the foam product to the cart, about 30–60 minutes, depending on the size of the foam product. The workers then call the AGV to take the full cart out and replace it with the new empty cart. Fig. 10 shows the time that the worker at the front machine has to wait for the AGV to start taking the full cart out. The problem that can occur is when the workers call the AGV at the same time. Then the AGV has to complete the tasks in the queue one by one. The tasks that are not assigned to the AGVs have to wait up to 20 minutes, which is a very long time.

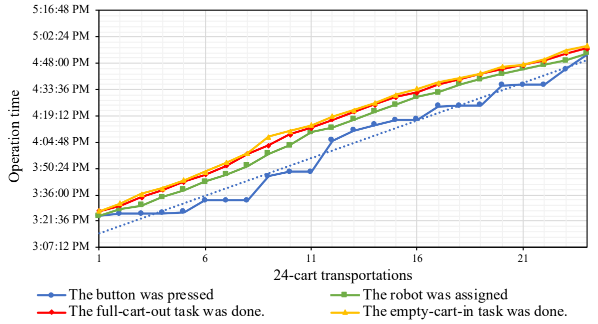


Fig. 9. All complete tasks to place 24 full carts in zone 3 and 24 empty carts in zone 2.

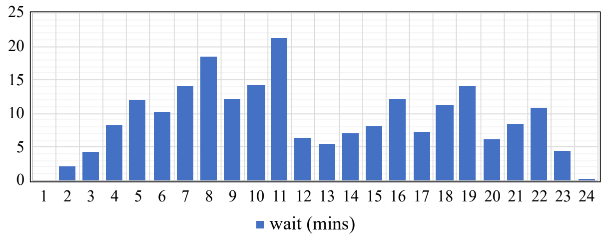


Fig. 10. The AGV waited to start working and take the full cart out until after the worker pressed a button on the front of the machine.

6. Conclusion

On average, the transportation of one cart takes 3.52 minutes to finally replace a new empty cart and ready for a worker to load the foam product to the cart. However, the system needs to add more AGVs to decrease the waiting time in the queue. To increase work efficiency, the speed of the AGVs should be adjustable when they do not hook the cart; this can speed up as well as slow down the subtasks.

Acknowledgements

This work is supported by the Thai Research Fund (TRF) and Polyfoam Manufacturing Co., LTD under the contract no. RDG6050047.

References

1. Pérez-Rubio, M.C., Losada-Gutiérrez, C., Espinosa, F., Macias-Guarasa, J., Tiemann, J., Eckermann, F., Wietfeld, C., Katkov, M., Huba, S., Ureña, J. and Villadangos, J.M., 2019. A realistic evaluation of indoor robot position tracking systems: The IPIN 2016 competition experience. Measurement, 135, pp.151-162.
2. Amornphun Phunopas, Wisanu Jitviriya, Noppadol Pudchuen, Sutee Kumjaikong, Songklod Tunsiri and, Eiji Hayashi. Automated Guided Vehicle System Analysis in Foam Manufacturing Plant Using Petri-net, The International Conference on Artificial Life and Robotics, Jan. 10-13.

Object Tracking Method Considering Time Series Information Using Re3 with Stochastic Depth

Taichi Kitayama, Hyoungseop Kim
Department of Mechanical and Control Engineering, Kyushu Institute of Technology, 1-1, Sensui-cho, Tobata-ku, Kitakyushu-shi, Fukuoka, JAPAN

Abstract

In recent years, there has been an increasing demand for automation of cargo handling operations in harbors. However, its automation has not been realized in Japan since there are many fundamental technologies to be solved until now. In this study, we propose a tracking method of the container gripping area based on Re3 for the purpose of automation of cargo handling work. Re3, an object tracking method used in the conventional method, has a problem that global features cannot be extracted well. In order to solve this problem, our method incorporates a model called Stochastic Depth.

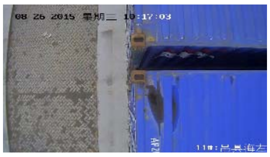
Keywords: Container Terminal, Gantry Crane, Convolutional Neural Network, Real-Time Recurrent Regression Networks for Visual Tracking of Generic Objects, Stochastic Depth.

1. Introduction

In recent years, container transportation has played an important role in international cargo transportation. The container cargo movement in the world has tended to increase over the years. Container movement in the world will reach 146 million TEU in 2018 [1]. Here, one 20-foot container is called 1 TEU. This is used to indicate the loading capacity of the container ship and the cargo handling quantity at the port terminal. From this kind of circumstance, there is a need to improve the efficiency of cargo handling operations in container yards around the world. This is because, in addition to an increase in the amount of cargo movement, the working environment is improved and safety is ensured. Along with the evolution of IT technology, de facto standardization and de facto stepping of cargo handling work automation are progressing. As a result, an environment that is easy to introduce into container terminals has been established. For these reasons, competition for container automation is intensifying around the world.

In Japan, there is a high interest in remote operation and automation of container terminals. However, semi-automated container handling has not been achieved in Japan. One of the reasons is that there is very little information and literature on automatic cargo handling. In particular, the required technology for automating container handling has not been established. It is a technology that accurately measures the relative position of the container crane suspension and the container being handled. Among them, there are very few previous studies on container position measurement for movies. There is a method using template matching in previous research [2]. However, the number of experiments is small datasets. Furthermore, there are few types of datasets. Since there are various situations in fact, it seems that generalization is lacking in practical use.

Now a day, image recognition technology has been introduced in various situations. Its application is very wide such as a face recognition system and an obstacle detection in automatic driving. Therefore, this field is used for container position measurement in cargo handling work.



(a)Blue container

(b)Distant container

(c)Green container

(d)Yellow container

Fig.1 Images of containers

Furthermore, this method uses an object tracking method using deep learning. In deep learning, the machine itself captures the characteristics of the tracking target. In this way, we can output effective position information for images under various conditions (Fig.1).

Based on the above, this method proposes a method for measuring the relative position of a hanger and a container using an object tracking method based on deep learning for the purpose of automation of cargo handling work.

2. Methods

Real-time recurrent regression networks for visual tracking of generic objects (Re3) [3], which incorporates Stochastic Depth [4], was used to track the target object from the image. Re3 is an object tracking method using deep learning, and Stochastic Depth is a model that improves the learning efficiency of deep learning.

2.1. Real-Time Recurrent Regression Networks for Visual Tracking of Generic Objects

Re3 is a hybrid tracker considering time series information. It performs a learning in advance and learn expressions that capture important features of the tracked object. In addition, a tracking is performed with information on the object being tracked. This makes it possible to adapt to changes during the tracking of objects. This model performs object tracking at 150[fps].

As a network structure, input is performed from two images of the present and one frame past. Prepare two identical models and process their inputs in parallel. The output is concatenated and fully connected to the Long short-term memory [5]. By directly comparing the difference between these two frames, it helps to predict the future. In addition, skip layers are introduced in the first, second and fifth layers. This is to preserve high-resolution spatial information.

2.1.1. Long Short-Term Memory

Re3 has the structure as its internal structure. LSTM is a model that can hold long-term time series data internally. Features include a memory cell and three gates. Memory cells store time-series information. The three gates include an input gate $\mathrm{i_t}$, a forgetting gate $\mathrm{f_t}$, and an output gate $\mathrm{o_t}$. In particular, the input gate and forgetting gate play a major role. The input gate selects the information input to the memory cell. The forgetting gate discards useless information from the state inside the memory cell. This makes it possible to process long-term time-series data. The following formula is used to calculate the output of each gate and the memory cell $\mathrm{c_t}$.

$$i_t = \sigma\left(W_i x_t + R_i h_{t-1} + b_i + p_i \odot c_{t-1}\right) \tag{1}$$

$$f_t = \sigma\left(W_f x_t + R_f h_{t-1} + b_f + p_f \odot c_{t-1}\right) \tag{2}$$

$$o_t = \sigma\left(W_o x_t + R_o h_{t-1} + b_0 + p_o \odot c_{t-1}\right) \tag{3}$$

$$c_t = o_t \otimes z_t + c_{t-1} \otimes f_t \tag{4}$$

$$h_t = o_t \otimes tanh(c_t) \tag{5}$$

Here, W, R, b and p are weights, x_t is the input to LSTM, and h is the final output.

2.2. Stochastic Depth

Stochastic Depth is a structure that improves learning efficiency. Probabilistically changing the number of layers has the effect of improving generalization performance and reducing computation time.

(a)Proposal method (b)Re3

Fig.2 Experimental result(case 1)

(a)Proposal method (b)Re3

Fig.3 Experimental result(case 2)

Dropout reduces the network horizontally. In contrast, Stochastic Depth shrinks the network vertically. To shorten it, set the function layer output in the residual block to 0. The identity map is handled as the output of the residual block. In other words, the identity mapping is performed. Function layer consists of a sequence of layers: Conv-BN-ReLU-Conv-BN, where Conv and BN stand for convolution and batch normalization respectively.

The establishment of whether or not to shorten is called the survival probability. The survival probability is set manually for the input layer and manually for the final layer. The intermediate layer is calculated by linear attenuation. In this method, the survival probability in the final layer was set at 0.9. The formula for survival probability p_l is shown below.

$$p_l = 1 - \frac{l}{L}(1 - p_L) \qquad (6)$$

3. Experimental Result

This method evaluates accuracy with three indicators of intersections over union (IoU) measurement, detection rate, and processing speed. The object to be detected is the corner bracket of the container. The following formula is used to evaluate the IoU and detection rates.

$$IoU = \frac{|A \cap B|}{|A \cup B|} \times 100 \qquad (7)$$

$$r = \frac{n_d}{N} \times 100 \qquad (8)$$

Where r is the undetected rate, n_d is the number of undetected images, and N is the number of test images. The number of learning images is 965. The number of detected images was 345, and the number of test images was 480. In addition to Re3 alone, the container tracking method using single shot multibox detector[6], which is the conventional method [7], is also compared. Table.1 shows the contents of the experimental results. The experimental results are shown in Fig.2 and Fig.3.

4. Discussion

In this study, object tracking was performed using a set generated from container animation. From Table.1, the results were lower than the conventional method for any accuracy index. However, as shown in Fig.2 and Fig.3, there seems to be almost no difference between the two methods. The background is that it is no longer possible to retrace lost objects. Fig.4 shows the output results of tracking again after losing sight. This means that the LSTM has lost its memory function. This is thought to be due to the compatibility of LSTM and Stochastic Depth. Stochastic Depth outputs different structures to the LSTM for each epoch. In other words, the output becomes unstable. By repeating this unstable output, the LSTM has learned to place importance on the latest information so that it can respond flexibly. There was also a problem before losing track. It was observed that the IoU decreased when the tracking target became larger on the screen.

Table.1 Experimental result

	IoU[%]	Undetected rate[%]	Frame rate[fps]
Proposal method	70.73	12.50	87.07
Re3	79.89	3.75	112.89
Single Shot Multibox Detector	70.39	5.86	45.67

(a)Proposal method (b)Re3

Fig.4 Experimental result(case 3)

It can be seen that the ability to extract global features is reduced. The cause is thought to be the increase in the number of layers to benefit from Stochastic Depth. This is because the specific gravity of local features in the output increases. However, the processing speed of this method exceeded 30 [fps]. This can be said to be effective in real-time video processing.

5. Conclusion

In this paper, the object tracking method for the purpose of automation in the container terminal was proposed. The experimental results are as follows: IoU are 70.73% and 12.50% for the undetected rate, which was below the conventional method. Future issues include the introduction of a global feature extraction model that replaces Stochastic Depth and a comparison with other object tracking methods.

References

1. Japanse maritime center, "Sekai no Kontena Niugokiryou no Suii" (Changes in container carriage amount in the world), http://www.jpmac.or.jp/relation/trend_graph/25_1_1.pdf(2019/9/27 access). (in Japanese)
2. S. Kunimitsu et al., ″Kontenakure-n Jidouka no tame no Gazou wo Motiita Kontena Soutaiiti Sokutei″ (Relative position measurement using images for container crane automation), https://www.jstage.jst.go.jp/article/ieejeiss1987/121/5/121_5_882/_pdf (2019/11/20 access). (in Japanese)
3. D. Gordon et al., "Real-Time, Recurrent, Regression Networks for Visual Tracking of Generic Objects", IEEE Robotics and Automation Letters, vol.3, pp.788-795, 2017
4. H. Gao et al., "Deep Networks with Stochastic Depth", European Conference on Computer Vision, pp.646-661, 2016.
5. K. Greff et al., "LSTM: A Search Space Odyssey", IEEE Transactions on Neural Networks and Learning Systems, vol.28, pp.2222-2232, 2016.
6. W.Liu et al., "SSD:Single Shot MultiBox Detector", European Conference on Computer Vision, pp.21-37, 2016.
7. T. Kitayama, et al., "Detection of Grasping Position from Video Images Based on SSD", International Conference on Control, Automation and Systems, pp.1472-1475, 2018.

Neural Network and Internal Resistance based SOH classification for lithium battery

Jong-Hyun Lee
School of Electronics Engineering, Kyungpook National University
80 Daehak-ro, Buk-gu, Daegu, Korea

Hyun-Sil Kim
Naval Combat System PMO Agency For Defense Development
Jinhae P.O. BOX 18, Chanwon, Gyeongnam, Korea

In-Soo Lee
School of Electronics Engineering, Kyungpook National University
80 Daehak-ro, Buk-gu, Daegu, Korea
E-mail: whdugs8428@knu.ac.kr, shyla@add.re.kr, insoolee@knu.ac.kr

Abstract

This paper presents a battery state of health (SOH) monitoring system to diagnose fault in battery using a multilayer neural network state classifier (MNNSC) and an internal resistance state classifier (IRSC). In this system, the MNNSC utilizes discharge voltage data from operating the lithium battery at high temperatures. Whereas, the IRSC uses the open circuit voltage, terminal voltage, and current to calculate the internal resistance. From experimental results, it is noted that the proposed battery SOH monitoring method diagnoses the battery status very well.

Keywords: Lithium battery, State of health, Fault diagnosis system, Multilayer neural network, Internal resistance.

1. Introduction

Lithium batteries are used in applications such as cell phones, electric vehicles, unmanned submarines and energy storage systems (ESS) as the main sources of power.[1][2] However, faults in the battery system can degrade device performance and even cause serious operational faults.[3] Therefore, as the risk of battery system faults have increased in recent years, research on fault diagnosis and safety management of the battery system have become particularly important.[4]

There are various methods to estimate battery SOH. The resistance of the battery equivalent circuit can be evaluated using the Kalman filter method to predict the SOH.[5] However, the Kalman filter method has the disadvantage of being difficult to apply owing to the complexities of the parameters and algorithms. The open-circuit voltage (OCV) method measures the output voltage of the battery under no-load conditions and predicts the SOH according to changes in the OCV.[9] However, it is difficult to apply the OCV method to real-time systems as measurements need to be obtained when the battery is in a state of internal chemical equilibrium. The proposed system uses neural network. The neural network can be learned complex nonlinear input and output relationship.

In this work, we propose a SOH monitoring method for a battery module using multilayer neural network (MNN) and the internal resistance of the battery. The lithium battery characteristics are affected by and the performance changes with temperature.[6] In order to diagnose the battery state, it is operated at a high temperature (40 ℃) for an extended period. Through experiments, the battery faults are diagnosed by the MNNSC and IRSC.

2. Monitoring System Based on a State Classifier by MNNSC and IRSC

2.1. *Lithium battery experiment setup*

The energy storage capacity of a battery is defined as the total amount of charge discharged when a constant current is dissipated from the fully charged state of the battery to the fully discharged state. The fully charged state is defined as the state in which no current flows for charging with a constant voltage. The fully discharged state is defined as the state in which all dischargeable electric charges are dissipated in a range such that the battery is not damaged.

The test method exposes a fully charged battery to a high temperature (40 °C) for 8 hours. The battery is then exposed to room temperature for 3 hours and subsequently completely discharged.

2.2. *Battery SOH algorithm*

The proposed lithium battery SOH monitoring system consists of a battery controller system, a MNNSC, an IRSC, and a coordinator. As shown in Fig. 1. The coordinator finally determines diagnosis. The final diagnosis results as normal, warning and fault using output from the MNNSC and IRSC. The diagnosis rules for the coordinator are shown in Table 1.

2.2.1. *IRSC (Internal resistance state classifier)*

In this work, the internal parameters of the battery are estimated using its equivalent circuit model shown in Fig. 2.[7]

Here, V_{OCV} is the open circuit voltage of the battery, V_t is the battery terminal voltage, V_{R0} and V_{R1} are the voltages across R_0 and R_1.

At the moment when the circuit is closed, V_{R1} becomes 0 by the capacitor C1. The battery state is

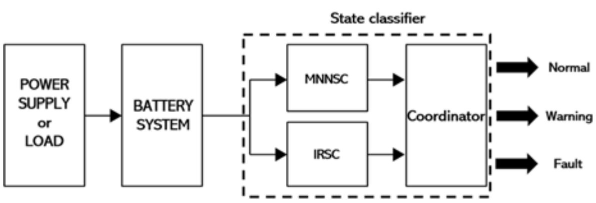


Fig. 1. Configuration of the battery SOH monitoring system

estimated by measuring its internal resistance R_0, which is calculated as follows.

Table 1. Diagnosis rules for classification at the coordinator

MNNSC result	IRSC result	Coordinator result
Normal	Normal	Normal
Normal	Abnormal	Warning
Warning	Normal or Abnormal	Warning
Fault	Normal or Abnormal	Fault

$$R_0 = \frac{V_{OCV} - V_{R0}}{I} \tag{1}$$

The internal resistance is defined as normal for the range of 0.12–0.18 Ω and abnormal for any value exceeding this range.

2.2.2. *MNNSC (Multilayer neural network state classifier)*

The MNNSC used in this study consists of one input layer, two hidden layers, and one output layer. The

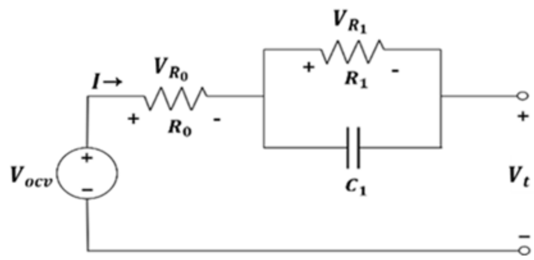


Fig. 2. Equivalent circuit model of the battery

rectified linear unit (ReLU) activation function is used for each hidden layer, and the softmax function is used for the output layer. The ReLU is expressed in Eq. (2).

$$f(x) = \begin{cases} x, & for\ x > 0 \\ 0, & for\ otherwise \end{cases} \tag{2}$$

The RMSprop algorithm does not uniformly add all the past gradients but reflects only the information of the new gradients such that the learning rate does not become 0. The RMSprop algorithm is numerically expressed by Eq. (3) and (4).

$$G_t = \gamma G + (1 - \gamma)(\nabla_\theta f_t(\theta_t))^2 \tag{3}$$

$$\theta_t = \theta_{t-1} - \sqrt{\frac{\eta}{G_{t-\epsilon}}} \cdot \nabla_\theta f_t(\theta_{t-1}) \tag{4}$$

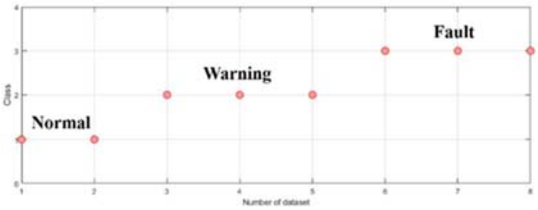


Fig. 5. SOH diagnostic test results for battery operated at high temperature by MNNSC

where γ is the forgetting factor, which is typically 0.9, θ is a network parameter, and $f(\theta)$ is the cost function. In this work, softmax is used. $\nabla_\theta f(\theta)$ is the gradient of the network and t is a time step.

The softmax function is used for the output layer and for transforming the class classification problem. The softmax function is mathematically expressed as Eq. (5).

$$p_i = \frac{\exp(x_i)}{\sum_k \exp(x_k)} \tag{5}$$

where k is the number of dimensions of the input and output vectors, p_i is the i-th output value, and x_i is the i-th input value.

3. Experiment and Results

The experimental setup configuration for conducting the experiment is shown in Fig. 3. Fig. 3-① is an electronic load, and when charging, replace it with a power supply. Fig. 3-② is a battery system and consists of battery, current and voltage sensor and MCU (Micro Controller Unit) to communicate with PC. Fig. 3-③ is PC and serial communication with battery system to receive battery voltage and current data and monitor battery SOH.

The MNNSC consists of two hidden layers and an output layer; the input layer consists of 3600 nodes, and the two hidden layers each consist of 256 nodes. The output layer consists of 3 nodes. The output is 1 when the battery state is normal, 2 when the battery is a warning state, and 3 when the battery is in the fault state.

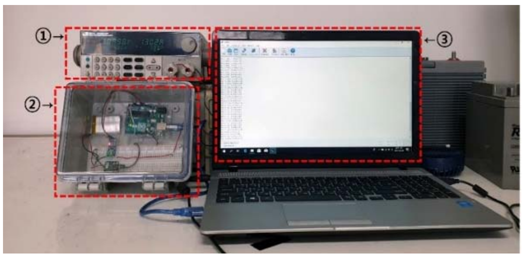
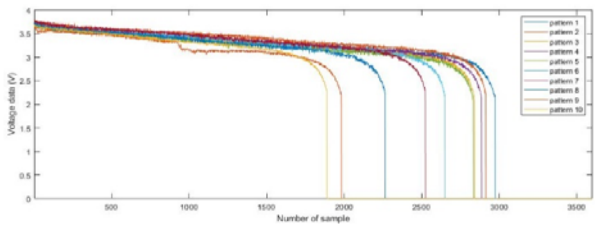


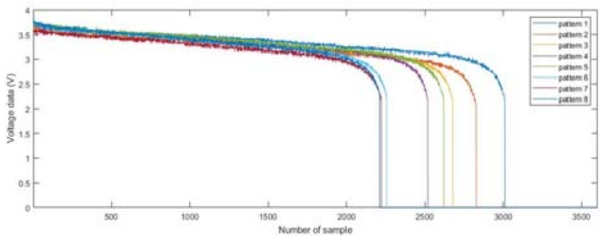
Fig. 3. Experimental environment configuration

In order to evaluate the performance of the proposed system, the discharge data of the lithium battery tested in the high temperature environment were used. The data used in the learning are shown in Fig. 4 (a). When the capacity of the battery reaches 80% of the rated capacity, it is defined as a fault. In this work, we defined 100–90% of the rated capacity as normal, 90–80% of the rated capacity as warning, and less than 80% of the rated capacity as fault. Therefore, patterns 1 to 5 were learned as normal, patterns 6 and 7 as warning, and patterns 8 to 10 as fault states.

Fig. 4 (b) shows the discharge graphs of another battery operated in the same environment as that of the test. In Fig. 4 (b), patterns 1 and 2 depict the normal state with SOH of 90% or more; patterns 3, 4, and 5 depict the



(a) Battery discharge voltage data used in learning



(b) Battery discharge voltage data used in test

Fig. 4. Lithium battery discharge graph for operation at high temperature

warning state with SOH close to 80%; and patterns 6, 7, and 8 are for fault state with SOH less than 80%.

Fig. 5 shows the test results for patterns 1 through 8 of Fig. 4 (b). The proposed system classifies patterns 1 and 2 as normal; patterns 3, 4, and 5 as warning; and patterns 6, 7, and 8 as fault states. The proposed system therefore accurately diagnoses the battery status using its SOH data.

The internal resistances are obtained from IRSC using a shown in table 3.

The final diagnosis is determined by the coordinator in the state classifier. When the result of the MNNSC is normal and that of the IRSC is normal, the output is normal; when the MNNSC result is a warning and IRSC result is normal or abnormal, a warning is generated at the output; when the MNNSC result is a fault and IRSC result is abnormal, the output generated is a fault state. The final diagnosis results from the coordinator are shown in Fig. 6.

4. Conclusion

In this study, a battery module SOH monitoring system was developed and implemented using MNNSC and IRSC. The battery was charged and discharged under a high temperature (40 °C) operating condition, and the fault state of the battery was diagnosed using the proposed MNNSC and IRSC along with the discharge data of the battery. From the experimental results, we demonstrate that the proposed diagnosis method for the lithium battery module has good performance

Table 3. Battery internal resistances used in the tests

Pattern	Internal resistance	Result
1	0.173	Normal
2	0.143	Normal
3	0.208	Abnormal
4	0.157	Normal
5	0.257	Abnormal
6	0.227	Abnormal
7	0.268	Abnormal
8	0.252	Abnormal

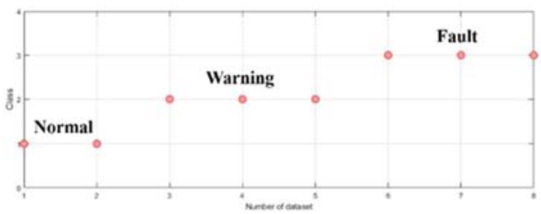


Fig. 6. Final diagnosis results from the coordinator

Acknowledgements

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. 2019R1I1A3A01058319).

This study was also supported by the BK21 Plus project funded by the Ministry of Education, Korea (No. 21A20131600011).

References

1. H. R. Hwang, B. S. Kim, J. H. Kim, R. Xu and I. S. Lee, Development of a fault diagnosis algorithm for Solar Panel, in *Proceedings of IEIE Summer Conference* (2016) pp. 1463–1465.
2. T. H. Cho, H. R. Hwang, J. H. Lee and I. S. Lee, Comparison of Intelligent Methods of SOC Estimation for Battery of Photovoltaic System, *International Journal of Advanced Computer Science and Applications* **9**(9) (2018), pp. 49–56.
3. A. Sidhu, A. Izadian and S. Anwar, Adaptive nonlinear model-based fault diagnosis of Li-ion batteries, *IEEE Transactions on Industrial Electronics* **62**(2) (2014), pp. 1002–1011.
4. Y. Zhao, P. Liu, Z. Wang, L. Zhang and J. Hong, Fault and defect diagnosis of battery for electric vehicles based on big data analysis methods, *Applied Energy* 207 (2017), pp. 354–362.
5. D. Andre, C. Appel, T. Soczka-Guth and D. U. Sauer, Advanced mathematical methods of SOC and SOH estimation for lithium-ion batteries, *Journal of Power Sources* 224 (2013), pp. 20-27.
6. A. Burke and M. Miller, Life cycle testing of lithium batteries for fast charging and second-use applications, *2013 World Electric Vehicle Symposium and Exhibition (EVS27)*. IEEE (2013), pp. 1-10.
7. L. Zhang, H. Peng, Z. Ning, Z. Mu, and C. Sun, Comparative Research on RC Equivalent Circuit Models for Lithium-Ion Batteries of Electric Vehicles, *Appl. Sci.* **7**(10) (2017), 1002.

Estimation of Self-Posture of a Pedestrian Using MY VISION and Deep Learning

Tomoyuki Kurosaki
Graduate school of Engineering, Kyushu Institute of Technology,
1-1 Sensui-cho, Tobata-ku, Kitakyushu-si, Fukuoka 804-8550

Joo Kooi Tan
Faculty of Engineering, Kyushu Institute of Technology,
1-1 Sensui-cho, Tobata-ku, Kitakyushu-si, Fukuoka 804-8550

E-mail: etheltan@cntl.kyutech.ac.jp, kurosaki-tomoyuki@ss10.cntl.kyutech.ac.jp

Abstract

A system is proposed that performs gait analysis of a pedestrian to prevent fall. In the system, a user walks with a chest-mounted camera. His/her walking posture is estimated using a pair of images obtained from the camera. Normally it is difficult to estimate the camera movement, when the parallax of the image pair is small. Therefore, the system uses a convolutional neural network. Optical flow and camera movement, and depth images are estimated alternately. Satisfactory results were obtained experimentally.

Keywords: Posture, Posture analysis, Pedestrian, MY VISION, Deep learning, DeMoN.

1. Introduction

In recent years, an increasing aging rate has become a serious problem worldwide. Due to the fatal fall of elderly people over the age of 65, the elderly need care. However, there are relatively few young people who can care for elderly people and there is a shortage of human resources. Therefore, elderly people need to take measures by themselves to prevent falls. It is then useful to know own walking posture. This paper proposes a system for analyzing one's walking posture from chest-mounted camera images. The goal is to improve the walking posture by feedback of the analyzed results to help prevent falls.

Researches on estimating human walking posture have been done for the purpose of creating animation motion and virtual reality games in many cases. Many studies have proposed the installation of a single camera [1] or multiple cameras [2, 3] around a pedestrian. There are two main methods for estimating a posture of a pedestrian. One is to attach markers to a pedestrian’s joints and estimate their positions by deep learning [4], for example. A study has also been proposed in which a pedestrian wears multiple cameras and estimates the walking posture using the geometric relationship between the cameras [5]. However, they are rather large systems and no one can use it easily.

This paper proposes a method for analyzing one's walking posture from a video provided by a monocular chest-mounted camera, which we call a MY VISION system. DeMoN (Depth and Motion Network) [6] is used to determine the camera motion. The obtained camera motion data is corrected based on a walking motion. The

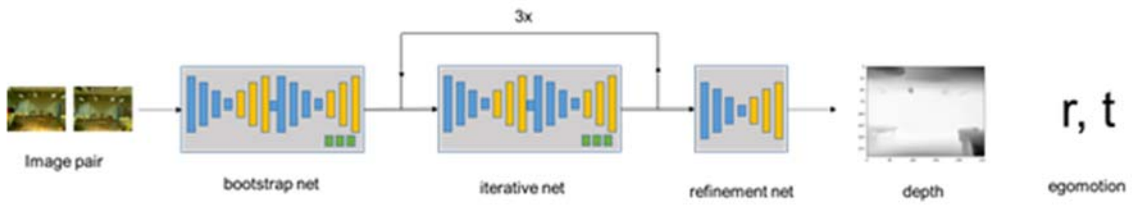


Fig. 1. Overview of the used network.

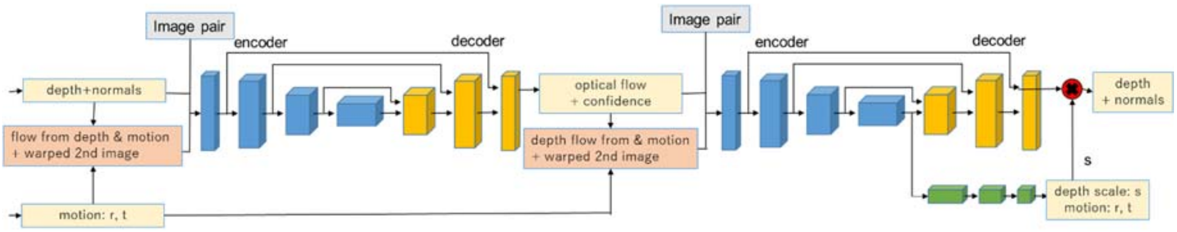


Fig. 2. Schematic of the networks used in the bootstrap net and in the iterative net

difference between each walking motion pattern is quantified from the experiment considering multiple walking patterns, and an index is obtained which represents whether the walking posture is normal or not. This method is distinguished from existing methods by adopting a monocular self-mounted camera. It is also claimed that the proposed system can be used easily by elderly people with a small, light-weight camera.

2. Methods

This section describes a method for estimating the rotation angle of a chest-mounted camera for posture estimation. DeMoN (Depth and Motion Network) is used for camera rotation angle estimation.

2.1. *Camera motion estimation*

Two images are fed into the used neural network shown in **Fig. 1**. The two image frames are taken out every n frames on a video. The relative angle from the initial camera is obtained by calculating the product of the relative angles of the output at times t and t+1. In addition, trend removal is performed for the purpose of removing accumulated errors. The left network in **Fig. 2** predicts the optical flow and its reliability, and the right network predicts the depth image, camera motion and surface normal. The bootstrap net only forwards the fed image pairs to the left network of the iterative net, while the iterative net also inputs depth images, camera motion, and surface normals output from the right network in Fig. 2 (i.e., the iterative net inputs its own outputs by feedback, other than the input from the bootstrap net). The iterative net is executed three times to improve the estimation accuracy of the optical flow, depth image and camera motion at the same time. The refinement net shown in the right network of Fig. 1 generates depth images from low resolution (64 × 48) to high resolution (255 × 192), as shown in Fig. 1. This corrects the incorrect depth prediction from the edge of the depth image.

2.2. *Correction based on camera motion*

Trend removal calculates the sum of squares with a plausible line for the data by simple regression analysis, and calculates the least square approximation line that minimizes it. This is done by finding the difference between the obtained least squares approximation line and the y=0 line and subtracting the difference from the data. In addition, since the initial angle is not 0 due to the trend removal, correction based on the camera motion is performed. The straight line connecting the start point and the end point of the data from which the trend has been removed is used as the straight line before correction. Also, the starting point being 0, the end point is obtained from the camera motion obtained by inputting the first and last image pairs of the video to the network, and the straight line connecting them is taken as the corrected straight line. Once the difference between these two lines are found, it is subtracted from the trend-removed data.

2.3. *Indicators for walking analysis*

The extrema for the corrected data is found. Then, in order to consider only extreme values corresponding to significant posture changes excluding camera vibrations, only those extreme values whose absolute value of the difference between successive extreme values is greater than or equal to a threshold value are used. The average of the obtained maximum and minimum value is used as an index of gait analysis.

3. Experiment

This chapter describes the evaluation method and results of the posture estimation experiment using the proposed method.

3.1. *Experimental environment*

In the experiment, we fabricated and used the device shown in **Fig. 3** that is equipped with a camera and a sensor and can be mounted on the chest. As experimental data, some images were taken by the chest-mounted camera in an indoor environment. At the same time, the Raspberry Pi was used to obtain the values of a 9-axis sensor consisting of acceleration, gyro, and geomagnetism. Equation (1) shows the internal parameter matrix of the camera obtained by camera calibration. Experiments were performed with five different walking patterns. The details of the walking patterns are shown below.

- Normal walk as usual.
- Bow back walk as if dragging the upper body.
- Warp walk with the upper body down.
- Drag leg walk with one leg dragged.
- Tilt sideway walk with the upper body leaning to the left or right.

$$K = \begin{bmatrix} 631.68 & 0 & 314.64 \\ 0 & 629.62 & 240.94 \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

3.2. *Experimental method*

The video obtained from the camera is cut into images every 10 frames and inputted to the neural network. Then the camera rotation angle between the two input images is output. By accumulating these angles, the camera rotation angle from the initial frame to the current frame is obtained. The true value is obtained by integrating the values obtained from the 9-axis sensor using the RTQF algorithm [7]. However, as the sampling interval of the sensor is normally different from that of the camera, the sensor information obtained at the time closest to the camera sampling time is used in the experiment. In addition, the initial posture of the sensor is set to 0 [deg].

3.3. *Experimental Results*

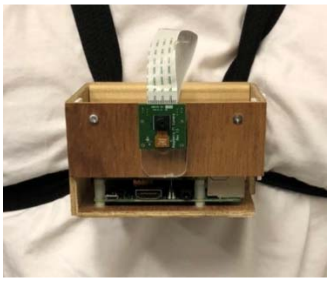


Fig. 3. Experimental equipment (a camera & a 9-axis sensor)

Fig. 4 shows comparison of the true and the estimated rotation angles of the chest-mounted camera motion. The accuracy of the estimation is evaluated using RMSE (root mean square error). This index is expressed by Eq. (2), where N is the number of data, x_i is an estimated value and X_i is a true value. The results are shown in **Table 3**. **Table 4** shows the accuracy evaluation after the correction based on trend removal and camera motion. **Table 5** shows the average of the maximum and the minimum effective values for gait analysis.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - X_i)^2} \quad (2)$$

Table 3. Accuracy of the camera motion estimation by RMSE [deg].

	roll	pitch	yaw
Normal	3.00	7.00	5.53
Bow back	10.26	9.78	13.02
Warp	4.40	27.27	25.21
Drag leg	4.52	12.15	13.42
Tilt sideway	4.71	5.46	8.07

Table 4. Accuracy of the camera motion estimation after correction (RMSE)

	roll	pitch	yaw
Normal	1.42	1.20	3.27
Bow back	6.83	3.74	11.94
Warp	3.29	6.17	10.45
Drag leg	1.71	1.30	3.53
Tilt sideway	3.40	3.23	3.67

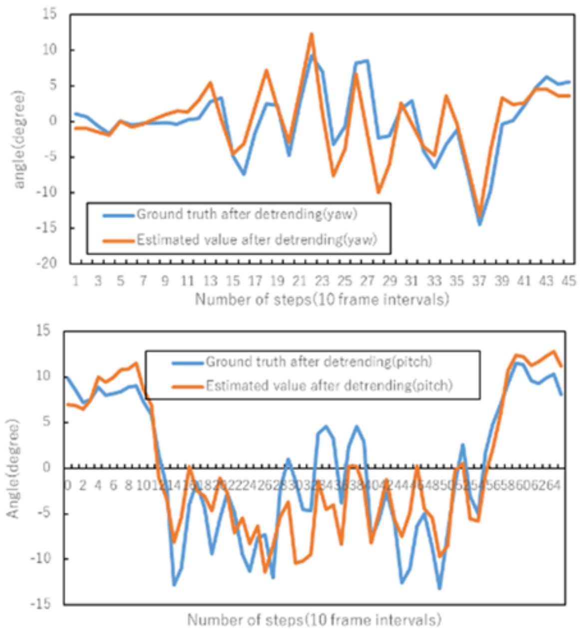


Fig. 4. Comparison of the true and the estimated camera rotation angles.

4. Discussion

The results in Table 4 are more accurate than the results in Table 3, indicating that correction based on trend removal and camera motion is effective. Looking at Table 5, one can see that the results are related to walking patterns. For example, when comparing 'normal' and 'drag leg', the maximum and the minimum values of roll average and pitch average differ only by about 1 degree. However, in the yaw average, the maximum and the minimum values differ by about 10 degrees. This is probably because ‘normal’ was good for the left and the right balance, but ‘drag leg’ was biased towards either side, because it was dragged. From these results, there is a relation between the walking pattern and the rotation angle of the chest-mounted camera. Therefore, it is considered possible to improve the walking posture by analyzing the motion of the chest-mounted camera.

5. Conclusion

In this paper, we proposed a method of analyzing a pedestrian’s walking posture using MY VISION, the images provided from a chest-mounted camera. The data obtained by estimating the camera motion using DeMoN was processed by removing the trend. To evaluate the accuracy, a device containing a 9-axis sensor and a camera was manufactured and worn on the chest of a subject. The 9-axis sensor and the camera were synchronized and the data provided from them were analyzed. Experiments using five walk patterns were conducted in an indoor environment and the results showed the effectiveness of the proposed method.

Acknowledgement

This work was supported by JSPS KAKENHI Grant Number 16K01554

6. References

1. D. Mehta, S. Sridhar, O. Sotnychenko, H. Rhodin, M. Shafiei, H. P. Seidel, W. Xu, D. Casas, C. Theobalt: “VNect: Real-time 3D human pose estimation with a single RGB camera”, *ACM Transactions on Graphics*, Vol.36, Issue 4, Article No. 44, 2017.
2. Tan, J. K., Ishikawa, S., Yamaguchi, I., Naito, T., Yokota, M.: “3-D recovery of human motion by mobile stereo cameras”, *Artificial Life and Robotics*, Vol.10, No.1, pp.64-68, 2006.
3. Tan, J. K., Ishikawa, S.: “Deformable shape recovery by factorization based on a spatiotemporal measurement matrix”, *Computer Vision and Image Understanding*, Vol.82, No.2, pp.101-109, 2001.
4. B. Bodenheimer, C. Rose, S. Rosenthal, J. Pella: “The process of motion capture: Dealing with the data”, *Computer Animation and Simulation’97*, pp.3-18, 1997.
5. T. Shiratori, H. S. Park, L. Sigal, Y. Sheikh, J. K. Hodgins: “Motion capture from body-mounted cameras”, *ACM Transactions on Graphics*, Vol. 30, No. 4: Article No. 31, 2011.
6. B. Ummenhofer et al.: “DeMoN: Depth and motion network for learning monocular stereo”, *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.
7. RTIMULib2(https://github.com/richardstechnotes/RTIMULib2).

Simultaneous Space Object Recognition and Pose Estimation by Convolutional Neural Network

Roya Afshar
School of Instrumental Science and Opto-electronics Engineering, Beihang University, Beijing,100191, China

Zhongyi Chu
School of Instrumental Science and Opto-electronics Engineering, Beihang University, Beijing, 100191, China

Shuai Lu
Department of Mathematics, Beijing University of Chemical Technology, Beijing, 100191, China
E-mail: roya_afshar@buaa.edu.cn, chuzystar@gmail.com, lushuaie@gmail.com,
www.buaa.edu.cn

Abstract

The increasing population of orbital debris is considered as a growing threat to space missions. For this purpose, Convolutional Neural Network was implemented based on transfer learning and data augmentation in order to conduct satellite classification and pose regression. In addition, the effects of un-centered and noisy images as well as different illumination conditions were analyzed by implementing different pre-trained networks. Based on the results, the present method could identify satellites and evaluate their poses against different space conditions effectively.

Keywords: Space debris; Non-cooperative satellite; Recognition space target; pose estimation; convolutional neural network

1. Introduction

The earth orbit is based on a critical situation due to a large number of population in space debris, which is defined as unproductive manmade objects of all sizes. The risk of failure for functional satellites and manned space missions is decreasing because of an increase in the number of orbital debris[2]. To this end, a large number of studies have focused on many enabling space debris for identifying, capturing, and removal technologies. Thus, automated recognition and on-board pose estimation of an uncooperative target spacecraft by using passive sensors such as monocular cameras is regarded as a major task for the removing missions, see Refs. 1-4 for more details.

In another study,[5] found the number of published solutions of automated space target recognition. The performance of traditional algorithms like scale-invariant feature transform (SIFT), local binary pattern (LBP), and histogram of oriented gradient (HOG) rely on extracting hand engineered features, as well as engineering the related feature.[4] However, producing effective identification data is difficult in space application due to the robust less features and the significant gap between visual features and targets. Thus, discriminatory method is necessary for identifying the target, for more information see Refs. 4 and 5.

Further, the studies focused on monocular pose determination for space applications can be divided into 3D model-based methods and 2D image-based methods like in Refs. 7 and 8. It is worth noting that 3D model-based methods necessitate a prior 3D CAD model or 3D point cloud, which seem difficult to be obtained in practice. Most of the image-based methods are based on classical image processing algorithms, which are prevented by the computational elaboration leading to the evaluation of many pose hypotheses.[4] Further, it results in creating an a-priori knowledge of the pose which is always inaccessible and inflexible to different structural and physical kinds of space crafts.[9] Furthermore, monocular navigation is not robust enough due to the low image acquisition rate, low signal-to-noise ratio, as well as illumination conditions.[10]

Recently, machine learning-based methods were suggested without considering the above-mentioned limitations some studies could be found in Refs. 11-14. Zhang and Jiang[11] solved multi-view space object recognition by using kernel regression. Haopeng Zhang et al.[12] implemented Gaussian Process Regression, and in another study utilized homeomorphic manifold analysis for recognizing satellite and estimating relative poses space objects. However, these methods are useful for 1D and 2D pose variations, while they cannot be implemented for 3D attitude variations.[13]

In addition, CNN plays a significant role in classifying, detecting, and estimating pose. Further, CNN can learn the extraction and classification of the related feature. A large body of research in the area of terrestrial and space application used deep convolutional neural network in order to eliminate the above-mentioned shortcomings. Zeng et al. implemented a nine-layer deep CNN in order to recognize space target, and data augmentation was used to overcome overfitting due to a small size of dataset[5], while a deeper network is more effective. Sharma et al. proposed an image synthetic pipeline to produce a massive image dataset for any spacecraft with 3D model, and then CNN method was utilized for an initial guess of real time pose.[4] In another study Regional Proposal Network (RPN) based on CNN was used in order to detect bounding box around satellite and identify the coarse attitude by classification. Finally, a post refinement was conducted for evaluating the relative attitude.[14] In addition, another author trained the CNN by implementing transfer learning method, and discretized 3D space in some regions. Then, a label of the camera where located was considered as the output of CNN.[10] Although some studies such as Refs. 4, 10 and 14 reported using CNN in space navigation, no study, to the best of our knowledge, focused on a unified CNN to recognize satellite and estimate pose estimation simultaneously.

By considering the above-mentioned studies, the present paper aimed to present a CNN architecture for simultaneous satellite recognition and pose estimation. In fact, the contributions of this paper are as follows:

- A two-stage CNN architecture is proposed for classifying type of space target, and regression for posing the satellite. Inception CNN pre-trained with Imagenet dataset is regarded as the first stage. During the second stage, the related network is trained independently for classification and regression with BUAA-SID dataset made by Ref. 15,
- Faster training and better extracting features are done by adopting the transfer learning mechanism,
- A number of augmentation techniques are employed on the training images with KerasDataGenerator Tool through using a number of random transformations like rotating, shifting, rescaling, and zooming due to limited size of BUAA-SID for increasing the accuracy of classification.

The performance of the proposed CNN-based method is evaluated by using BUAA-SID dataset under different noisy and un-centered pictures, as well as focusing on illumination conditions of space environment. The total accuracy and Mean Absolute Error is better than that of the state-of-the-art methods in Ref. 13.

2. Method and Dataset

In the following section, the learning problem is formulated for detecting and estimating pose jointly. Then, the proposed architecture is provided, along with the details of loss functions in our architecture. Finally, the dataset is described. The main problem is related to learning a strong visual representation allowing the model to classify objects and estimate the attitude of the camera frame (C), by considering the body frame of the target spacecraft (B). As shown in Figure 1, R indicates the rotation matrix which aligns the target body reference frame with that of the camera which is estimated by CNN.[16]

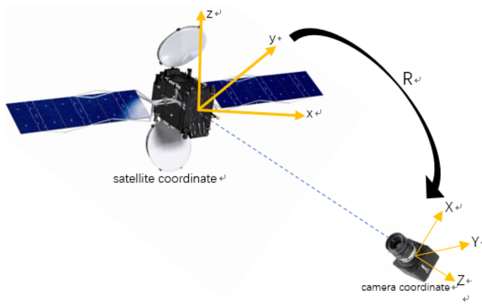


Fig. 1. A representation of the camera and satellite coordinate system

Thus, the learning process starts from a training set $S = \{(x_i, t_i)\}_{i=1}^{N}$ where N indicates the number of training samples. For each sample i in the dataset, $x_i \in X$ represents the input image, and $t_i \in T$, which $t_i = (y_i, \emptyset_i)$, encodes the annotations for solving classification (y_i) and pose estimation $(\emptyset_i)$. $y_i \in Y$ with $Y = [1,2, \ldots, C, C+1]$ describes the object class, where C shows the total number of satellite categories (Fig. 2). Finally, $\emptyset_i \in \mathbb{R}^3$ encodes the 3D viewpoint annotation for a special space object with respect to the camera position as a tuple of azimuth, elevation, and zenith angles.[17]

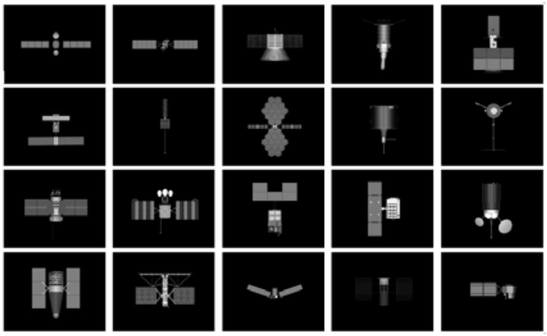


Fig. 2. 20 different categories of satellites

It is proposed to learn a convolutional neural network (CNN) for detecting satellite and estimating pose simultaneously. Technically, each neuron in a layer receives inputs from a set of neurons placed in previous layer called "receptive field". In the forward pass, the output characteristic of each neuron is considered as the convolution of the input features extracted from the previous layer. Thus, CNN transforms input image into the features related to the map utilized to solve the target tasks.

Prediction $\hat{t}$ is defined as Eq.(1) in order to the problem related to the detection of objects and estimation of viewpoint simultaneously:

$$\hat{t}_{\theta,W} = F_W . z_\theta(x_i). \quad (1)$$

$z_\theta : X \rightarrow \mathbb{R}^D$ indicates the D-dimensional feature mapping performed by the network for the input images. Technically, it involves the transformation of the input image x_i into the features used to feed the output layers of our models. The trainable weights of the deep architecture are encoded in θ which lets the network to perform the mapping. The weights in θ define the hidden layers which are shared by all of the tasks that the deep network needs to solve.[17]

F_W is related to the set of functions of the output layers, which take the deep feature map $z_\theta(x_i)$ as input. Regarding the problem stated in this paper, classification (y) and viewpoint estimation (∅) should be addressed in the present set of functions. Therefore, $F_w = (f^y_{Wy}, f^\emptyset_{W\emptyset})$, where f^y_{Wy} with weights Wy produces the predictions for the object category, i.e.$\hat{t}^y$, and $f^\emptyset_{W\emptyset}$ is in charge of predicting the viewpoint $\hat{t}^\emptyset$.

Based on the prediction model in Eq. (1), the following objective function is defined for learning our multi-task neural network:

$argmin_{\theta,W} \mathcal{L}(\theta, w, s)$,

where the loss function follows the Eq. (2),

$$\mathcal{L}(\theta, W, S) = \lambda_1 \mathcal{L}_y(\theta, W^y, S) + \lambda_2 \mathcal{L}_\phi(\theta, W^\phi, S). \quad (2)$$

where λ indicates the scalar value which controls the significance of a special loss during the training process. Regarding the classification loss, a Categorical Cross-entropy function is used, and the pose estimation loss was traditionally considered from continuous and discrete perspectives. In discrete formulation, pose estimation is regarded as a classification problem while it is solved by regression in continues problem. In the present study, Mean squared error was used to solve the regression problem due to coarse quantization of our dataset.

2.1. *Network architecture*

In this architecture, pose estimation task is slightly separated from the satellite class detection. As shown in Figure 3, the proposed extension for these two objectives includes two stages. During the first stage, a pre-trained network is used with hyper-parameter of Imagenet dataset for extracting the feature. However, independent FC layers connected to the pose regression and satellite classification layers in the second stage. The model is learned by solving the objective function displayed in Eq. (2). During the training process for the detection path, $\lambda_2 = 0$ and $\lambda_1 = 0$ for the pose path.

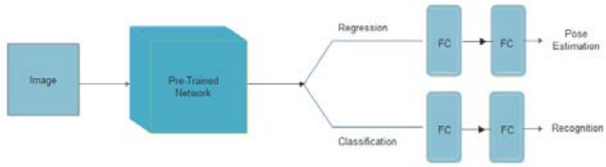


Fig. 3. An overview of the proposed CNN structure

2.2. *Dataset*

In the present study, BUAA-SID 1.0 and BUAA-SID 1.5 were used for training and testing CNN.

Train Set 1: In BUAA-SID 1.0, there is a subset including 4600 gray images among the 20 satellites from 230 viewpoints sampled on a viewing sphere as used in Refs. 11 and 15 and 18.

Train Set 2: BUAA-SID 1.5 consists of four subsets. The 1D subset as the first subset includes 3600 grayscale images of ten satellites captured from 360 viewpoints uniformly sampled on a circle with the pitch angle $\varphi = 0$ and the yaw angle $\theta \in [0,2\pi)$, as shown in Refs. 11 and 13 and 8.

Test Set 1: a test set was made by adding Zero Mean Gaussian White (ZMGW) noise with variances from 0.01 to 0.1 steps of 0.01 to the pictures of Train Set 1 as displayed in part 'a' of Fig. 4.

Test Set 2: The lighting subset is considered as the third subset. This subset includes 10080 gray images of one satellite from the same viewpoints as the 1D subset,

which were simulated in different lighting conditions. In other words, the phase angle of the light ranged from 0° to 90° in the steps of 10° while the altitude angle of the light was in the range of 0° , 90°, and 180° , separately.
Test Set 3: Un-centered pictures from Train Set 1 were used to created another test set. In this regard, Matlab was used to put each satellite in image randomly in image plane as shown in Fig. 4b.

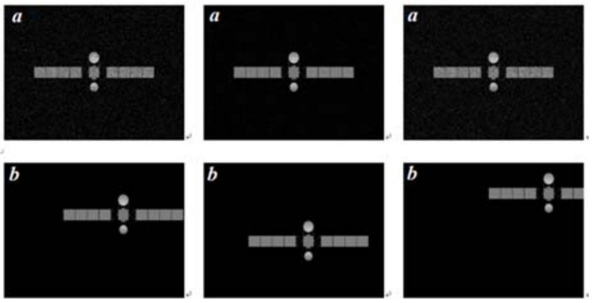


Fig. 4. a) ZMWG noisy pictures with variances 0.1, 0.01, 0.05, and b) Un-centered pictures

2.2.1 Augmentation of Dataset

In general, a number of data augmentation techniques are used for the training images in order to increase the size of dataset for classification task in which more instances are needed for training. A number of random transformations like rotating, shifting, rescaling, zooming related to image matrices were applied through KerasDataGenerator Tool, the link of website is available in Ref. 19.

3. Experiments

In order to conduct the experiment, the models and loss functions were implemented by using the deep learning framework of Keras.[19] The Stochastic Gradient Descent algorithm with the momentum of 0.9, weight decay of 0.0001, and learning rate of 0.001 was used for optimization in order to assure that the network converges properly.
InceptionResNetV2 was utilized after trying pre-trained popular networks with Imagenet Dataset like VGG19, InceptionV3, Xception, ResNet, and InceptionResNetV2 the popular CNN proposed in Ref. 20 due to the better results in our case. The remainder of structure is trained by BUAA-SID dataset. The batch size involves 32 samples during 20 epochs. Accuracy and Mean Absolute Error(MAE) are the evaluation metric for satellite classification and pose regression, respectively. The accuracy increased significantly by initializing the second stage of the network with 'random_uniform' function in Keras. Then, by 'ImageDataGenerator', as the image preprocessing Tool of Keras, our shallow dataset was augmented via a number of random transformations like rotating, shifting, rescaling, zooming in order to prevent under-fitting the network. Finally, the FCs layers were fine-tuned by the augmented dataset. Figure 5 displays the accuracy and loss function changes during 20 epochs. The accuracy of training and testing increased when they are close to each other, which reached to 100% accuracy. In addition, loss curves on the train and test set decreased, which reached close to zero at the last epochs. Based on the behavior of these curves, network was trained correctly without any overfitting or under fitting. The accuracy for classification was 99.24%, while the value decreased to 97.07% for un-centered pictures.

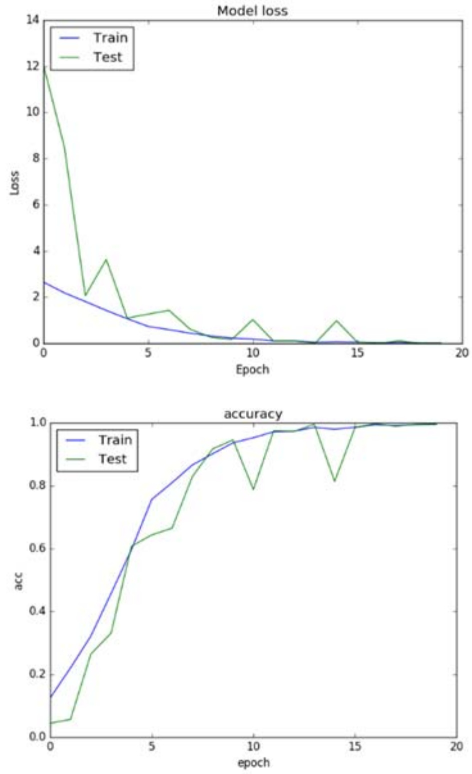


Fig. 5. Evaluation of training procedure of CNN, a decrease in loss function to very close to zero, and an increase in accuracy to around 100% emphasizing a fine-tuned CNN

3.1. *Robustness against noise*

Sub-headings Practically, the images captured by a real space-based imaging system may have noise due to the effect of the space environment and imaging system itself. Gaussian noise is usually regarded as modelling such noise for theoretical analysis and simulation [16]. To evaluate robustness against noise, the experiments were conducted with and without noisy pictures. In other words, CNN was first trained by dataset including pictures without noise. Then, the trained CNN was tested by noisy pictures with different variances and zero mean as shown in Figure 6. In this experiment, the accuracy decreased to less than 10% significantly. During the second experiment, CNN was trained by implementing subset containing noisy pictures with the variances of 0.01, 0.05, and 0.1. Finally, the CNN was tested like the previous experiment, the results of which are illustrated in Figure 6. In this case, the total accuracy increased significantly for lower variances, while it decreased gradually for the variances more than 0.05, where the accuracy was similar to that of the first experiment in the variance of 0.1. Based on the experimental result in Figure 6, CNN learned with trainset containing noisy pictures is more robust against noise compared to the case when trainset is devoid of noisy pictures. Figure 7 shows the experimental results for pose estimation with noisy test-sets. A slight change was observed in mean absolute error against increasing the variance, which emphasizes the robustness of CNN for pose regression with noisy pictures.

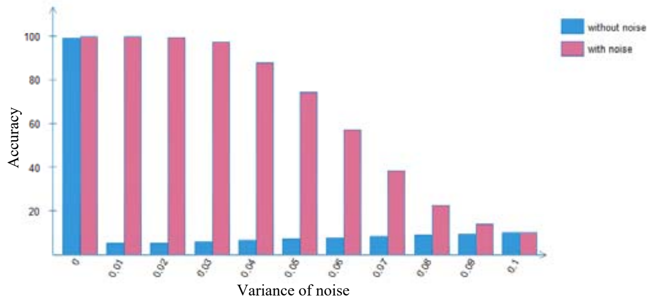


Fig. 6. Results of recognition accuracy on noise subset for with and without noisy pictures

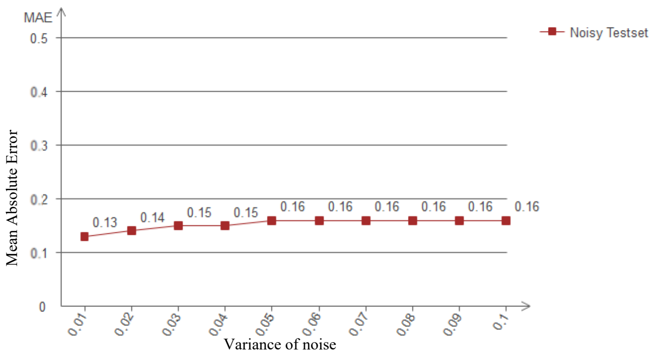


Fig. 7. Results of pose estimation on noise subset

3.2. *Robustness against lighting*

In this section, the network is trained with the pictures not affected by lighting conditions. In fact, the light angle of train set becomes zero. Then, the trained network was evaluated with test-sets with different light angles. Figure8 displays the accuracy of classification in different light angles. As shown, the accuracy of the recognition of satellite decreases by increasing the light angles, especially for light angle more than 30° . Further, the results were satisfactory for the mean absolute error of pose regression. Furthermore, slight changes were observed by increasing light angle (Figure 9). The value in Ref. 13 is reported less than 10 by using Homeomorphic Manifold Analysis, while it is less than 1 in the present case. It is worth noting that CNN was taught with train set without including the pictures with different lighting angles. The recognition accuracy and mean absolute error were improved in real applications by boosting the train set with the pictures in different lighting conditions.

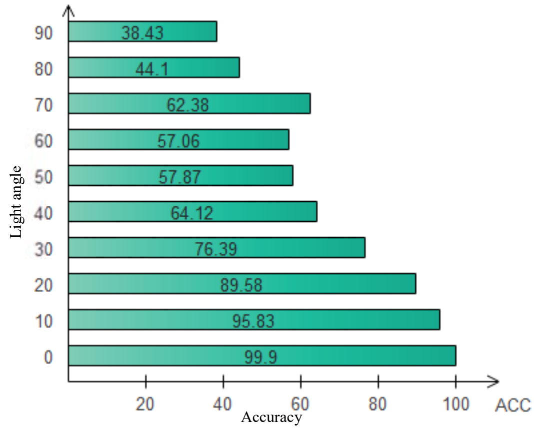


Fig. 8. Recognition accuracy on light subset

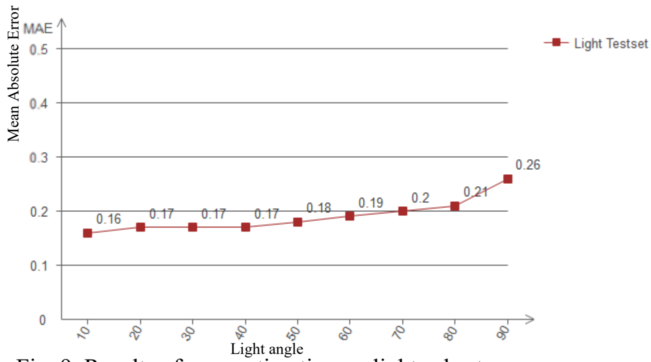


Fig. 9. Results of pose estimation on light subset

4. Conclusion

In the present study, a CNN-based approach was proposed for recognizing satellite and estimating pose simultaneously by implementing classification and regression method, respectively. Regarding the CNN training, transfer learning method was used for the first stage of the network, while data augmentation method was utilized for the rest of network trained with synthetic images of BUAA-SID dataset in the case of classification in order to increase the accuracy. Finally, the robustness of CNN was evaluated in the space environment conditions like noisy and un-centered pictures and illumination conditions. Based on the experimental results, CNN can deal with space conditions appropriately when this situation was considered in train-set. Further research can be conducted for evaluating the performance related to the mitigation of the noise in the pictures to input clean pictures to the network.

References

[1] Minghe Shan, Jian Guo,Eberhard Gill, “Review and comparison of active space debris capturing and removal methods,” Progress in Aerospace Sciences, 2015.

[2] Rong-yu Sun, Jin-wei Zhan, Chang-yin Zhao, Xiao-xiang Zhang “Algorithms and applications for detecting faint space debris in GEO,” Acta Astronautica, 2015.

[3] Bikram Pradhan, Paul Hickson, Jean Surdej, “Serendipitous detection and size estimation of space debris using a survey zenith-pointing telescope,” Acta Astronautica, 2019.

[4] Sumant Sharma,Connor Beierle,Simone D’Amico, “Pose Estimation for Non-Cooperative Spacecraft Rendezvous Using Convolutional Neural Networks,” in IEEE Aerospace Conference , 2018.

[5] Haoyue Zeng, Yong Xia, “Space Target Recognition based on Deep Learning,” in 20th International Conference on Information Fusion, 2017.

[6] Tae Ha “Jeff” Park, Simone D’Amico, “ESA Pose Estimation Challenge 2019,” Space Rendezvous Laboratory Department of Aeronautics and Astronautics Stanford Universityhttps://slab.stanford.edu/, 2019.

[7] Tae Ha Park, Sumant Sharma, Simone D’Amico, “Towards Robust Learning-Based Pose Estimation Of Noncooperative Spacecraft,” in AAS/AIAA Astrodynamics Specialist Conference, 2019.

[8] Haopeng Zhang, Zhiguo Jiang, Ahmed Elgammal, “Vision-based pose estimation for cooperative space objects,” Acta Astronautica, 2013.

[9] Sumant Sharma, “Pose Estimation Of Uncooperative Spacecraft Using Monocular Vision and Deep Learning,” Phd Thesis, 2019.

[10] Sumant Sharma,Connor Beierle,Simone D’Amico, “Towards Pose Determination For Non-Cooperative Spacecraft Using Convolutional Neural Networks,” in 1st IAA Conference on Space Situational Awareness (ICSSA), 2019.

[11] Zhang Haopeng, Jiang Zhiguo, “Multi-view space object recognition and pose estimation based on kernel regression,” Chinese Journal of Aeronautics, 2014.

[12] Haopeng Zhang ,Zhiguo Jiang ,Yuan Yao ,Gang Meng, “Vision-based pose estimation for space objects by Gaussian process regression,” in IEEE Aerospace Conference, 2015.

[13] Haopeng Zhang , Zhiguo Jiang , Ahmed Elgammal, “Satellite Recognition and Pose Estimation Using Homeomorphic Manifold Analysis,” IEEE Transactions on Aerospace and Electronic Systems, 2015.

[14] Sumant Sharma, Simone D’Amico, “Pose Estimation For Non-Cooperative Spacecraft Rendezvous Using Neural Networks” in AAS/AIAA Astrodynamics Specialist Conference, 2019.

[15] Zhang, Haopeng and Liu, Zhengyi and Jiang, Zhiguo and An, Meng and Zhao, Danpei, “{BUAA-SID1.0} space object image dataset,” Spacecraft Recovery \& Remote Sensing, 2010.

[16] Y. Xiang, T. Schmidt, V. Narayanan, and D. Fox, “PoseCNN: A Convolutional Neural Network for 6D Object Pose Estimation in Cluttered Scenes.,” 2017.

[17] Daniel Oñoro-Rubio, Roberto J. López-Sastre, Carolina Redondo-Cabrera, Pedro Gil-Jiménez, “The challenge of simultaneous object detection and pose estimation: A comparative study,” Image and Vision Computing, 2018.

[18] Meng, Gang and Jiang, Zhiguo and Liu, Zhengyi and Zhang, Haopeng and Zhao, Danpei, “Full-viewpoint 3D Space Object Recognition Based on Kernel Locality Preserving Projections,” *Chinese Journal of Aeronautics*, vol. 23, pp. 563-572, 2010.

[19] Keras: The Python Deep Learning library, https://keras.io/

[20] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, Alex Alemi, “Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning,” in *arXiv:1602.07261 [cs.CV]*, 2016.

An error correction mechanism for reliable chemical communication systems

Masashi K. Kajita
Institute of Industrial Science, The University of Tokyo,
4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan
E-mail: mkajita@sat.t.u-tokyo.ac.jp

Abstract

Chemical communication systems, such as bio-inspired chemical sensory systems or biological cells, sense the environment by detecting target ligand molecules, which convey environmental information. However, non-target ligands, similar to the target ones, are ubiquitous in the environment and can hamper accurate information transmission. In this work, we investigate an error correction mechanism for reliable chemical communications and find an intuitive understanding of how the mechanism can amplify the small difference between the target and non-target ligands. We also demonstrate that the mechanism can balance accuracy and output intensity. Our approach may provide a method to design reliable chemical communication systems.

Keywords: Bio-inspired communication, chemical reaction network, chemical sensory system, error correction, kinetic proofreading

1. Introduction

Chemical communication systems, such as bio-inspired chemical sensory systems or biological cells, sense the environment by detecting target ligand molecules, which convey environmental information[1]. The chemical communication system detects the target ligands using receptor molecules, which act as sensors for environmental information. However, non-target ligands, similar to the target ones, are ubiquitous in the environment. Due to the structural similarity, the non-target ligands have affinities to the receptors and can attach to the receptor molecules. This non-specific and undesired interaction may send an erroneous signal into the system and hamper accurate information transmission. Even under the effect of non-target ligands, biological cells such as immune T cells or chemical reaction systems for the translation and transcription of genetic information have high fidelity to the target ligands[2-6].

The series of experimental observations suggest that biological systems have some error correction mechanisms to detect the target ligands for reliable chemical communication. To reveal the underlying mechanism of the biological error correction systems for chemical communications, the zero-order proofreading model[7] was recently proposed as an extended model of the kinetic proofreading model, which is the pioneering works by Hopfield and Ninio for a biological error correction mechanism[8,9].

In this paper, we first model chemical sensing. Then we introduce the zero-order proofreading model7 and investigate the intuitive understanding of how the model can precisely discriminate the target ligand from the similar non-target one based on their affinity parameters. We further investigate the target detection performance of the zero-order proofreading model under the condition that the target and non-target ligands exist in the environment simultaneously.

2. Zero-order proofreading mechanism for ultra-specificity

2.1. *Modeling of chemical sensing*

We start with considering the simplest situation that only a single type of ligand molecule exists in the environment (Fig. 1). In this situation, a chemical communication system or a cell distinguishes whether the ligand molecule is the target ligand or non-target one by using receptor molecules, which work as chemical sensors for the environmental signal. The receptor R can detect the ligand L, and the ligand-receptor binding and unbinding reactions are described as follows:

$$\mathrm{R}+\mathrm{L} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{C}_0, \qquad (1)$$

where k_1[s^{-1}M^{-1}] and k_{-1}[s^{-1}] are the reaction constants of the binding and unbinding reactions, respectively. The complex of the receptor R and the ligand L is denoted as C_0.

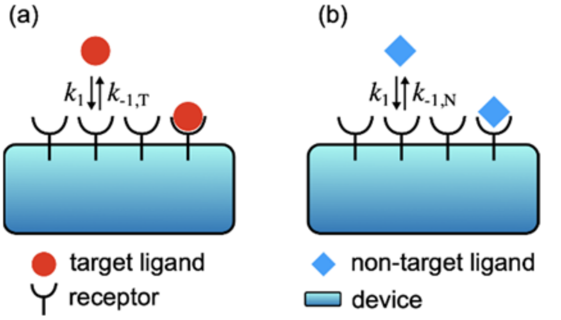


Fig. 1. Schematics of a chemical communication system when (a) the target ligand molecule or (b) non-target one exists in the environment. The reaction constants of the binding and unbinding reactions are respectively k_1[s^{-1}M^{-1}] and k_{-1}[s^{-1}] where * is T and N for the target ligand and non-target one, respectively. We assume that $k_{-1,\mathrm{T}}$[s^{-1}] < $k_{-1,\mathrm{N}}$[s^{-1}] because the non-target ligand is more likely to dissociate from the receptor.

We assume that the target and non-target ligands have similar sizes and structures. Then, we can assume that the difference between the target and non-target ligands is only the unbinding rate k_{-1} (Fig. 1). If we denote the unbinding rates of the target ligand $k_{-1,\mathrm{T}}$[s^{-1}] (Fig. 1 a) and that of the non-target one $k_{-1,\mathrm{N}}$[s^{-1}] (Fig. 1 b), the unbinding rates satisfy $k_{-1,\mathrm{T}}$[s^{-1}] < $k_{-1,\mathrm{N}}$[s^{-1}] because the non-target ligand is more likely to dissociate from the receptor.

2.2. *Zero-order proofreading model*

If the target and non-target ligands are similar molecules, the difference in the unbinding rates among $k_{-1,\mathrm{T}}$[s^{-1}] and $k_{-1,\mathrm{N}}$[s^{-1}] is small. For reliable chemical communications, the system must amplify the small difference of the unbinding rate k_{-1}. To this end, the following phosphorylation and dephosphorylation cycle was proposed in Ref. 7 as an intracellular chemical reaction network that can amplify the small difference in k_{-1} (see also Fig. 2):

$$\mathrm{R}+\mathrm{L} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{C}_0 \xrightarrow{w} \mathrm{C}_1 \xrightarrow{w} \mathrm{R_P}+\mathrm{L}, \quad (2)$$

$$\mathrm{R_P}+\mathrm{P} \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \mathrm{D} \xrightarrow{k_3} \mathrm{R}+\mathrm{P}.$$

After the formation of the ligand-receptor complex C_0, the receptor can be converted into the intermediate state of the complex denoted by C_1 with rate w[s^{-1}] via an irreversible reaction. The product $\mathrm{R_P}$ can be generated from C_1 with rate w. However, some of C_1 can be converted back to R+L due to the ligand dissociation from the receptor with rate k_{-1}. Note that we assume that the two unbinding reactions, $\mathrm{C}_0\to$R+L and $\mathrm{C}_1\to$R+L, have the same rate k_{-1}[8-10]. This assumption is valid if the chemical modification of the receptor, which is carried out by the reaction $\mathrm{C}_0\to\mathrm{C}_1$, only affects the intracellular part of the receptor and does not affect the unbinding rate k_{-1}. We also note that the unbinding reaction $\mathrm{C}_1\to$R+L is an irreversible reaction for an error correction function, which is called kinetic proofreading[8-10]. The product $\mathrm{R_P}$ is covalently modified and activated receptor R, and $\mathrm{R_P}$ can be deactivated by an enzyme P. D denotes the complex of $\mathrm{R_P}$ and P. The reaction constants of the binding and unbinding reactions are k_2[s^{-1}M^{-1}] and k_{-2}[s^{-1}], respectively. The receptor R is generated from D with rate k_3[s^{-1}].

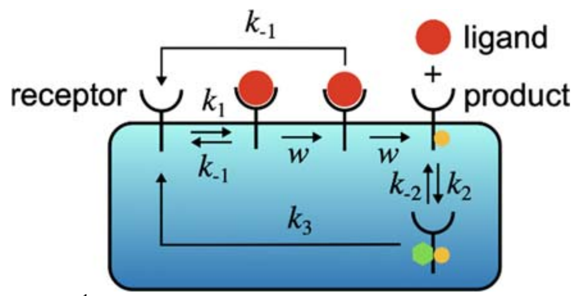


Fig. 2. Schematic diagram of the zero-order proofreading model (eq. 2) in a chemical communication system. The receptors detect the ligands in the environment, then transmit the signal into the inside of the system using the chemical reactions. The system distinguishes whether the ligand is the target ligand or not by converting the unbinding rate k_{-1} into the amount of the

product molecules $[R_P]$ through the zero-order proofreading model.

2.3. *Model reduction based on Michaelis-Menten approximation*

By assuming that each reaction follows the law of mass action, the dynamics of the zero-order proofreading model (Eq. 2) can be described by ordinary differential equations (ODEs). To obtain an intuitive understanding of how the zero-order proofreading model amplifies the small difference among the target and non-target ligands, we perform a model reduction based on Michaelis-Menten approximation[11].

If the total concentration of receptor $[R]_{total} := [R] + [R_P] + [C_0] + [C_1] + [D]$ is much larger than those of ligand and enzyme, $[L]_{total} := [L] + [C_0] + [C_1]$ and $[P]_{total} := [P] + [D]$, $[R_P]_{total} \approx [R]_{total} \approx [R] + [R_P]$ holds, where $[X]$ denotes the concentration of molecule X, whose unit is [M]. At the quasi steady state, that is, $d[C_1]/dt = d[C_0]/dt = d[P]/dt = 0$, the dynamics of $[R_P]$ can be described by the following ODE:

$$\frac{d[R_P]}{dt} = V_1 \frac{[R]_{total} - [R_P]}{K_{m,1} + [R]_{total} - [R_P]} - V_2 \frac{[R_P]}{K_2 + [R_P]}, \tag{3}$$

where $V_1 := [L]_{total} w\alpha/(1+\alpha)$, $V_2 := k_3 [P]_{total}$, $K_{m,1} := K_1/(1+\alpha)$, $K_1 := (k_{-1}+w)/k_1$, $K_2 := (k_{-2}+k_3)/k_2$, and $\alpha := w/(k_{-1}+w)$.

The production speed of $[R_P]$ can be decomposed into the positive and negative fluxes, which are denoted as J_1 and J_2, respectively. The definitions are

$$J_1 := V_1 \frac{[R]_{total} - [R_P]}{K_{m,1} + [R]_{total} - [R_P]}, \tag{4}$$

$$J_2 := V_2 \frac{[R_P]}{K_2 + [R_P]}. \tag{5}$$

Thus, the net flux is given by $J := d[R_P]/dt = J_1 - J_2$. These equations Eqs. 4 and 5 are called Michaelis-Menten equations[11].

The parameters $K_{m,1}$ and K_2 are the effective Michaelis-Menten constants of the positive and negative fluxes, respectively. Note that $K_{m,1}$ and K_2 respectively control the ligand and enzyme saturation levels. At the unsaturated condition, when $K_{m,1} \gg [R_P]_{total}$ and $K_2 \gg [R_P]_{total}$, Eqs. 4 and 5 become

$$J_1 \approx \frac{V_1}{K_{m,1}}([R]_{total} - [R_P]), \tag{6}$$

$$J_2 \approx \frac{V_2}{K_2}[R_P]. \tag{7}$$

These fluxes are approximately the first-order reactions with respect to $[R_P]$, which are demonstrated in Fig. 3 (a).

$$J_1 \approx V_1, \tag{8}$$

$$J_2 \approx V_2. \tag{9}$$

On the other hand, at the saturated condition, when $K_{m,1} << [R_P]_{total}$ and $K_2 << [R_P]_{total}$, Eqs. 4 and 5 become These approximated equations do not depend on $[R_P]$, thus these are called zero-order reactions with respect to $[R_P]$, which are also demonstrated in Fig. 3 (b).

2.4. *An intuitive understanding of the zero-order proofreading mechanism*

Next, we clarify the mechanism of how the saturation level affects the response of the zero-order proofreading model to the unbinding rate k_{-1}. At the steady state, $J_1 = J_2$ holds. By solving $J_1 = J_2$ for $[R_P]$, we can obtain the steady state concentration of $[R_P]$, denoted as $[R_P]^*$. Because J_1 depends on the unbinding rate k_{-1}, $[R_P]^*$ also depends on k_{-1}. This means that $[R_P]^*$ changes depending on whether the ligand is the target or not. To investigate how the saturation level changes the difference in the steady-state concentrations $[R_P]^*$ for the target and non-target ligands, we analyze the dependence of the fluxes, Eqs. 4 and 5, on the unbinding rate k_{-1}. When the system is in unsaturated condition (Fig. 3 a), the small change in k_{-1} slightly affects J_1. Because of the small effect, the steady-state concentrations $[R_P]^*$ for the target and non-target ligands do not differ largely. On the other hand, when the system is saturated (Fig. 3 b), even though the small difference in k_{-1} slightly changes the positive flux, $J_1 \approx V_1 = w\alpha[L]_{total}/(1+\alpha) = w^2/(k_{-1}+2w)$, due to the zero-order reactions, the steady state concentration $[R_P]^*$ largely moves. This is an intuitive understanding of how the zero-order proofreading model (Eq. 2)[7] can amplify the small difference of the unbinding rate k_{-1} between the target and non-target ligands by changing the saturation level.

Note that the highly nonlinear response due to the cycle reaction composed of the two zero-order reactions is called zero-order ultra-sensitivity[11,12]. Although the original model proposed by Goldbeter and Koshland[12] has the ultra-sensitivity to concentration variables ($[L]_{total}$ and $[P]_{total}$ in the zero-order proofreading model), the model cannot amplify the unbinding rate k_{-1}. Our representation in Fig. 3 clarifies how the zero-order proofreading model obtains the ultra-specificity to the target ligand by the nonlinearity to the unbinding rate k_{-1} based on the mechanism of the zero-order reaction.

3. The case of two ligands in the environment

Let us consider a more realistic condition for chemical communication systems, that is, the system is surrounded by more than one type of ligand. As the simplest case, we consider that there are two types of ligands, the target and non-target ligands, exist in the environment simultaneously (Fig. 4). In this situation, we investigate how the zero-order proofreading mechanism contributes to the reliable detection of the target ligand.

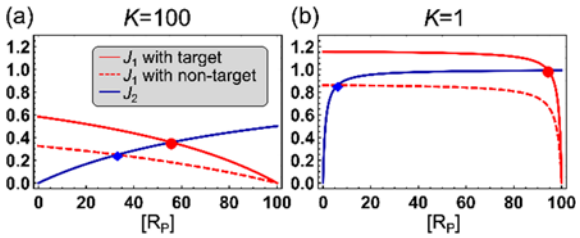


Fig. 3. Fluxes of the activation and deactivation reaction cycle of the zero-order proofreading model (Eq. 2) when the system is in an unsaturated condition (a) and a saturated condition (b). The positive flux J_1 (Eq. 4) with the target ligand (red lines), with the non-target ligand (red dashed lines), and the negative flux J_2 (Eq. 5) (blue lines) are plotted as functions of $[R_P]$. The steady states with the target ligand are plotted as red circles, and the steady states with the non-target one are plotted as blue rhombuses in the figures. Here, K denotes the effective Michaelis-Menten constants, which controls the unsaturation level of J_1 and J_2, respectively, where $K=K_{m,1}=K_2$. The displayed values of K are set for the target ligand, and k_{-1} and k_2 are obtained from the values of K. The other parameters are $k_{-1}=1$ for the target ligand, $k_{-1}=2$ for the non-target one, $w=1$, $k_{-2}=10$, $k_3=1$, $[R]_{total}=100$, $[L]_{total}=3.5$, and $[P]_{total}=1$.

4. The case of two ligands in the environment

Let us consider a more realistic condition for chemical communication systems, that is, the system is surrounded by more than one type of ligand. As the simplest case, we consider that there are two types of ligands, the target and non-target ligands, exist in the environment simultaneously (Fig. 4). In this situation, we investigate how the zero-order proofreading mechanism contributes to the reliable detection of the target ligand.

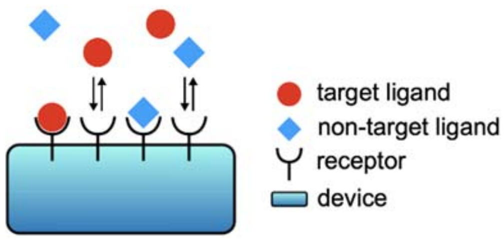


Fig. 4. Schematics of a chemical communication system when the target and non-target ligands exist in the environment simultaneously.

$$\frac{d[R_P]}{dt} = V_{1,T}\frac{[R]_{total}-[R_P]}{K_{m,T}+[R]_{total}-[R_P]} + V_{1,N}\frac{[R]_{total}-[R_P]}{K_{m,N}+[R]_{total}-[R_P]} - V_2\frac{[R_P]}{K_2+[R_P]}, \qquad (11)$$

$$R + L_N \underset{k_{-1,N}}{\overset{k_1}{\rightleftharpoons}} C_{0,N} \xrightarrow{w} C_{1,N} \xrightarrow{w} R_P + L_N,$$

$$R_P + P \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} D \xrightarrow{k_3} R + P,$$

When there are two ligands, the zero-order proofreading model can be extended as follows (see also Fig. 5):

where $[L_T]$ and $[L_N]$ are the target and non-target ligands, respectively. Each ligand has intermediate states, $[C_{0,*}]$ and $[C_{1,*}]$, when it forms a complex with the receptor R, where $* \in \{T, N\}$. Here, $k_{-1,T}$ and $k_{-1,N}$ denote the unbinding rate of the target ligand and that of the non-target one, respectively.

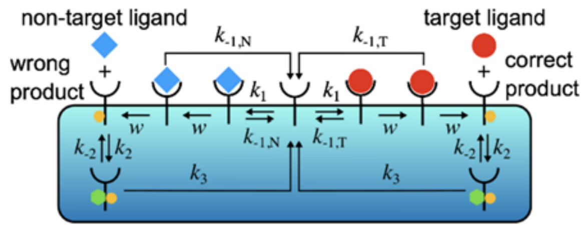


Fig. 5. Schematics of the zero-order proofreading model when the target and non-target ligands exist simultaneously in the environment.

3.1. Model reduction based on Michaelis-Menten approximation

As we did the model reduction of the zero-order proofreading for the single type of ligand, we apply the Michaelis-Menten approximation to the zero-order proofreading model for two types of ligands (Eq. 10).

Assume that the dynamics of Eq. 10 follows the law of mass action, then the dynamics can be described by ODEs. If we also assume that the total concentration of receptor $[R]_{total} := [R]+[R_P]+\sum_{i,*}[C_{i,*}]+[D]$ is much larger than those of the target ligand, $[L_T]_{total} := [L_T]+\sum_{i,*}[C_{i,T}]$, non-target ligand, $[L_T]_{total} := [L_T]+\sum_{i,*}[C_{i,T}]$, $[L_N]_{total} := [L_N]+\sum_{i,*}[C_{i,N}]$, and enzyme, $[P]_{total} := [P]+[D]$, then $[R]_{total} \approx [R]+[R_P]$ holds. At the quasi-steady-state, that is, $d[C_{1,*}]/dt = d[C_{0,*}]/dt = d[P]/dt = 0$ for $* \in \{T, N\}$, the dynamics of R_P can be described by the following ODE with respect to $[R_P]$:

where $V_{1,\mathrm{T}}$:=$[\mathrm{L_T}]_{\mathrm{total}}\alpha_{\mathrm{T}}/(1+\alpha_{\mathrm{T}})$, $V_{1,\mathrm{N}}$:=$[\mathrm{L_N}]_{\mathrm{total}}w\alpha_{\mathrm{N}}/(1+\alpha_{\mathrm{N}})$, $K_{\mathrm{m,T}}$:=$K_{1,\mathrm{T}}/(1+\alpha_{\mathrm{T}})$, $K_{\mathrm{m,N}}$:=$K_{1,\mathrm{N}}/(1+\alpha_{\mathrm{N}})$, $K_{1,\mathrm{T}}$:=$K_{1,\mathrm{N}}/(1+\alpha_{\mathrm{T}})$, $K_{1,\mathrm{N}}$:=$K_{1,\mathrm{N}}/(1+\alpha_{\mathrm{N}})$, K_2:=$(k_{-2}+k_3)/k_2$, α_{T}:=$w/(k_{-1,\mathrm{T}}+w)$, and α_{N}:=$w/(k_{-1,\mathrm{N}}+w)$.

3.2. The reliability of the zero-order proofreading model

To evaluate the performance of the target ligand detection, we introduce the following quantities. First, we introduce the true positive TP and false positive FP by using the partial flux to produce the product molecule $\mathrm{R_P}$ by the target ligand and that by the non-target one, which are defined respectively by

$$TP := V_{1,\mathrm{T}} \frac{[\mathrm{R}]_{\mathrm{total}} - [\mathrm{R_P}]}{K_{\mathrm{m,T}} + [\mathrm{R}]_{\mathrm{total}} - [\mathrm{R_P}]}, \qquad (12)$$

$$FP := V_{1,\mathrm{N}} \frac{[\mathrm{R}]_{\mathrm{total}} - [\mathrm{R_P}]}{K_{\mathrm{m,N}} + [\mathrm{R}]_{\mathrm{total}} - [\mathrm{R_P}]}. \qquad (13)$$

By using TP and FP, we introduce the error η to evaluate the reliability of the target detection, which is defined as

$$\eta := \frac{FP}{TP + FP}. \qquad (14)$$

For a reliable chemical communication, the balance between the accuracy and the intensity of output is necessary. For example, even if the error is low, but the output intensity from the sensory system is quite small, there is a risk that the output signal may disappear before the subsequent reactions are induced inside of the communication system. To evaluate the intensity of output, we introduce the output intensity O, which is defined by

$$O := \frac{[\mathrm{R_P}]^*}{[\mathrm{R}]_{\mathrm{total}}}, \qquad (15)$$

where $[\mathrm{R_P}]^*$ is the steady-state concentration of $[\mathrm{R_P}]$ obtained from Eq. 11 at the steady-state condition $d[\mathrm{R_P}]/dt$=0. Hereafter, by evaluating the error η and the output intensity O, we investigate the performance of the zero-order proofreading model.

To analyze the effect of the absence of non-target ligands, we introduce the fraction of the target ligand[13],

$$r := \frac{[\mathrm{L_T}]_{\mathrm{total}}}{[\mathrm{L_T}]_{\mathrm{total}} + [\mathrm{L_N}]_{\mathrm{total}}}, \qquad (16)$$

and evaluate the dependency of the error η and the output intensity O on r (Fig. 6). Although the error η in the saturated condition (Fig. 6 a) is slightly higher than those of the unsaturated conditions (Fig. 6 b), there is a common tendency that the error η decreases with the increase of the target ligand fraction r. On the other hand, the output intensity O shows a qualitatively different dependency on r. The output intensity O in the saturated condition shows all-or-none response to r (Fig. 6 a). However, the response becomes gradual with a decrease in the saturation level (Fig. 6 b). Consequently, when the target ligand fraction is high (r>1/2), the saturated system (K=1) can balance the accuracy and output intensity (Fig. 6 a).

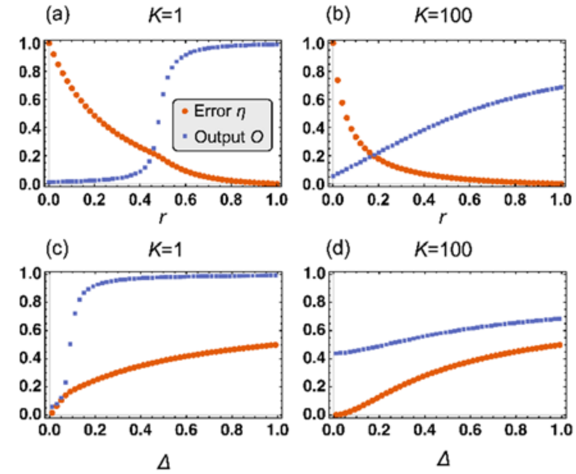


Fig. 6. The error η (red dots) and the output intensity O (blue squares) are plotted as functions of the target ligand fraction r (a, b) and the unbinding rate ratio Δ (c, d) for different values of K, where K=$K_{\mathrm{m,T}}$=K_2. The parameters are $k_{-1,\mathrm{T}}$=1, Δ=1/10 (a, b), r=1/2 (c, d), w=1, k_{-2}=10, k_3=1, $[\mathrm{R}]_{\mathrm{total}}$=100, and $[\mathrm{L}]_{\mathrm{total}}$=5. Note that k_1 and k_2 are obtained from the value of K, $[\mathrm{L}_*]_{\mathrm{total}}$ for $*\in\{\mathrm{T,N}\}$ is obtained from r (a, b), and $k_{-1,\mathrm{N}}$ is obtained from Δ (c, d).

To investigate the effect of similarity between the target and non-target ligands, we introduce the unbinding constant ratio

$$\Delta := \frac{k_{1,\mathrm{T}}}{k_{1,\mathrm{N}}}. \qquad (17)$$

As same as in the Fig. 6 (a, b), the error of saturated condition (Fig. 6 c) is always slightly higher than those of unsaturated condition (Fig. 6 d). However, these two cases have the same tendency that the error decreases when the difference between the target and non-target ligands Δ becomes large. In contrast, the response of the output intensity O to Δ again shows a qualitatively different dependency on the unsaturation level K. When the system is saturated (Fig. 6 c), the response of O to Δ shows an all-or-none response. In addition, in this saturated condition, the system can generate large output O even when Δ is small, which means that the system can generate a large output without the help of the non-target ligand. With an increase in the unsaturation level (Fig. 6 d), the response becomes more gradual, and the system cannot generate large output O in the range where the error η is small (Δ∈[0.0, 0.3] for example). As the results, at the saturated condition (Fig. 6 c), the system can balance the high accuracy and large output O around $\Delta \in [0.1, 0.2]$.

We also investigate the dependencies of the error η and output intensity O on both the target ligand fraction r and the unbinding constant ratio Δ for various saturation levels K (Fig. 7), and confirm that the qualitatively same results in Fig. 6 hold for the various values of K. Together with the results in Figs. 6 and 7, the zero-order proofreading mechanism with saturating condition can balance the high accuracy and large output O at the slight cost of the error η.

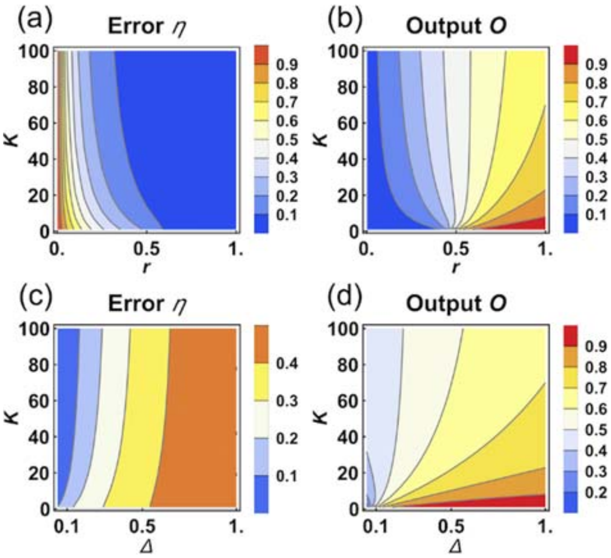


Fig. 7. The error η (a, c) and the output intensity O (b, d) are plotted as functions of the target ligand fraction r and the unsaturation level K (a, b), and the unbinding rate ratio Δ and the unsaturation level K (c, d), where $K=K_{m,T}=K_2$. The parameters are $k_{-1,T}=1$, $w=1$, $k_{-2}=10$, $k_3=1$, $[R]_{total}=100$, and $[L]_{total}=5$. The others are $\Delta=1/10$ for (a, b), and $r=1/2$ for (c, d). Note that k_1, k_2, $[L*]_{total}$ for $* \in \{T,N\}$ (a, b), and $k_{-1,N}$ (c, d) are obtained in the same manner as in Fig. 6.

5. Summary & discussion

In this work, we analyzed the zero-order proofreading model7 and obtain an intuitive understanding of how the mechanism amplifies the unbinding rate from the viewpoint of the balance of the positive and negative fluxes in a saturated condition. We further demonstrated that the zero-order proofreading model is valid for the case that there are the target and non-target ligands in the environment simultaneously, by revealing that the model with saturated conditions can balance both the accuracy and output intensity. The zero-order proofreading mechanism may not only give an insight into how biological cells perform reliable chemical communications in the complex mixtures of ligands, but also provide a method to design bio-inspired reliable chemical communication systems.

Acknowledgements

I would like to thank Tetsuya J. Kobayashi and Kazuyuki Aihara for fruitful discussion. This work was supported by JSPS KAKENHI Grant Numbers 15H05707, and 18K18147.

References

1. P. R. ten Wolde, N. B. Becker, T. E. Ouldridge, and A. Mugler, Fundamental limits to cellular sensing, J. Stat. Phys., 162, (2016) 1395-1424.
2. M. K. Kajita, R. Yokota, K. Aihara, and T. J. Kobayashi, Experimental and theoretical bases for mechanisms of antigen discrimination by T cells, Biophysics, 11, (2015) 85-92.
3. M. J. Taylor, K. Husain, Z. J. Gartner, S. Mayorand, and R. D. Vale, A DNA-based T cell receptor reveals a role for receptor clustering in ligand discrimination, Cell, 169, (2017) 108-119.
4. M. V. Rodnina, and W. Wintermeyer, Ribosome fidelity: tRNA discrimination, proofreading and induced fit, Trends Biochem Sci., 26, (2001) 124-130.
5. S. C. Blanchard, R. L. Gonzalez Jr, H. D. Kim, S. Chu, and J. D. Puglisi, tRNA selection and kinetic proofreading in translation, Nat. Struct. Mol., 11, (2004) 1008.
6. K. W. Ieong, Ü. Uzun, M. Selmer, and M. Ehrenberg, Two proofreading steps amplify the accuracy of genetic code translation, Proc. Natl. Acad. Sci., 113, (2016) 13744-13749.
7. M. K. Kajita, K. Aihara, and T. J. Kobayashi, Balancing specificity, sensitivity, and speed of ligand discrimination by zero-order ultraspecificity, Physical Review E, 96, (2017) 012405.
8. J. J. Hopfield, Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity, Proc. Natl. Acad. Sci., 71, (1974) 4135-4139.
9. T. W. Mckeithan, Kinetic proofreading in T-cell receptor signal transduction, Proc. Natl. Acad. Sci., 92, (1995) 5042-5046.
10. D. A. Beard and H. Qian, Chemical biophysics: quantitative analysis of cellular systems. (Cambridge University Press, 2008).
11. A. Goldbeter and D. E. Koshland, Reducibility among combinatorial problems, Proc. Natl. Acad. Sci., 78, (1981) 6840-6844.

A Reinforcement Learning-Based Path Planning Considering Degree of Observability

Yong Hyeon Cho, Chan Gook Park
Department of Mechanical and Aerospace Engineering, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul, 08826 Republic of Korea
E-mail: 201929291@snu.ac.kr, chanpark@snu.ac.kr

Abstract

This paper presents a novel way to find a path using a degree of observability as a reward for the reinforcement learning in the INS/GNSS loosely coupled system. In the proposed algorithm, an agent follows the Dubin's car model in a grid map, using a degree of observability to update Q-value. Various ratios between the penalty and reward at each step show different trajectories, and the degree of observability at the endpoint of a grid map is compared with one another. The progress is shown by computer simulation.

Keywords: Degree of Observability, Reinforcement Learning, Q-Learning, INS/GNSS integrated navigation

1. Introduction

Given several measurements from GPS, the system's states can be estimated in the INS/GNSS integrated navigation system. Loosely coupled (LC) navigation system typically estimates the system's 15 states, including position, velocity, attitude, bias of accelerometer and gyroscope. The observability of the system is a crucial factor in determining whether a specific state can be estimated correctly. And for the case of piece-wise constant systems, it is shown that the observability of error states can be determined by testing the rank of stripped observability matrix.[1] This method can only tell whether the system is observable or not, and don't specify which states are more or less observable. Therefore, the concept of the degree of observability was introduced by Ham and Brown[2], who considered eigenvalues of the covariance matrix. It is shown that the estimation results of the system's specific state is more uncertain if its eigenvalue is bigger than that of other states.

The system's observability can be affected by several factors, and the navigator's maneuver is one of them.[3] Therefore, one can expect the better observability of the specific state if the trajectory of agents is planned in a specific way. Meanwhile, reinforcement learning can find an optimal path by using dynamic programming, and there are several ways to utilize it for path planning.[4,5] Q-learning is one of the traditional ways to teach agents in the model-free environment. The environment's policy can be determined by setting values of learning rate and gamma, which acts as the main factor to make a solution converge. This paper aims to find the trajectory that makes a specific state more observable by using reinforcement learning.

In this paper, a novel way of planning a specific trajectory is introduced. The Q-learning algorithm is modified by adding a degree of observability as an instant reward. The paper is organized as follows. Section II explains preliminary requirements that can make the degree of observability as a valid parameter for the instant reward. Section III discusses the observability of a yaw state, comparing results of the new algorithm by the ratio of reward and penalty. Conclusions and discussions are made in Section VI.

2. System Environment

2.1. *Environment model for the Dubin's car*

Usually, INS/GNSS integrated navigation system estimates states of the navigator following Dubin's car model, like airplanes and vehicles. But the critical point is that the agent never search paths in a way that a car or an airplane moves in the early stage of Q-learning. It just merely searches for the highest Q-value across the states, zig-zagging and moving back and forth through the entire process. In that case, it is not reasonable to take a degree of observability of certain states as an instant reward, as the agent doesn't move like a real car. This problem can be solved by changing an original algorithm as follows.

$s := currentstate()$
Choose a based on Q
Take action a, observe r, s'
$Q(s,a) := Q(s,a) + \alpha[R(s) + \gamma \max_{a'} Q(s',a') - Q(s,a)]$
$\mathrm{s} := \mathrm{s}$

↓

$s := currentstate()$
Repeat
Choose a(action) based on Q
Exclude a that causes recurring movement
Take action a, observe r, s'
$Q(s,a) := Q(s,a) + \alpha[R(s) + \gamma \max_{a'} Q(s',a') - Q(s,a)]$
$\mathrm{s} := \mathrm{s}$

Fig. 1. Modified Q-Learning Algorithm

As the agent moves along a grid map, every step it moves from one grid cell to another cell is regarded as one of these three movements under a constant speed: Going straight, turning left, turning right. For example, let us consider a following 5x5 grid map like Fig. 2., setting upper-left as a coordinate (1,1) and down-right as a coordinate (5,5). When it moves in series of (4,3) - (3,3) - (3,4), it means that the agent is currently turning right, eventually change its heading from north to east. And for the agent at a coordinate (3,4), it cannot go back to the (3,3) at that time, because its heading is pointing to the east direction, and going back to (3,3) is impossible maneuvering for the Dubin's car. Also, when it moves in series of (5,2) - (4,2), it means that the agent is going straight, heading north.

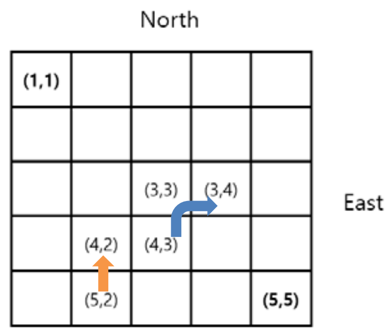


Fig. 2. 5x5 grid map

The agent navigates through each grid cell for 10 seconds, maintaining a constant speed of 10m/s. The entire process is under the control of loosely coupled system, using the agent's position as GPS measurements. The degree of observability can be calculated by deriving eigenvalues of the state's covariance matrix. When the agent goes through each grid cell, there will be a list of different eigenvlaues as time passes by. However, the eigenvalue at the last point of the time is used so that the difference of eigenvalues between a cell to another is utilized as a reward for Q-learning.

$s := currentstate()$
Repeat
Choose a(action) based on Q
Exclude a that causes recurring movement
Take action a, observe r, s'
$Q(s,a) := Q(s,a) + \alpha[\mathrm{DObs} + M + \gamma \max_{a'} Q(s',a') - Q(s,a)]$
(DObs: *Degree of Observability as a reward*
M : *Penalty by moving*)
$\mathrm{s} := \mathrm{s}$

Fig. 3. Modified Q-Learning Algorithm – Setting the Degree of Observability as a Reward

2.2. *Applying the proposed algorithm*

For the application of the propsed Q-learning algorithm in the 5x5 grid map, the process begins from the coordinate (5,1) until it reaches the coordinate (1,5). The degree of observability of a yaw state which is driven as a normalized form[2] is used for the algorithm, and the reward for improving the observability is set to 1.5. The penalty for moving from a cell to another is set to -1. The result of path planning, which is considering a yaw state's degree of observability, is shown below.

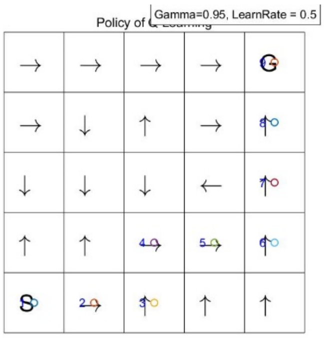


Fig. 4. Result of a proposed algorithm(S:start, G:goal)

3. Additional Simulation results

3.1. *Case I : Varying observability reward*

In the previous result, the reward for improving the observability was set to 1.5. In addition to the result, simulations are performed by varying the reward value.

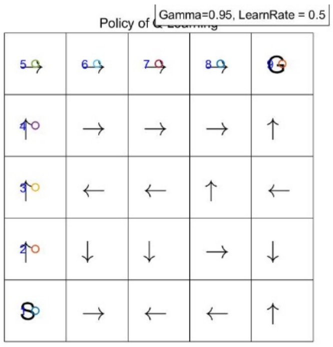


Fig. 5. Case I - 1(Reward: 1, Penalty: -1)

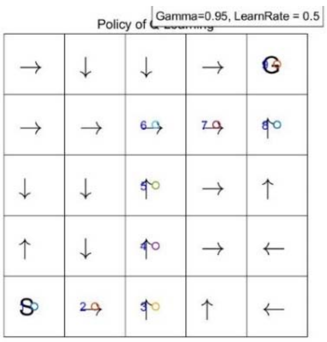


Fig. 6. Case I - 2(Reward: 2.5, Penalty: -1)

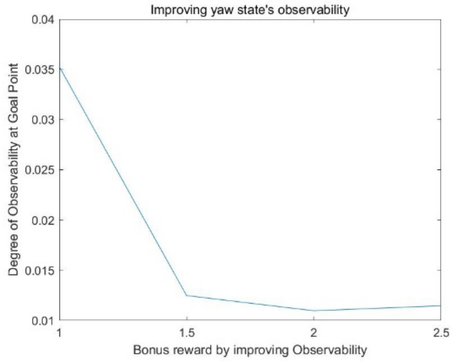


Fig. 7. Comparison of the yaw state's observability at the goal point with varying reward.

The process tend to show decreased eigenvalues (which menas increasing of the degree of observability) as the reward for improving observability increases.

3.2. *Case II : Varying movement penalty*

In this case, the reward is set to 1, and the penalty for moving through other cells varies from -0.7 to -0.3. Results show that as penalty decreases, agents tend to search longer path to get better eigenvalues at the goal point.

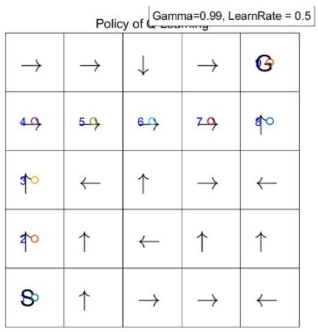


Fig. 8. Case II – 1 (Reward: 1, Penalty: -0.5)

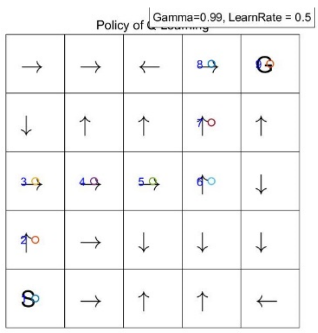


Fig. 10. Case II – 2 (Reward: 1, Penalty: -0.4)

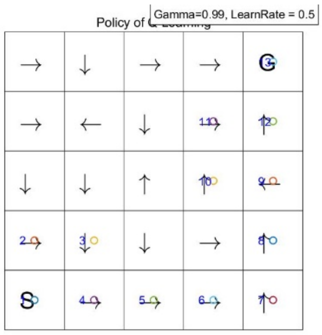


Fig. 9. Case II – 3 (Reward: 1, Penalty: -0.3)

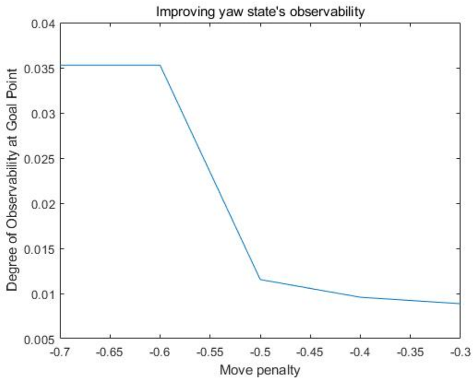


Fig. 10. Comparison of the yaw state's observability at the goal point with varying penalty

Eigenvalues tend to decrease as the penalty for moving thorugh gird cell decreases, and in Fig. 9., the agent even takes further route to reach the goal point.

4. Conclusion

In this paper, the modification of the Q-learning algorithm was made by 1) changing the searching method of Q-value to the Dubin's car model and by 2) adding the difference of the degree of observability as an instant reward. The improvement of the observability of a yaw state was compared by making differences of ratio between the penalty and reward.

There are other factors that can affect the observability of states like acceleration, pitch, roll maneuvers, but these factors make things too complicate to be considered in the view of grid map. Although there still remain several questions (the optimized ratio between reward and penalty, the maximum degree of observability that system can reach in this algorithm, etc.), the observability applied in Q learning showed several 'desirable trajectories' that have an enhancement in the degree of observability at goal point. Further study is required to optimize the observability thorough the process.

Acknowledgements

This work was supported by the R&D program of MOTIE/KEIT. [10084665, Development of IMU Embedded 6-axis, 10-axis compound navigation system integrating highly reliable IMU, GNSS, Magnetometer and altimeter for manned/unmanned aircraft]

References

1. D. Goshen-Meskin and I. Y. Bar-Itzhack, Observability analysis of piece-wise constant systems - Part I: theory, in *IEEE Transaction on Aerospace Electronic Systems*, **28**(4) (1992) pp. 1056-1067
2. F. M. Ham and R. G. Brown, Observability, eigenvalues, and Kalman filtering, in *IEEE Transaction on Aerospace Electronic Systems*, **AES-19**(2) (1983) pp. 269-273
3. Y. M. Yoo, J. G. Park, D. H. Lee, and C. G. Park, A theoretical approach to observability analysis of the SDINS/GPS in maneuvering with horizontal constant velocity, in *International Journal of Control, Automation, and Systems,* **10**(2) (2012) pp. 298-307
4. Y. Song, Y. B. Li, C. H. Li, and G. F. Zhang, An efficient initialization approach of Q-learning for mobile robots, in *International Journal of Control, Automation, and Systems*, **10**(1) (2012) pp. 166-172
5. S. CHENG'EN, and H. Jun, A hybrid path planning algorithm for indoor mobile robot using hierarchy reinforcement learning, in *International Journal of Control, Automation, and Systems*, **9**(5) (2016) 351–362

A Performance Analysis of Pose Estimation Based on Two-View Tracking and Multi-State Constraint Kalman Filter Fusion

Tae Ihn Kim, Jae Hyung Jung, Chan Gook Park
Department of Mechanical and Aerospace Engineering, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul, 08826, Republic of Korea
E-mail: natigerya@snu.ac.kr, lastflowers@snu.ac.kr, chanpark@snu.ac.kr

Abstract

This paper presents a performance analysis of two-view tracking and Multi-State Constraint Kalman Filter (MSCKF) fusion for a pose estimation. The system and measurement model of both two-view tracking and MSCKF are derived based on the fusion condition. The simulation result of the fused algorithm using the Drone Racing dataset, collected from an aggressive flight of micro aerial vehicle (MAV), shows the performance improvement of both attitude and position estimation compared to the performance of MSCKF.

Keywords: Visual-Inertial Odometry, Two-View Tracking, Multi-State Constraint Kalman Filter, Sensor Fusion

1. Introduction

Visual-Inertial Odometry (VIO) is an algorithm for a pose estimation using images from a camera sensor and linear accelerations and angular velocities from an Inertial Measurement Unit (IMU). The pose estimation algorithms using each sensor are known as Visual Odometry (VO) and Inertial Navigation System (INS). Both VO and INS suffer from drift over time because the pose is estimated incrementally in both algorithms.[1] While the pose estimation of VO results in high precision in a slow motion, that of INS results in high accuracy in a rapid motion. Therefore, VIO, which is the fusion of VO and INS, complements the strengths of each sensor and improves the performance of pose estimation.

Among various VIO algorithms, Multi-State Constraint Kalman Filter (MSKCF) is chosen as the main algorithm. Different from other general VIO algorithms such as Extended Kalman Filter (EKF) based algorithms, MSCKF does not include three-dimensional feature positions in the filter states, which result in a drop of computation complexity. MSCKF uses a geometric constraint obtained from the poses of previous camera frames included in a sliding window of filter state to estimate the pose.[2] However, the performance of MSCKF pose estimation is comparatively inaccurate in a rapid motion. To overcome this weakness, a fused algorithm of two-view tracking and MSCKF is proposed. Two-view tracking uses an optical flow measurement obtained from consecutive image frames in the measurement update. The strength of two-view tracking is that the pose estimation is comparatively accurate even in a rapid motion. Therefore, the fusion of two algorithms compensates for the weakness of MSCKF.

This paper presents the performance analysis of two-view tracking and MSCKF fusion for pose estimation throughout an application to the Drone Racing dataset, which images, linear accelerations and angular velocities are collected in a rapid motion.[3]

2. System Model

The error state of the fused algorithm is a combination of two-view tracking and MSCKF error state. The error state of two-view tracking is described as Eq.(1).[4]

$$\tilde{X}_{TV} = \begin{bmatrix} \tilde{p}_{GB}^{G\,T} & \tilde{v}_{GB}^{B\,T} & \tilde{\theta}_{GB}^{\;T} & \tilde{b}_a^{\;T} & \tilde{b}_g^{\;T} & \tilde{\alpha}^T \end{bmatrix}^T \quad (1)$$

where $\tilde{p}_{GB}^{G}$ is a position error expressed in a global-frame, $\tilde{v}_{GB}^{B}$ is a velocity error expressed in a body-frame, $\tilde{\theta}_{GB}$ is an attitude error expressed in a global-frame, $\tilde{b}_a$ is an accelerometer bias error, $\tilde{b}_g$ is a gyroscope bias error and $\tilde{\alpha}$ is an inverse-scene-depth error. The error state of MSCKF system model is described as Eq.(2).[5]

$$\tilde{X}_{MSCKF} = \begin{bmatrix} \tilde{X}_{IMU}{}^{T} & \tilde{X}_{slw}{}^{T} \end{bmatrix}^{T} \tag{2}$$

where

$$\tilde{X}_{IMU} = \begin{bmatrix} \tilde{p}_{GB}^{G}{}^{T} & \tilde{v}_{GB}^{G}{}^{T} & \tilde{\theta}_{GB}{}^{T} & \tilde{b}_a{}^{T} & \tilde{b}_g{}^{T} \end{bmatrix}^{T} \tag{3}$$

$$\tilde{X}_{slw} = \begin{bmatrix} \tilde{p}_{GC_1}^{G}{}^{T} & \tilde{\theta}_{GC_1}{}^{T} \vdots \ \dots \ \vdots \tilde{p}_{GC_N}^{G}{}^{T} & \tilde{\theta}_{GC_N}{}^{T} \end{bmatrix}^{T} \tag{4}$$

$\tilde{X}_{IMU}$ is IMU error state, described as Eq.(3), $\tilde{X}_{slw}$ is temporarily included previous camera pose error state in the sliding window, described as Eq.(4), $\tilde{v}_{GB}^{G}$ is a velocity error expressed in a global-frame, $\tilde{p}_{GC_N}^{G}$ and $\tilde{\theta}_{GC_N}$ are N-th camera position error and attitude error in the sliding window. The error state of the fused algorithm is described as Eq.(5). For the convenience of Jacobian derivation, $\tilde{v}_{GB}^{B}$ is chosen instead of $\tilde{v}_{GB}^{G}$.

$$\tilde{X}_{fused} = \begin{bmatrix} \tilde{X}_{TV}{}^{T} & \tilde{X}_{slw}{}^{T} \end{bmatrix}^{T} \tag{5}$$

The system model is described as Eq.(6).

$$\begin{aligned}
\dot{\tilde{p}}_{GB}^{G} &= \hat{R}_B^G \tilde{v}_{GB}^{B} - \lfloor \hat{R}_B^G \hat{v}_{GB}^{B} \times \rfloor \tilde{\theta}_{GB} \\
\dot{\tilde{v}}_{GB}^{B} &= \hat{R}_G^B \lfloor g^G \times \rfloor \tilde{\theta}_{GB} - \lfloor (\omega_m - \hat{b}_g) \times \rfloor \tilde{v}_{GB}^{B} \\
&\quad - \tilde{b}_a - \lfloor \hat{v}_{GB}^{B} \times \rfloor \tilde{b}_g - n_a - \lfloor \hat{v}_{GB}^{B} \times \rfloor n_g \\
\dot{\tilde{\theta}}_{GB} &= -\hat{R}_B^G \tilde{b}_g - \hat{R}_B^G \tilde{n}_g \\
\dot{\tilde{b}}_a &= n_{wa} \\
\dot{\tilde{b}}_g &= n_{wg} \\
\dot{\tilde{\alpha}} &= n_\alpha
\end{aligned} \tag{6}$$

where n_a and n_g are zero-mean, white Gaussian noise of accelerometer and gyroscope, respectively, and n_{wa} and n_{wg} are random walk rate of accelerometer and gyroscope, respectively.

3. Measurement Model

The measurement model of the fused algorithm is divided into two-view tracking and MSCKF measurement model. The measurement update related term of two-view tracking is described as Eq.(7).

$$\begin{aligned}
\tilde{y}_i = M_i \{ &\hat{\alpha} R_B^C \tilde{v}_{GB}^{B} + (\lfloor \bar{P}_{Cf_i}^{C} \times \rfloor R_B^C + \hat{\alpha} R_B^C \lfloor P_{BC}^{B} \times \rfloor) \tilde{b}_g \\
&R_B^C (\hat{v}_{GB}^{B} + \lfloor (\omega_m - \hat{b}_g) \times \rfloor P_{BC}^{B}) \tilde{\alpha} \} + \tilde{n}_{v_i}
\end{aligned} \tag{7}$$

where, $\tilde{y}_i$ is a two-dimensional innovation term, M_i is a nullspace reprojection matrix and ω_m is a gyroscope measurement. The measurement update related term of MSCKF is described as Eq.(8).[5]

$$r_i^{(j)} = z_i^{(j)} - \hat{z}_i^{(j)} \cong H_{X_i}^{(j)} \tilde{X} + H_{f_i}^{(j)} \tilde{p}_{Gf_i}^{G} + n_i^{(j)} \tag{8}$$

$$r^{(j)} \cong H_X^{(j)} \tilde{X} + H_f^{(j)} \tilde{p}_{Gf}^{G} + n^{(j)} \tag{9}$$

$$r_0^{(j)} = A^T (z^{(j)} - \hat{z}^{(j)}) \cong A^T H_X^{(j)} \tilde{X}^{(j)} + A^T n^{(j)} \tag{10}$$

$$= H_0^{(j)} \tilde{X}^{(j)} + n_0^{(j)} \tag{11}$$

$$r_0 = H_X \tilde{X} + n_0 \tag{12}$$

where $r_i^{(j)}$ is a measurement residual, $z_i^{(j)}$ is a measurement, $\hat{z}_i^{(j)}$ is a reference, $H_{X_i}^{(j)}$ and $H_{f_i}^{(j)}$ are state and feature position Jacobians of $z_i^{(j)}$, respectively, $\tilde{X}$ is the error state, $\tilde{p}_{Gf_i}^{G}$ is a feature position error and $n_i^{(j)}$ is a noise vector. Residuals for each previous camera frames in Eq.(8) are stacked up to form Eq.(9) and reprojected on the left nullspace to form Eq.(10) and Eq.(11). Residuals for each feature are finally stacked up and H_X is used in the EKF update.

The keypoint of the fused algorithm is that both two-view tracking and MSCKF measurement updates proceed during the pose estimation. MSCKF proceeds the measurement update when one of the two conditions satisfies. Those two conditions are a failure of feature tracking and an excess of the size of the sliding window. However, the minimum number of tracks is also assigned in order to avoid an error in residuals. The fused algorithm follows the same measurement trigger as MSCKF, but two-view tracking measurement update proceeds when there are exactly two tracks in the sliding window. For example, as shown in Fig. 1., when N equals two, the optical flow is measured from C_N and C_{N-1} camera frames included in the sliding window.

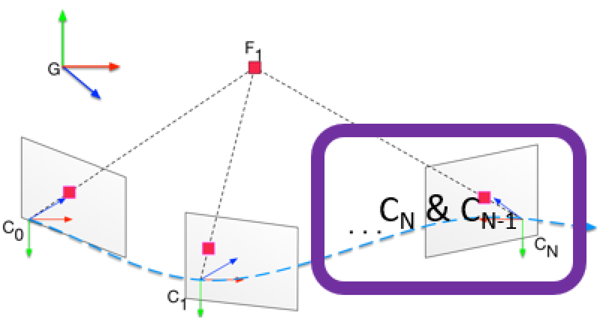


Fig. 1. Two-View Tracking Frames in the Sliding Window

4. Performance Analysis

The performance analysis of the fused algorithm is proceeded using the Drone Racing dataset, which provides images, accelerometer and gyroscope measurements and ground truth collected using a laser tracking system.[3] The total distance traveled by MAV in the dataset is 270.7448 m and the maximum instantaneous velocity is 25.0499 m/s. Other drone flight datasets such as EuRoC and Zurich Urban MAV dataset exist, but the Drone Racing is chosen because only this dataset provides measurements collected by an aggressive drone flight. The performance difference between MSCKF and the fused algorithm is expected to be clearly observable in bad condition in terms of the visual environment.

For the performance analysis, the altitude error and the position error of the fused algorithm are compared to those of MSCKF. The result is shown in Table 1, Fig.2 and Fig.3. The attitude error has improved from 0.2670 rad to 0.2306 rad, and the position error has improved from 6.9418 m to 6.4576 m, which are 13.6 % and 7.0 % improvement, respectively.

Table 1. Root Mean Square Error (RMSE)

	MSCKF	Fused
Attitude [rad]	0.2670	0.2306
position [m]	6.9418	6.4576

The improvement of both attitude and position error is a result of the difference in the EKF update. In MSCKF, the measurement update does not proceed when the number of tracks in the sliding window is less than three. However, in the fused algorithm, the measurement update proceeds even when there are two tracks in the sliding window and two-view measurement update proceeds instead of MSCKF measurement update. Compared to other drone flight datasets, the Drone Racing dataset provides lower sized images, which fewer features are detected. MSCKF measurement update using a few features results in low accuracy. However, in the same condition, two-view tracking measurement update results in more accurate pose estimation compared to that of MSCKF. Therefore, the additional two-view tracking measurement update in MSCKF measurement update improves the performance of pose estimation of MAV in an aggressive flight condition.

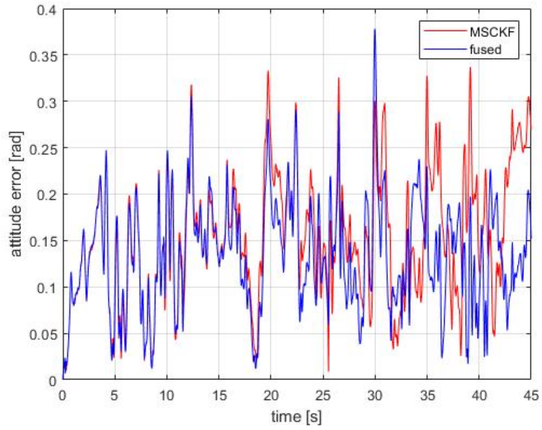


Fig. 2. Three-dimensional Attitude Error of Drone Racing

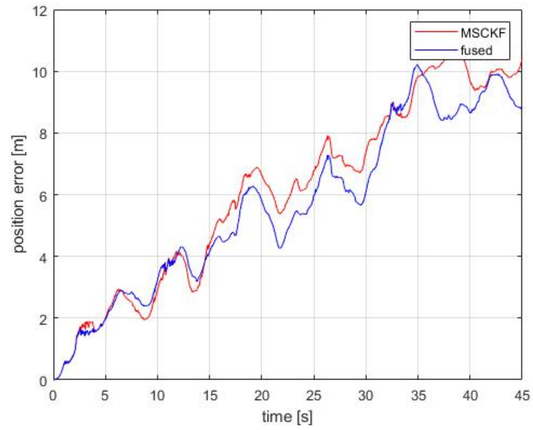


Fig. 3. Three-dimensional Position Error of Drone Racing

5. Conclusion

This paper presents the performance analysis of the fusion of two-view tracking and MSCKF for the pose estimation. The error state of the fused algorithm is selected to be the combination of two-view tracking and MSCKF error state. The system model and the measurement update related terms are newly derived since the velocity error is expressed in a body-frame. MSCKF measurement update proceeds when one of two conditions is triggered and two-view tracking measurement update proceeds when there are two tracks in the sliding window. The fused algorithm results in an improvement of the pose estimation.

Acknowledgements

This work was supported by the Ministry of Science and ICT of the Republic of Korea through the Space Core Technology Development Program under Project NRF-2018M1A3A3A02065722.

References

1. J. H. Jung, S. Heo and C. G. Park, Patch-based Stereo Direct Visual Odometery Robust to Illumintation Changes, *International Journal of Control, Automation and System*, 17(3), (2019), 743-751.
2. P. Kim, H. Lim and H. J. Kim, Visual Inertial Odometry with Pentafocal Geometric Constraints, *International Journal of Control, Automation and System*, 16(4), (2018), 1962-1970.
3. M. Bloesch, S. Omari, P. Fankhauser, H. Sommer, C. Gehring, J. Hwangbo, M. A. Hoepflinger, M. Hutter and R. Siegwart, Fusion of Optical Flow and Inertial Measurements for Robust Egomotion Estimation, *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, (2014), 3102-3107.
4. J. Delmerico, T. Cieslewski, H. Rebecq, M. Faessler and D. Scaramuzza, Are We Ready for Autonomous Drone Racing? The UZH-FPV Drone Racing Dataset, *IEEE International Conference on Robotics and Automation*, (2019).
5. A. I. Mourikis and S. I. Roumeliotis, A Multi-State Constraint Kalman Filter for Vision-aided Inertial Navigation, *Proceedings 2007 IEEE International Conference on Robotics and Automation* (2007), 3565-3572.

ORB-SLAM based Sensor Fusion Algorithm for Real-Time Precision Driving

Yong-Jin Ock
Dep. of Electronic Eng., Pusan National University, 2, Busandaehak-ro 63beon-gil Geumjeong-gu Busan, 46241, Republic of Korea

Zhan-Ming Gu
Dep. of Electronic Eng., Pusan National University, 2, Busandaehak-ro 63beon-gil Geumjeong-gu Busan, 46241, Republic of Korea

Jang-Myoung Lee
Dep. of Electronic Eng., Pusan National University, 2, Busandaehak-ro 63beon-gil Geumjeong-gu Busan, 46241, Republic of Korea

E-mail: yongjin7379@pusan.ac.kr, guzhanming7379@pusan.ac.kr, jmlee@pusan.ac.kr
www.pusan.ac.kr

Abstract

In this paper, we propose a position correction method through SLAM (Simultaneous Localization And Mapping)-based sensor fusion for precise driving in the indoor. There was a problem that it was not possible to determine the exact posture and position with a single image alone. To compensate for this, additional IMU (Inertial Measurement Unit) sensor and encoder sensor should be installed and calibrated. At this time, the encoder sensor acquires information about the distance traveled and the attitude of the mobile robot. The IMU sensor measures the attitude error caused by the sliding and friction of the mobile robot and acquires the slope information of the current terrain. As a result, by combining the location information acquired by using the SLAM and the complex location information of the IMU sensor and the encoder sensor, precise position control is possible even in a space without many feature points.

Keywords: SLAM, Indoor Localization, Mobile Robot

1. Introduction

Recently, with the development of various image processing-related technologies, SLAM technology is in the spotlight as an algorithm for autonomous driving, which is a key keyword of the fourth industrial revolution. SLAM is a technology that maps the surrounding environment and estimates its relative position simultaneously based on sensors attached to the robot. Sensors used are Mono Camera, Stereo Camera, RGB-D Camera, IMU, LiDAR (Light Imaging Detection and Ranging), etc., and are used in various spaces such as AGV (Automated Guided Vehicle) in factory environments, vehicles and drones in outdoor environments, and service robots in indoor environments. Becomes Various algorithms [1-3] have been studied according to the number of cameras and the sensor used, and the advantage of using LiDAR, which enables precise environmental measurement, has the advantage of high algorithm accuracy, but it is difficult to commercialize due to the high price of LiDAR. Has a big disadvantage. Therefore, the need for an alternative technology to replace the LiDAR-based SLAM is highlighted. Therefore, through this research, IMU and Encoder intend to develop algorithm that can measure and control precise position in the environment where there are not many feature points.

2. Mobile Robot Location Estimation

Dead Reckoning is a method of obtaining the position and direction of a mobile robot using only the amount of wheel movement. If there is no sliding of the mobile robot and there are no structural errors such as wheel size and rotation angle, dead reckoning can be used to estimate the position of the robot. Due to the error of the wheel, the sliding of the wheel during movement, and the structural error of the mobile robot, it is difficult to estimate the exact position only by the odometry information through the encoder. Because of this, the odometry information should be corrected and used. At this time, the correction acquires the posture information (Roll, Pitch, Yaw) of the mobile robot by using the IMU sensor and corrects the posture error of the mobile robot based on this.

2.1. *Position Estimation Using Encoder Values*

The robot used in the paper is shown in Fig. 1. The driving method of the mobile robot is driven without a separate steering device and determines the moving direction of the robot by the sum of the forces of four wheels, respectively. In order to calculate the forward and angular velocities of the mobile robot, each wheel rotation must be measured. In order to measure the angular velocity of the robot, each motor was equipped with an incremental encoder, and the angular velocity was measured by calculating the RPM value as a pulse is input every time the motor rotates. The robot is equipped with four encoders and uses the average of the left and right encoders.

$$E_L = \frac{E_{FL} + E_{BL}}{2} \qquad E_R = \frac{E_{FR} + E_{BR}}{2} \tag{1}$$

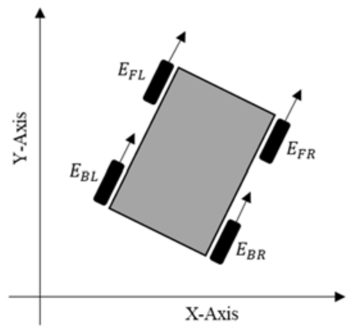


Fig. 1. 4-Wheel mobile robot.

Where E_L and E_R are the average of the left and right encoder rotations. And the velocity and angular velocity can be obtained by using the encoder rotation values obtained above.

$$v = \frac{r(\Delta E_R + \Delta E_L)}{2} \qquad \omega = \frac{r(\Delta E_R - \Delta E_L)}{d} \tag{2}$$

Where r is the diameter of the wheel and d is the width between the left and right wheels. For the time [t_i, t_{i+1}], when the robot is given a constant speed input v, ω and knows p_i as the position of the robot at the current time, the method to calculate the position p_{i+1} of the robot at the next time is as follows.

$$\begin{aligned} x_{i+1} &= x_i + v_i \Delta t \cos\theta_i \\ y_{i+1} &= y_i + v_i \Delta t \sin\theta_i \\ \theta_{i+1} &= \theta_i + \omega_i \Delta t \\ \Delta t &= t_{i+1} - t_i \end{aligned} \tag{3}$$

Eq. (3) can be used to find the position p_{i+1} at time t_{i+1}.

2.2. *IMU-based pose measurement*

Fig. 2. shows the IMU sensor coordinate system used in the paper. The gyroscope sensor needs an integration process to find the rotation angle. At this time, an integration error occurs, and as the number of rotation angle calculations increases, rotation angle drift due to error accumulation occurs. The accelerometer sensor extracts abnormal values for fast rotation and fast

direction changes. However, instead of these drawbacks, the gyroscope showed a relatively stable value change during the movement, and the accelerometer has a characteristic that maintains a constant value at the beginning and the end, so the merits of the two should be fused through the complementary filter.

Accelerometer is a sensor that shows the derivative value of velocity for unit time in the linear direction. It should be converted into Roll and Pitch angle values as follows.

$$\rho = \arctan\left(\frac{A_x}{\sqrt{A_y^2 + A_z^2}}\right)$$
$$\phi = \arctan\left(\frac{A_y}{\sqrt{A_x^2 + A_z^2}}\right) \quad (4)$$

The accelerometer passes the low pass filter and the gyro sensor passes the high pass filter and adds up the result to get the corrected angle. Fig. 3. shows the complementary filter.

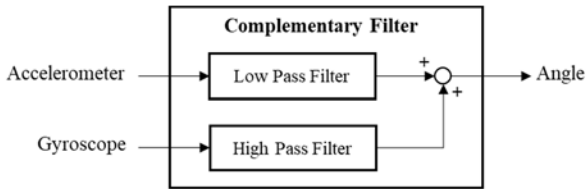


Fig. 3. Block diagram of the complementary filter.

The corrected angle value in Fig. 3. can be expressed as the following equation.

$$\rho_i = \alpha * (\rho_{i-1} + gyro\_x * dt) + (1-\alpha) * (acc\_x)$$
$$\phi_i = \alpha * (\phi_{i-1} + gyro\_y * dt) + (1-\alpha) * (acc\_y) \quad (4)$$

α is the filter constant, ρ_{i-1} is the previous roll angle and ϕ_{i-1} is the previous pitch angle. $gyro$ is the angular velocity measured by the gyro sensor, dt is the sampling time of the gyro sensor, and acc is the x-axis measurement of the acceleration sensor.

Yaw angle can be obtained by using a geomagnetic field sensor, and correction according to tilt change should be applied. The equation can be expressed as follows.

$$X' = X\cos(\phi) + Y\sin(\rho) - Z\cos(\rho)\sin(\phi)$$
$$Y' = Y\cos(\rho) + Z\sin(\phi) \quad (5)$$
$$\theta = \arctan\left(\frac{Y'}{X'}\right)$$

3. ORB-SLAM Calibration

Fig. 4. shows the flow chart for the ORB-SLAM [4-5] calibration. According to the procedure, the incremental encoder is used before correction to collect information on the moving distance and the current driving direction by changing the encoder value according to the movement. And using the IMU sensor to acquire the attitude information (Roll, Pitch, Yaw) of the mobile robot. Finally, the posture information and the position information in the image acquired by ORB-SLAM are compared with the information obtained from each sensor and corrected to improve the accuracy.

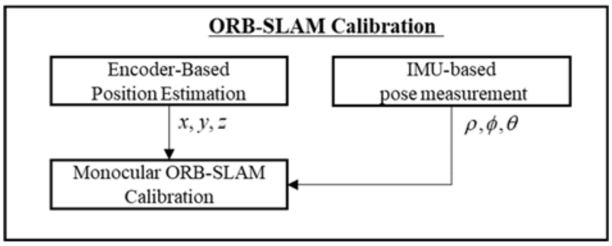


Fig. 4. Block diagram of the ORB-SLAM calibration

4. Experiment

Experiments were performed to verify the ORB-SLAM calibration method using the Encoder and IMU sensors. In this experiment, the camera uses a webcam with 1280 * 720 resolution for the robot's position estimation, and is equipped with four encoders, an IMU including an acceleration sensor, a gyro sensor, and a geomagnetic sensor. Fig. 5. shows the mobile robot used in the actual experiment. In the mobile robot, sensor values and images acquired in real time are transmitted to the host PC using TCP / IP communication based wireless communication, and the host PC estimates the position of the mobile robot with improved accuracy by correcting the ORB-SLAM based on the acquired information.

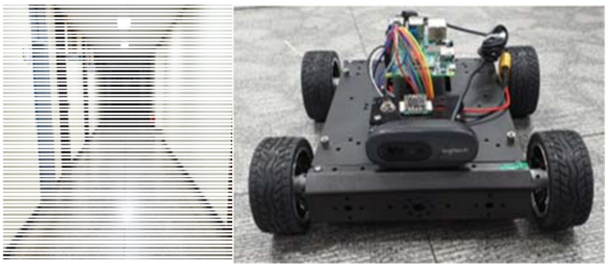


Fig.5. Experiment Environment and Mobile Robot

Fig. 6. is a graph comparing the difference between before and after applying the algorithm when driving indoors at 9m width and 14m height.

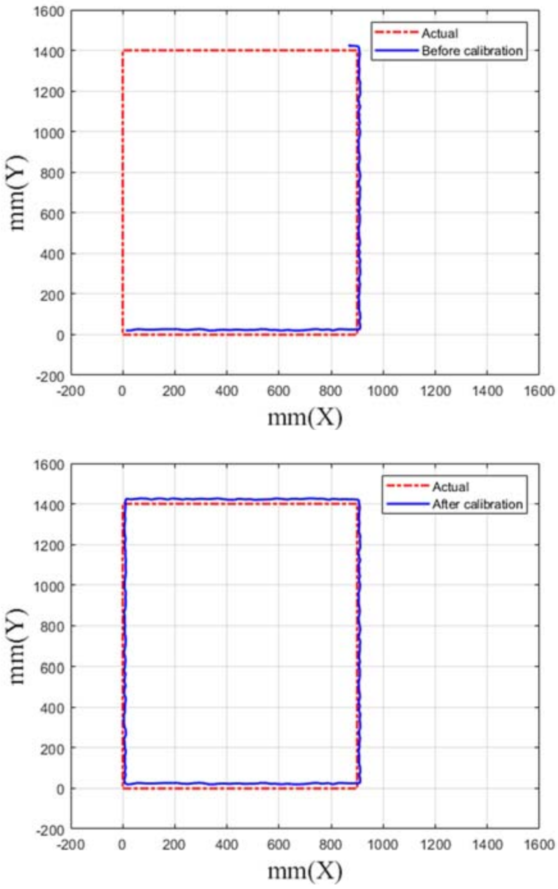


Fig.6. Results before and after calibration

5. Conclusion

In this paper, we proposed an algorithm to correct the position by acquiring the position and posture information of encoder and IMU sensor to solve the problem of position error when driving indoors with less feature points when using Monocular ORB-SLAM. As a result of the experiment, it was able to recognize the path not recognized by the existing algorithm and create a path similar to the experimental environment.

References

1. Guivant, Jose E., and Eduardo Mario Nebot, Optimization of the simultaneous localization and map-building algorithm for real-time implementation, *IEEE transactions on robotics and automation*, Vol. 17, No. 3 pp.242-257, 2001.
2. J.A. Castellanos, R. Martinez-Cantin, J.D. Tardós, J. Neira, Robocentric map joining: Improving the consistency of EKF-SLAM, *Robotics and Autonomous Systems*, Vol. 17, No. 1, pp.21-29, 2007.
3. Engel, J., Schöps, T., & Cremers, D., LSD-SLAM: Large-scale direct monocular SLAM, *European conference on computer vision*, pp.834-849, 2014.
4. Mur-Artal, R., Montiel, J. M. M., & Tardos, J. D., ORB-SLAM: a versatile and accurate monocular SLAM system, *IEEE transactions on robotics*, Vol. 31, No. 5, pp.1147-1163, 2015.
5. Mur-Artal, R., & Tardós, J. D., Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras, *IEEE Transactions on Robotics*, Vol. 33, No. 5, pp.1255-1262, 2017.

Gait Control of A Four-legged Robot with Fuzzy-PID Controller

Arphakorn Kunha, Amornphun Phunopas, Wisanu Jitviriya
Faculty of Engineering, Production Engineering King Mongkut's University of Technology North Bangkok
Bang Sue, Bangkok, 10800, Thailand
s6101021810566@email.kmutnb.ac.th, amornphun.p@eng.kmutnb.ac.th, wisanu.j@eng.kmutnb.ac.th

Abstract

The mobile robot has developed as the legged mobile robot. Which has a high locomotion performance on smooth and non-smooth surface Meanwhile, the legged mobile robot movement should be have algorithm for control on the legged. Because the legged robot have required self-balance and gait that imitated from the locomotion behavior of four-legged animal, such as dog cat horse, etc. Therefore, this paper aims to describes the control system design of the four-legged robot through the PID that it is a controlling of gait. The fuzzy model has used as a determined a range of error in system. But the experiment has emphasized on walking-trot in the gait of animal

Keywords: Gait Design, PID Controller, Quadruped Robot, Fuzzy Logic Controller.

1. Introduction

The legged robot has type of control different. Based on the characteristics animals have using the leg for movement. This legged robot has the ability to possess greater mobility flexibility and stair climbing when compared to a wheel type [2]. In addition, Legged locomotion is complicated and also all legs must be inextricably connected [1]. If focusing on four-legged animals will analyze the step of gait with a mathematical model with the kinematics equation. A four-legged animal has different gait patterns follow the locomotion of animals [4]. Once considered, many factors affect the movement of the legged especially the weight, structure and feet touching that continuous movement.

Nevertheless, for gait pattern can configure the angle and position control range of the pivot point of the legs with walking step by step. To be used to set the posture of the four-legged robot such as standing, sitting, crouch, etc. However, the method simple gait controlling the position any joint of legs by trial and error. However, the mathematical model is still an essential part of creating a control designed. The robot can move according to the selected gait pattern. The Control system is widespread and applied in a variety of applications, such as a PID control system. With a variable configuration, the structure is easy to experiment and convenient to recognize results that occur in the system with adjusting the gain of kp, ki, and kd then observed the behavior that occurred. If it does not go as expected, it is able to adjust the above criteria into the system and check the response from the output. In-dept some system PID controller, there will be a limitation for some of the results or the controlled device will affect if one of them has changed.

Especially nonlinear systems with relatively ambiguous controls will make that system always have problems [4]. Fuzzy Logic Control is another type of control system that is capable of non-linear systems. In which Fuzzy Logic is a processing technique that is flexible and the ability to assign concrete details to the system. In addition, the integration of fuzzy logic control systems with the PID system is another way to help control or manage the processes within this system. In this paper, we propose conceptually and have applied the research of Amornphun Phunopas [1]. To further develop a control system of a four-legged robot. It focuses on using PID to control of gait and using a fuzzy system to set a range of error from PID controlled with determining gain of value into the PID system before control of gait

2. The robot specification

The structure of the four-legged robot will be used according to Amornphun Phunopas[1], which has details of the kinematics equations already. Moreover, that has a real robot used for this research with details in Table 1.

TABLE 1. SPECIFICATION OF THE FOUR-LEGGED ROBOT

Description	Value
Width	250 mm.
Length	235 mm.
Height	220 mm.
Weight	0.8 kg. (without battery)
Controller	OpenCM V.9
Actuator	Digital Servo AX-18
DOF	12
Sensor	IMU 6 DOF

3. Control system design

In this controller design, that will be using a concept from Akekalak Supamanee[4]. The research with the concept to create a control system that responds to also the complexity within the robot. Before the design controller system robot that wants to control will walking gait move forward only. There is an IMU sensor used to measure the robot's tilt on each side, the angle of the robot, as shown in Fig. 1. In the three-axis of orientation, but for this step control, there is an axis The responsive PID and roll axis are used to create the control system because he robot does not control the rotation or movement by turning, so the yaw axis does not affect the system.

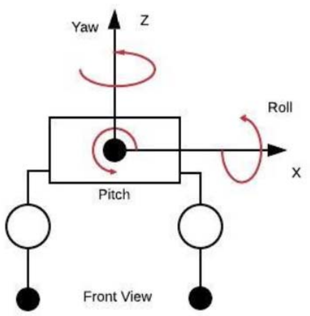


Fig. 1: The robot on front view and all orientation

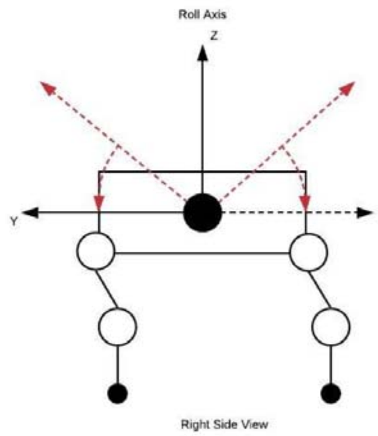


Fig. 2: The robot on right side view

After that, the control position of any legs on the robot with a mathematics equation of feedback control theory using the controller using the PID technique follow equation (1),(2), and (3). It is a gain of proportional, integral, and derivative. Moreover, that has fuzzy controller whose input is error value resulting from the PID controller. The into the fuzzy controller is a tool for managing error values to be compared with the errors from the PID system until to result. Before choosing the output, the hat comes from both types of control systems. Before going to control the plant. Follow Fig.3 is system diagram concept design for control this robot

$$P = k_p e(t) dt \quad (1)$$

$$\mathrm{I} = k_i \int_0^t e(t) dt \quad (2)$$

$$\mathrm{D} = k_d \frac{de(t)}{dt} \quad (3)$$

When combined all of gain

$$\mathrm{U}c(t) = k_p e(t) dt + k_i \int_0^t e(t) dt + k_d \frac{de(t)}{dt} \quad (4)$$

U$c(t)$ is the output of the PID controller and U$f(t)$ is the output of the fuzzy controller. The real output is input into a plant is U(t) is caused by combined as

$$\mathrm{U}(t) = \mathrm{U}c(t) + \mathrm{U}f(t) \quad (5)$$

Follow Fig.3 is system diagram concept design for control this robot

© The 2020 International Conference on Artificial Life and Robotics (ICAROB2020), Jan. 13-16, B-Con Plaza, Beppu, Oita, Japan

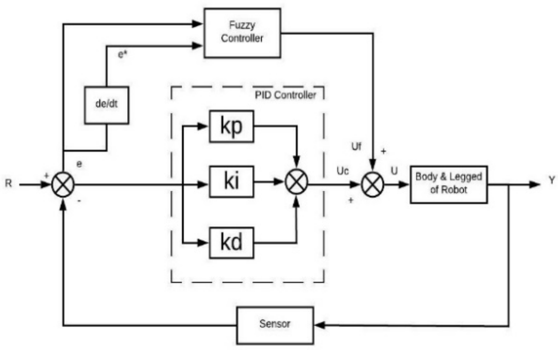


Fig. 3: Controls system diagram

Then the robot control will be moving forward There will always be moved along with the pitch and roll axis. The tilt angle of both axes will affect the walking gait. The fuzzy controller will receive input from the error caused by the forward movement of the robot always. Therefore is the source of error comparison. That comes from tilting on the robot. This part will choose to determine the membership level of the variable that must be using the triangular membership function consisting of 3 parameters a, b, and c, as equation (6). When the value a is the lowest range, b is the middle range and c is the maximum range of members

$$\textit{Triangular}\ (x:a,b,c) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ \frac{c-x}{c-b} & a \leq x \leq b \\ 0 & x \geq c \end{cases} \tag{6}$$

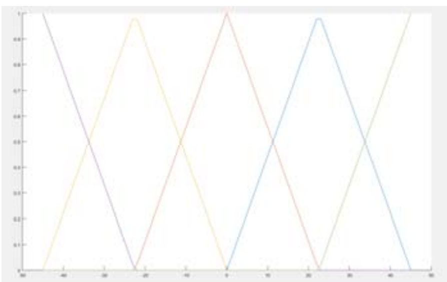


Fig. 4: Triangle Membership Function

4. Gait pattern design

The walking gait we chose to design the method follow the Fig.5. By using symmetrical gaits, footfalls of two feet of the fore pair and the hind pair are evenly spaced in time [5].

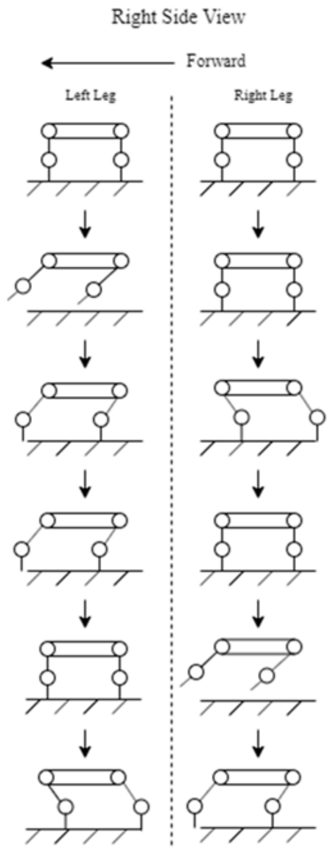


Fig. 5: Gait pattern design

5. Experiment

As shown in Fig. 5, a walking pattern has been designed. In the experiment, it begins with finding the range of the position control of the exciter of each leg. In the experiment, the researcher will control the position of the leg which is used to move according to the picture which can be seen that six steps in the design that will make the robot move forward For the angle of the robot legs. This is the result in Table 2. by the configuration frame shown in Fig. 6

TABLE 2. THE JOINT ANGLE)°(OF ALL LEG FROM STEP WALKING GAIT

	FR leg			**RR leg**			**FL leg**			**LL leg**		
t	θ_1	θ_2	θ_3	θ_1	θ_2	θ_3	θ_1	θ_2	θ_3	θ_1	θ_2	θ_3
1	*2*	*-90*	*90*	*178*	*-90*	*90*	*178*	*-90*	*90*	*-2*	*-90*	*90*
2	*2*	*-90*	*90*	*178*	*-90*	*90*	*157*	*-90*	*90*	*-19*	*-90*	*90*
3	*-20*	*-90*	*70*	*157*	*-90*	*70*	*157*	*-90*	*70*	*-19*	*-90*	*70.2*
4	*2*	*-90*	*90*	*178*	*-90*	*90*	*157*	*-90*	*70*	*-19*	*-90*	*70.2*
5	*20*	*-90*	*90*	*-157*	*-90*	*90*	*178*	*-90*	*90*	*-2*	*-90*	*90*
6	*-20*	*-90*	*119*	*157*	*-90*	*119*	*-157*	*-90*	*119*	*19*	*-90*	*118*

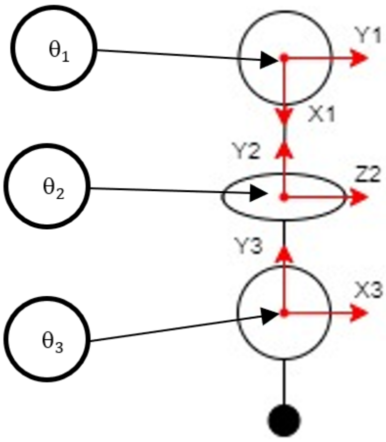


Fig.6: A one -legged configuration

While doing a walking gait experiment, the graph of the IMU sensor response on the robot was measured. The results have shown in Fig. 7 is a graph of the response axis of the pitch. The average initial measurement value is 0.4°

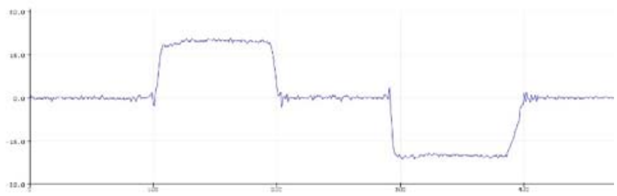


Fig.7: Pitch axis response from IMU sensor

Follow as Fig. 7 have results of the pitch angle response from tilting the attached sensor on the robot. Will get the values shown in Table 3.

TABLE 3. THE PITCH REPONSE TILTING FROM FIG. 7

Action	**AVG Angle response (°)**
Stand	0.4
Up Left/Down Right slope	22.04
Up Right/Down Left slope	-22.1

Meanwhile as shown in Fig.8 is a graph of the response axis, which the average initial measurement value is 4.2°

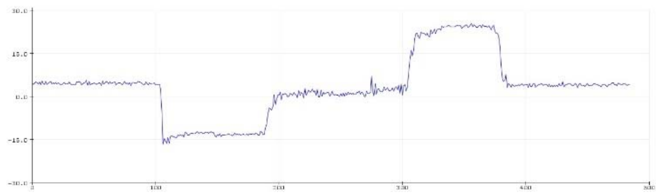


Fig.8: Roll axis response from IMU sensor

For the results of the response from tilting the attached sensor on the robot get the values shown in Table 4.

TABLE 4. THE ROLL REPONSE TILTING FROM FIG. 8

Action	**AVG Angle response (°)**
Stand	4.2
Up Front/Down Rear	25.6
Up Rear/Down Front	-14.8

6. Conclusion

However, this Gait control is still a complicated matter for controlling Which has to be collected and proven each time the experiment is performed, including the gait pattern chosen. There is a tendency for the control systems of this four-legged robot. Must use the dynamics of the control system to make the gait control stable. In the future of this research will still use the Fuzzy-PID control system but will change the design of a fuzzy set and control system diagrams difference from the previous. Because the Fuzzy-PID control system is able to define variables in the design, the fuzzy set defines basic rules. Moreover, any function format For the control of this robot. While also trying to collect comparative results in many format. whether using only Fuzzy Logic. compared to using Fuzzy-PID or using Fuzzy-PID compared to using Fuzzy-PID Auto-Tuning has been using all these control models. Will depend on the gait model used for this research

References

1. Amornphun Phunopas, “A Four-legged Robot’s Soft Feet Structural Design and Walking Gait Generated from Inverse Kinematics,” Journal of Robotics, Networking and Artificial Life Vol. 5(3); December (2018), pp. 161-164
2. Marc, Kevin, Gabriel, Rob and The BigDog Team “ BigDog, the Rough-Terrain Quadruped Robot,” The International Federation of Automatic Control, Seoul, Korea, July 6-11, (2008)
3. M. H. Raibert, Legged robots. MIT press, Vol. 29 No.6, June (1986), pp.499-514. Available from http://agents.sci.brooklyn.cuny.edu/cc30.03/papers/p499-raibert.pdf
4. Akekalak Supamanee, “Fuzzy - PID Controller for Heading Control of Small Helicopter,” The 21st Conference on Mechanical Engineering Network of Thailand, October 17-19, (2007)
5. L.Y.Y. Wang, Weijun Tao, “Gait Analysis and Implementation of a Simple Quadruped Robot,” The 2nd International Conference on Industrial Mechatronics and Automation, Wuhan, China, May 30-31, (2010)
6. Petrus Sutyasadi, Manukid Parnichun, “TROTTING CONTROL OF A QUADRUPED ROBOT USING PID-ILC,” IECON 2015 – 41st Annual Conference of the IEEE Industrial Electronics Society, Yokohama, Japan, November 17-20,(2015)

Hybrid Force/Position Teaching and Control Method for 6 DoF Manipulator Utilizing f-PAWTED

Quang-Trung Chu
Department of Electrical and Mechanical Engineering, Graduate School of Engineering, Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya, Aichi 466-8555, Japan

Hiroki Tanaka, Hideki Inuzuka, 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

Masao Sakai
Industrial Research Center, Aichi Center for Industry and Science Technology
Onda-cho 1-157-1, Kariya-shi Aichi 448-0013, Japan
E-mail: cmc13126@nitech.jp, h.tanaka.683@nitech.jp,
h.inuzuka.776@nitech.jp, morita@nitech.ac.jp, sakai@aichi-inst.jp
www.nitech.ac.jp

Abstract

The parallel wire-type teaching device with a force sensor (f-PAWTED) and a hybrid force/position teaching and control method developed in this study are one of the novel approaches to teach robot manipulators easily by human hands instead of teaching pendants. A direct teaching method based on hybrid force/position control was proposed to teach robots both desired position and force trajectories. The effectiveness of the developed device and method in teaching the robot manipulator was confirmed experimentally.

Keywords: Direct teaching method, Hybrid force/position control, f-PAWTED, Teaching pendant, Manipulator.

1. Introduction

Nowadays, robot teaching technology plays an intrinsic role to exploit robots adequately. A teaching pendant has been widely utilized in manufacturing industry to teach robots. Several teaching methods were proposed by extant studies.[1,2] A direct teaching method using impedance control was proposed so as to teach robot desired trajectories while directly moving the end effector.[3] Nevertheless, this method consumes a significant amount of time and the response characteristic of the robot during operation is not good. To cope with these difficulties, a parallel wire-type teaching device (PAWTED) and direct teaching position control method were proposed.[4] An operator can handle PAWTED to teach a six-degree-of-freedom (6-DoF) robot any arbitrary three-dimensional trajectories within working space in short period. Then, the robot can reproduce teaching position trajectories precisely.[5] However, the developed PAWTED is not able to teach and playback simultaneously both desired position and force trajectories.

In this study, the parallel wire-type teaching device with a force sensor (f-PAWTED) and a hybrid force/position teaching and control method are developed to overcome these problems. The effectiveness of the f-PAWTED and the mentioned teaching method is verified through experiment results.

2. Parallel Wire-Type Teaching Device with a Force Sensor

2.1. *f-PAWTED Mechanism*

The PAWTED consists of a moving platform, six wires and a base. The f-PAWTED comprises the PAWTED and a force sensor, as shown in Fig. 1. The force sensor is fixed on the top face of the moving platform. The base

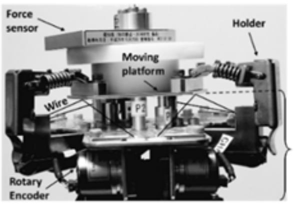


Fig. 1. f-PAWTED.

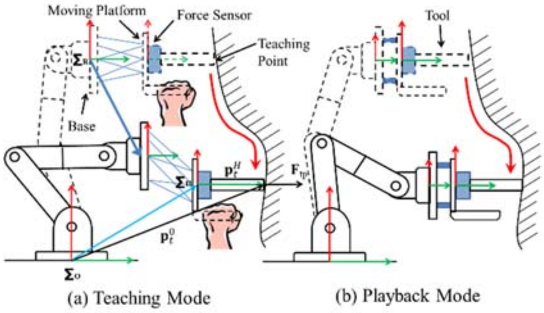


Fig. 2. The operation of f-PAWTED.

fastened firmly to the end effector. It comprises holders, rotary encoders and flat spiral springs. The moving platform is connected to the base by a system of six wires. Therefore, the moving platform can move freely in three-dimensional space. Figure 2 shows the operation of f-PAWTED. In the teaching mode, the f-PAWTED measures the drawn length of wires while the operator moves the moving platform. It computes the position and orientation of teaching points based on the drawn lengths. In addition, force teaching data is also measured by the force sensor. All teaching data is transformed to robot's reference frame $\sum_0$ and saved by the computer. During teaching, the robot is controlled to track and keep the moving platform at a constant distance. In the playback mode, the moving platform is fastened to holders and the robot reproduces the teaching trajectories.

2.2. *Theoretical Transformation*

2.2.1. Teaching mode

As shown in Fig. 2, $\sum_H$ and $\sum_R$ are the coordinate frames attached to the moving platform and the robot's tip, respectively. While moving the tool, teaching position and force data expressed in $\sum_0$ are saved to matrix $\mathbf{r}_t^0$ and $\mathbf{F}_t^0$, respectively. $\mathbf{r}_t^0$ is given as

$$\mathbf{r}_t^0 = [\mathbf{p}_t^0 \quad \boldsymbol{\eta}_t^0]^T, \tag{1}$$

in which $\mathbf{p}_t^0$ is a position matrix and $\boldsymbol{\eta}_t^0$ is an orientation matrix. They are given as follows

$$\mathbf{p}_t^0 = \mathbf{T}_H^0 \mathbf{p}_t^H, \qquad \boldsymbol{\eta}_t^0 = \mathbf{R}_R^0 \boldsymbol{\eta}_t^R,$$
$$\mathbf{p}_t^H = [0,0,l_h]^T, \qquad \boldsymbol{\eta}_t^R = [\phi, \theta, \psi]^T.$$

$\mathbf{T}_k^i$ and $\mathbf{R}_k^i$ denote the homogeneous transform matrix and rotation matrix from frame $\sum_k$ to $\sum_i$, respectively. Force data comprises force $\mathbf{f}_t^0$ and torque $\boldsymbol{\tau}_t^0$. Similar to the above transformations, the force matrix is transformed from $\sum_H$ to $\sum_0$ and saved as

$$\mathbf{F}_t^0 = [\mathbf{f}_t^0 \quad \boldsymbol{\tau}_t^0]^T. \tag{2}$$

2.2.2. Playback mode

In the playback mode, the moving platform is fasted to the base, so $\sum_R = \sum_H$ and $\mathbf{T}_H^R = \mathbf{I}, \mathbf{T}_H^0 = \mathbf{T}_R^0$. Position and force data are saved as

$$\mathbf{r}_p^0 = \left[\mathbf{p}_p^0 \quad \boldsymbol{\eta}_p^0\right]^T, \quad \mathbf{F}_p^0 = \left[\mathbf{f}_p^0 \quad \boldsymbol{\tau}_p^0\right]^T. \tag{3}$$

Basing on feedback data, adjustment is computed by hybrid control algorithm. Then, a robot controller solves the inverse kinematics and dynamics problem.

3. Hybrid Force/Position Teaching Control Method

Hybrid control method was researched on the principle that some directions should be used to control position and other ones in which force should be controlled. Upon the specific task, these directions change but a single direction is never used to control both position and force.[6] Our task is to teach the robot attached f-PAWTED position and force trajectories on a spherical surface. The operator pushes the endpoint of robot arm with a force against the working surface while moving the endpoint. Then, robot reproduces the teaching trajectories comprising both position and force. We assign a moving constraint frame $\sum_C$ with its origin O_C is the present point A_i on the working surface, its Z_C axis is normal to the surface, it's also the force-control direction, its Y_C axis in the direction which is the direction of the cross product of vector e_{zc} and e_z. They are unit vectors of Z_C and Z axis, respectively. Its X_C axis such that three axes form a right-hand coordinate frame. X_C and Y_C axis are position-control directions. The moving constraint frame is shown in Fig. 3. Teaching and playback data are transformed to frame $\sum_C$ as follows:

$$[\mathbf{p}_t^C \quad 0]^T = \mathbf{I}_1 \mathbf{T}_0^C [\mathbf{p}_t^0 \quad 1]^T, \quad \boldsymbol{\eta}_t^C = \mathbf{I}_2 \mathbf{R}_0^C \boldsymbol{\eta}_t^0, \tag{4}$$

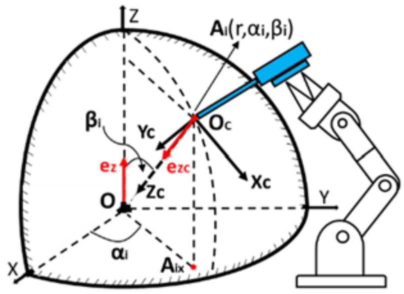


Fig. 3. Moving constraint frame on spherical surface.

$$\mathbf{f}_t^C = \mathbf{I}_3\mathbf{R}_0^C\mathbf{f}_t^0\,, \quad [\boldsymbol{\tau}_t^C \quad 0]^T = \mathbf{I}_4\mathbf{T}_0^C[\boldsymbol{\tau}_t^0 \quad 1]^T,$$
$$\mathbf{f}_p^C = \mathbf{I}_3\mathbf{R}_0^C\mathbf{f}_p^0\,, \quad \left[\boldsymbol{\tau}_p^C \quad 0\right]^T = \mathbf{I}_4\mathbf{T}_0^C\left[\boldsymbol{\tau}_p^0 \quad 1\right]^T, \quad (5)$$

where $\mathbf{I}_1 = \text{diag.}\,[1,1,0,0]$, $\mathbf{I}_2 = \text{diag.}\,[1,1,0,]$, $\mathbf{I}_3 = \text{diag.}\,[0,0,1]$, $\mathbf{I}_4 = \text{diag.}\,[0,0,1,0]$.

The compensatory for force error considering in frame $\sum_C$ applying PD (proportional and differential) control law is given as

$$\mathbf{r}_p^C(n) = \mathbf{r}_t^C(n) + \mathbf{K}_P\left[\mathbf{F}_t^C(n-1) - \mathbf{F}_p^C(n-1)\right] + \mathbf{K}_D\left[\frac{\mathbf{F}_t^C(n-1) - \mathbf{F}_p^C(n-1)}{dT}\right], \quad (6)$$

where $\mathbf{K}_P$ and $\mathbf{K}_D$ are the feedback gain matrices, dT is the sampling time.

4. Experiment

In order to verify the effectiveness of hybrid force/position teaching and control method using the f-PAWTED, experiments utilizing position teaching and control method with the PAWTED were conducted as comparative ones. The former and latter are referred to as the proposed and conventional methods, respectively. These experiments were performed with 6-DoF robot (DENSO WAVE, VS-060) installed the f-PAWTED and a solid ball with a radius of 85.5 mm. A pen was attached to the head of force sensor. A pen holder that an operator grasps is fastened to the moving platform. In teaching mode, the operator manipulated the pen holder to teach the robot a curve from point A_1 to A_n as shown in Figs. 4 and 5. Position of teaching point A_i expressed in spherical coordinate frame comprises three parameters (r, α_i, β_i), as shown in Fig. 3. r is sphere radius, α_i is the angle between axis X and vector $\overrightarrow{OA_{ix}}$, β_i denotes the angle between axis Z and vector $\overrightarrow{OA_i}$. We based the error computation on feedback parameters consisting of position, orientation

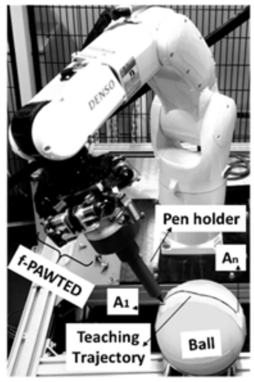


Fig. 4. Playback motion of robot attached to f-PAWTED.

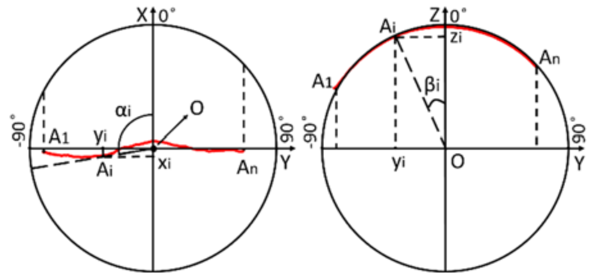


Fig. 5. Teaching trajectory on spherical surface.

and force from sensors attached to robot. The force errors with respect to axis Z_c between teaching and playback modes are shown in Fig. 6. It is seen that the force of playback mode of the proposed method followed the teaching force more closely than that of the conventional method. In the proposed and conventional methods, the ranges of the error were 0 to 0.76 N and 0.36 N to 3 N, respectively. They are illustrated by dark blue and brown lines in Fig. 6. The position and orientation trajectories are shown in Figs. 7 and 8, respectively. The playback position and orientation trajectories closely pursue the teaching ones when utilizing either proposed or conventional methods.

Hence, it proved the usefulness of developed teaching device and the applicability of hybrid force/position teaching and control method in robot teaching field.

5. Conclusion

This study proposed the hybrid force/position teaching control method and f-PAWTED for manipulators. We can foster the industrial processing in both accuracy and productivity by utilizing f-PAWTED. The above results demonstrated the efficacy as well as prospects of the direct teaching and control method for manipulators.

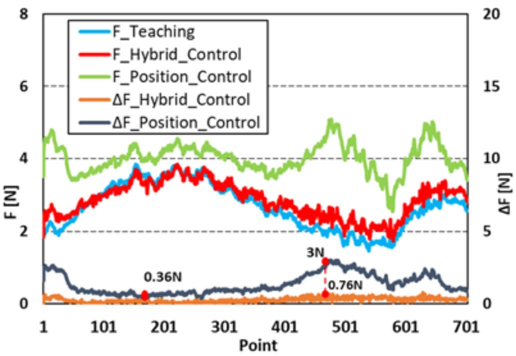
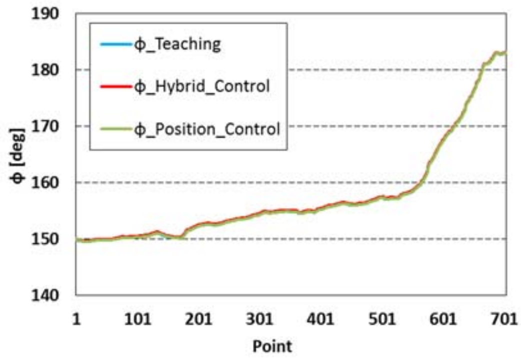
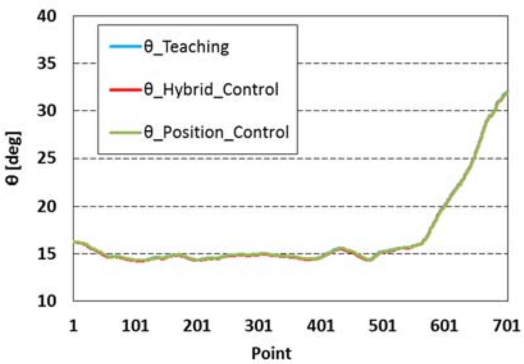


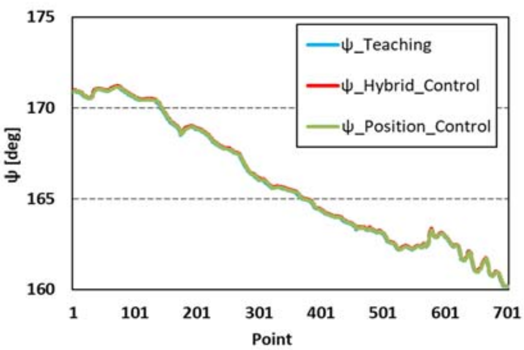
Fig. 6. Force errors expressed in Z_c axis.



(a) ϕ angle

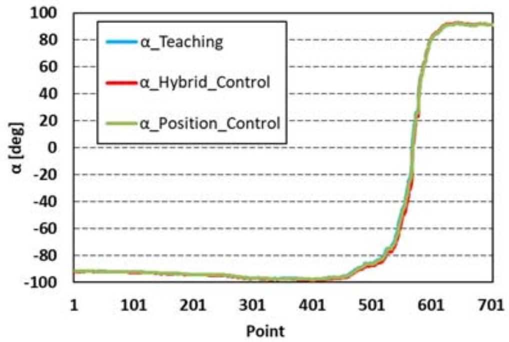


(b) θ angle

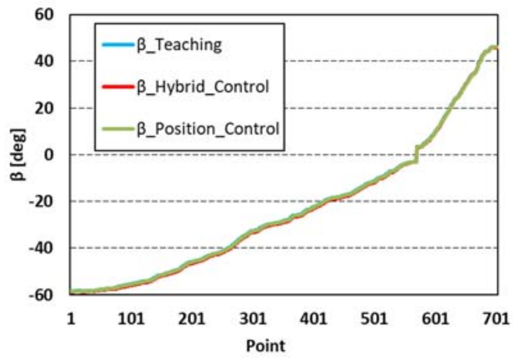


(c) ψ angle

Fig. 7. Teaching and playback orientation trajectories.



(a) α angle



(b) β angle

Fig. 8. Teaching and playback position trajectories.

References

1. S. Masaru, Direct Teaching for Robot, *J. Robotics Society of Japan*. **13**(5), 1995, pp. 627-628.
2. M. Kanji, Offline Teaching, *J. Robotics Society of Japan*. **13**(5), 1995, pp. 611-614.
3. Y. Mochizuki and T. Matsui, Direct Teaching Method for Hybrid Control Applying Impedance Control, *J. Robotics Society of Japan*. **12**(2), 1994, pp. 160-166.
4. S. Sakai, S. Noritaka, M. Yoshifumi, Development of Parallel Wire Mechanism for Robot Teaching, *J. Japan Society of Mechanical Engineers Series C*. **79**(807), 2013, pp. 4321-4329.
5. S. Sakai, S. Noritaka, M. Yoshifumi, Performance Evaluation of Robot Teaching Method with Parallel Wire Mechanism, *J. Society of Signal Processing Applications and Technology of Japan*. **16**(1), 2013, pp. 72-78.
6. T. Yoshikawa, *Foundations of Robotics-Analysis and Control*, The MIT Press, 1990.

A Study on Generalized Predictive Control in Consideration of Noise

Akira Yanou
Kawasaki University of Medical Welfare, 288 Matsushima
Kurashiki, Okayama 7014221, Japan
E-mail: yanou-a@mw.kawasaki-m.ac.jp
www.kawasaki-m.ac.jp

Abstract

Generalized Predictive Control (GPC) is one of the model-based control methods. The control law is derived through the performance index based on the sum about the squares of control input and the squares of the error between reference signal and output prediction. Although coprime factorization approach has been used in order to extend the conventional control law in the previous researches, there has been a possibility that the order of the derived control law becomes high. Therefore, this paper extends GPC through newly-defined output prediction and proposes the method to re-design the control law or the characteristic from noise to output with keeping the closed-loop transfer function. Numerical example is shown to check the characteristic of the proposed method.

Keywords: Generalized predictive control, Closed-loop characteristics, Noise, Output prediction.

1. Introduction

Generalized predictive control (GPC) [1] is one of the model-based design methods and the closed-loop system is designed through performance index. The performance index includes the error between the reference signal and the output prediction, and the control input. For consideration of designing safe systems, although coprime factorization approach [2] has been used in order to extend the control law in the previous researches [3], the derived controller will become high order because a stable polynomial is needed to derive the extended controller. Therefore, this research directly extends generalized predictive control [4] by defining new output prediction and proposes the scheme to re-design the controller without changing the closed-loop characteristics. A numerical example is shown to check the characteristics of the proposed method.

2. Extension of Generalized Predictive Control

A single-input single-output system is considered for $t = 0, 1, 2 \cdots$.

$$A[z^{-1}]y(t) = z^{-k_m}B[z^{-1}]u(t) + C[z^{-1}]\frac{\xi(t)}{\Delta} \quad (1)$$

$y(t)$ and $u(t)$ are output and input respectively. k_m is time delay, $\xi(t)$ is white Gaussian noise and $\Delta = 1 - z^{-1}$. $A[z^{-1}]$, $B[z^{-1}]$ and $C[z^{-1}]$ are the following polynomials with known degrees n, m and l.

$$A[z^{-1}] = 1 + a_1 z^{-1} + \cdots + a_n z^{-n} \quad (2)$$

$$B[z^{-1}] = b_0 + b_1 z^{-1} + \cdots + b_m z^{-m} \quad (3)$$

$$C[z^{-1}] = 1 + c_1 z^{-1} + \cdots + c_l z^{-l} \quad (4)$$

For (1), the following assumptions are hold.

(i) k_m is known.
(ii) The pairs of $(A[z^{-1}], B[z^{-1}])$ and $(A[z^{-1}], C[z^{-1}])$ are coprime.
(iii) $C[z^{-1}]$ is stable polynomial.

By minimizing the following performance index J, the conventional control law can be derived.

$$J = E_x\left[\sum_{j=N_1}^{N_2}\{y(t+j) - w(t+j)\}^2 + \sum_{j=1}^{N_u}\lambda\{\Delta u(t+j-1)\}^2\right] \quad (5)$$

$E_x[]$ means the expected value. $[N_1, N_2]$, $[1, N_u]$ and λ are design parameters named as prediction horizon, control horizon and weighting factor of control input. Their parameters are needed to determine a stable closed-loop system.

In order to derive the control law, the predicted output $\hat{y}(t+j|t)$ for $j = N_1 \cdots N_2$ is calculated by solving the following Diophantine equation.

$$C[z^{-1}] = \Delta A[z^{-1}]E_j[z^{-1}] + z^{-j}F_j[z^{-1}] \quad (6)$$

$E_j[z^{-1}]$ and $F_j[z^{-1}]$ are expressed as follows.

$$E_j[z^{-1}] = 1 + e_1 z^{-1} + \cdots + e_{j-1} z^{-(j-1)} \quad (7)$$

$$F_j[z^{-1}] = f_0^j + f_1^j z^{-1} + \cdots + f_n^j z^{-n} \quad (8)$$

Multiplying $z^j \Delta E_j[z^{-1}]$ to (1) and substituting (6) into it, the following equation is obtained.

$$\begin{aligned} C[z^{-1}]y(t+j) = E_j[z^{-1}]B[z^{-1}]\Delta u(t+j-k_m) \\ + F_j[z^{-1}]y(t) \\ + E_j[z^{-1}]C[z^{-1}]\xi(t+j) \end{aligned} \quad (9)$$

Although the output prediction is defined without including the future noise term $E_j[z^{-1}]C[z^{-1}]\xi(t+j)$ in conventional GPC, this paper proposes the use of the noise term to date, which can be calculated by (1). Concretely, the following output prediction is newly defined by introducing constant parameter s_e.

$$\begin{aligned} \hat{y}(t+j|t) = \frac{1}{C[z^{-1}]}\{E_j[z^{-1}]B[z^{-1}]\Delta u(t+j-k_m) \\ +F_j[z^{-1}]y(t) + s_e C[z^{-1}]\xi(t)\} \end{aligned} \quad (10)$$

Where

$$C[z^{-1}]\xi(t) = \Delta A[z^{-1}]y(t) - z^{-k_m}B[z^{-1}]\Delta u(t) \quad (11)$$

Moreover, the following equations are considered.

$$E_j[z^{-1}]B[z^{-1}] = C[z^{-1}]R_j[z^{-1}] + z^{-j}S_j[z^{-1}] \quad (12)$$

$$R_j[z^{-1}] = r_0 + r_1 z^{-1} + \cdots + r_{j-1} z^{-(j-1)} \quad (13)$$

$$S_j[z^{-1}] = s_0^j + s_1^j z^{-1} + \cdots + s_{n_3}^j z^{-n_3} \quad (14)$$

Where $n_3 = \max(m, l) - 1$. Then the output prediction (10) can be re-expressed as follows.

$$\begin{aligned} \hat{y}(t+j|t) = \frac{1}{C[z^{-1}]}\{C[z^{-1}]R_j[z^{-1}]\Delta u(t+j-k_m) \\ +(F_j[z^{-1}] + s_e \Delta A[z^{-1}])y(t) \\ +(S_j[z^{-1}] - s_e B[z^{-1}])\Delta u(t-k_m)\} \end{aligned} \quad (15)$$

The following equations are defined.

$$F_j'[z^{-1}] = F_j[z^{-1}] + s_e \Delta A[z^{-1}] \quad (16)$$

$$S_j'[z^{-1}] = S_j[z^{-1}] - s_e B[z^{-1}] \quad (17)$$

The past and present signals with output and input are also defined by $h_j'(t)$.

$$C[z^{-1}]h_j'(t) = F_j'[z^{-1}]y(t) + S_j'[z^{-1}]\Delta u(t-k_m) \quad (18)$$

Then the output prediction can be expressed as the following equation, which separates the input term $R_j[z^{-1}]\Delta u(t+j-k_m)$ in the current and future time from the other signal $h_j'(t)$.

$$\hat{y}(t+j|t) = R_j[z^{-1}]\Delta u(t+j-k_m) + h_j'(t) \quad (19)$$

The following vectors and coefficients at time t are defined for $j = N_1, \cdots, N_2$.

$$\hat{Y} = [\hat{y}(t+N_1|t) \cdots \hat{y}(t+N_2|t)]^T \quad (20)$$

$$U = [\Delta u(t) \cdots \Delta u(t+N_u-1)]^T \quad (21)$$

$$H = [h_{N_1}'(t) \cdots h_{N_2}'(t)]^T \quad (22)$$

$$W = [w(t+N_1) \cdots w(t+N_2)]^T \quad (23)$$

$$R = \begin{bmatrix} r_{N_1-k_m} & \cdots & r_0 & 0 \\ \vdots & & \ddots & r_0 \\ \vdots & & & \vdots \\ r_{N_2-k_m} & \cdots & \cdots & r_{N_2-N_u-k_m-1} \end{bmatrix} \quad (24)$$

Then the output prediction (15) can be expressed as following vector form.

$$\hat{Y} = RU + H \quad (25)$$

By using the above equation, the performance index J can be described as follows.

$$J = (RU + H - W)^T(RU + H - W) + \lambda U^T U \quad (26)$$

Minimizing the performance index J for the input vector U, the control input can be given.

$$\Delta u(t) = [1\ 0\ \cdots\ 0](R^T R + \lambda I)^{-1} R^T (W - H) \quad (27)$$

Where

$$[p_{N_1} \cdots p_{N_2}] = [10 \cdots 0](R^T R + \lambda I)^{-1} R^T \quad (28)$$

$$P[z^{-1}] = p_{N_2} + p_{N_2-1} z^{-1} + \cdots + p_{N_1} z^{-(N_2-N_1)} \quad (29)$$

$$F_p'[z^{-1}] = p_{N_1} F_{N_1}'[z^{-1}] + \cdots + p_{N_2} F_{N_2}'[z^{-1}] \quad (30)$$

$$S_p'[z^{-1}] = p_{N_1} S_{N_1}'[z^{-1}] + \cdots + p_{N_2} S_{N_2}'[z^{-1}] \quad (31)$$

Then the transfer function of control law is expressed as follows.

$$\Delta u(t) = \frac{C[z^{-1}]P[z^{-1}]}{C[z^{-1}] + z^{-k_m} S_p'[z^{-1}]} w(t+N_2)$$

$$-\frac{F_p'[z^{-1}]}{C[z^{-1}]+z^{-k_m}S_p'[z^{-1}]}y(t) \qquad (32)$$

Moreover, the following polynomials are defined.

$$D_p[z^{-1}] = \Delta A[z^{-1}]S_p'[z^{-1}] + B[z^{-1}]F_p'[z^{-1}] \qquad (33)$$

$$T[z^{-1}] = \Delta A[z^{-1}]C[z^{-1}] + z^{-k_m}D_p[z^{-1}] \qquad (34)$$

Then the closed-loop system can be derived as follows.

$$y(t) = \frac{z^{-k_m}B[z^{-1}]C[z^{-1}]P[z^{-1}]}{T[z^{-1}]}w(t+N_2) + \frac{C[z^{-1}]\left(C[z^{-1}]+z^{-k_m}S_p'[z^{-1}]\right)}{T[z^{-1}]}\xi(t) \qquad (35)$$

It means that the response characteristics for reference signal is the same as the conventional GPC. On the other hand, it can find that the response characteristics for noise can be tuned through the design polynomial $S_p'[z^{-1}]$ including s_e.

3. Numerical Example

The following controlled plant is considered [5].

$$A[z^{-1}] = 1 - 0.998775z^{-1}, B[z^{-1}] = 14.4$$
$$C[z^{-1}] = 1, k_m = 1$$

The number of simulation steps is 500, the initial values of input and output are set to be 0, and the variance of $\xi(t)$ is $\sigma^2 = 0.001$ (each data of noise $\xi(t)$ is the same for the conventional and the proposed method). The reference signal $w(t)$ is 0.1. The design parameters of GPC are given as follows.

$$N_1 = 1, N_2 = 6, N_u = 6, \lambda = 8500$$

In this example, the design parameter of the proposed method is given as $s_e = -1$.

Fig.1 and Fig.2 show the conventional output and input. In Fig.1 the dashed line and solid line mean the reference signal $w(t)$ and the output signal $y(t)$ respectively. In Fig.2 the upper figure shows the control input $u(t)$ and the lower one shows the input increment $\Delta u(t)$. Fig.3 and Fig.4 show the proposed output and input. Their lines are the same meanings as conventional method. From these figures, it can find that the proposed method can change the noise influence on output.

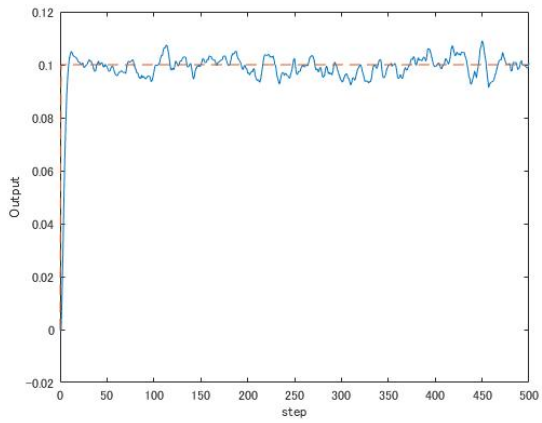


Fig.1 conventional output

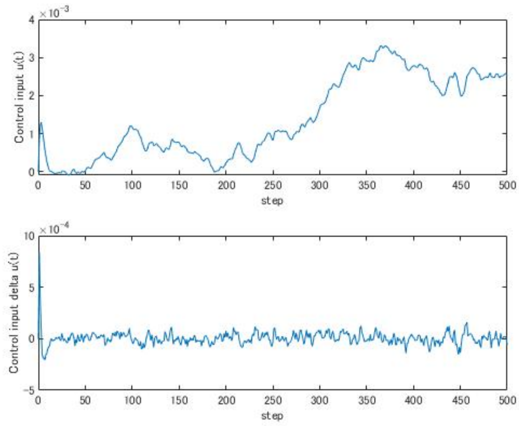


Fig.2 conventional input

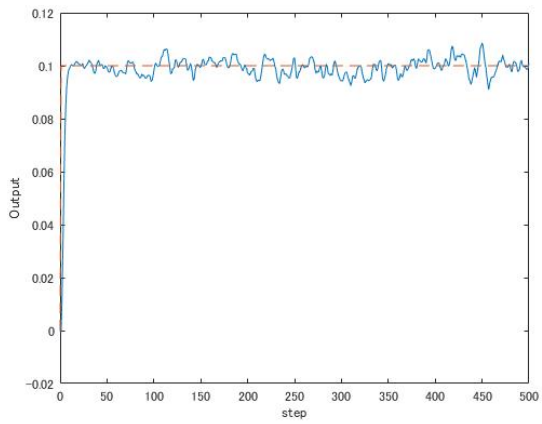


Fig.3 proposed output

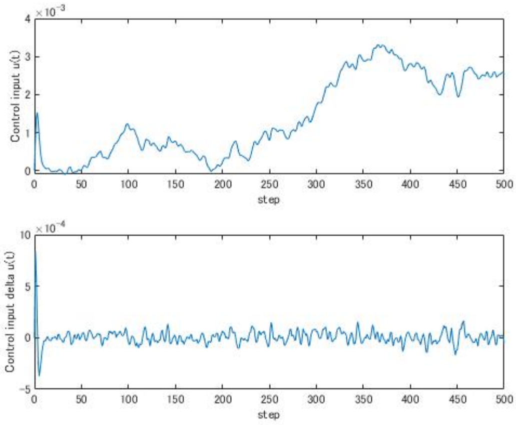


Fig.4 proposed input

4. Conclusion

This paper proposed a directly extended method of conventional GPC through newly-defined output prediction. The derived control law can re-design the characteristic from noise to output with keeping the closed-loop transfer function. Numerical example was given to check the characteristic of the proposed method.

As future work, the selection method of design parameter s_e should be considered.

References

1. S. Omatu and T. Yamamoto, *Self-tuning control*, Corona publishing, 1996.
2. M. Vidyasagar, *Control System Synthesis: A Factorization Approach*, The MIT Press, 1985.
3. A. Inoue, A. Yanou, and Y. Hirashima, A Design of a strongly Stable Self-Tuning Controller Using Coprime Factorization Approach, *Preprints of the 14th IFAC World Congress*, Vol. C, 1999, pp. 211–216.
4. A. Yanou, A Study on Extension of Generalized Predictive Control, *IEEJ Technical Meeting on Control*, 2019, CT–19–032.
5. S. Okazaki, J. Nishizaki, A. Yanou, M. Minami and M. Deng, Strongly Stable Generalized Predictive Control Focused on Closed-loop Characteristics, *Transactions of the Society of Instrument and Control Engineers*, Vol.47, No.7, 2011, pp.317–325.

Simulation Study on Emergency-Stopping Avoidance Control Due to Singularity During Teaching Operation with Parallel Wire-Type Teaching Device

Hideki Inuzuka
Department of Electrical and Mechanical Engineering, Nagoya Institute of Technology
Gokiso-cho, Showa-ku, Nagoya, Aichi 466-8555, Japan

Masao Sakai
Industrial Research Center, Aichi Center for Industry and Science Technology
Onda-cho 1-157-1, Kariya-shi Aichi 448-0013, Japan

Yoshifumi Morita
Department of Electrical and Mechanical Engineering, Nagoya Institute of Technology
Gokiso-cho, Showa-ku, Nagoya, Aichi 466-8555, Japan
E-mail: h.inuzuka.776@nitech.jp , masaosakai@pref.aichi.lg.jp, morita@nitech.ac.jp
www.nitech.ac.jp

Abstract

It is possible to teach industrial robots easily by using the parallel wire-type teaching device (PAWTED) developed in this study. One of several practical problems is the emergency stopping of the robot because of a singular configuration during teaching. A method is proposed to avoid these stops owing to the shoulder singular configuration. In the simulation, the robot motion was corrected by changing the target value following the PAWTED. The effectiveness of preventing the emergency stop from occurring was confirmed.

Keywords: Robot manipulator, Teaching device, Singular configuration, Emergency-stopping avoidance

1. Introduction

Robot teaching is generally done by using a teach pendant. However, this method consumes a significant amount of time and requires highly skilled operators. Therefore, the teaching operation can be greatly simplified by a direct teaching method using a force sensor. Nevertheless, the robot acts differently from the intended trajectory, because the response of the robot to the human's operation is delayed. [1] To cope with this difficulty, a direct teaching method utilizing a parallel wire mechanism is proposed. In addition, a parallel wire-type teaching device (PAWTED) and a direct teaching system were developed. [2,3] Because the work tool and the robot movement are not synchronized, it can solve the afore-mentioned malfunction. In practical experiments, the robot could fall into a singular configuration regardless of human intentions. At this time, the inverse kinematics problem cannot be solved, so the robot ceases operation for safety. Consequently, the teaching operation is interrupted. There are innumerable singular configurations in the robot action range, and measures to avoid emergency stopping of robot motion are necessary for smooth teaching operation.

In this study, a method is proposed to avoid the emergency stopping of the robot by avoiding the singular configurations without changing the desired trajectory of the work tool. The efficacy of the proposed control method was verified by conducting simulation to demonstrate that the emergency stop does not occur, even in a singular configuration or its vicinity.

2. Parallel Wire-Type Teaching Device

The PAWTED measures the relative position and orientation between the base and the moving platform by the drawn length of six wires, as shown in Fig. 1. It becomes a direct teaching system when this device is attached to an industrial six-degree-of-freedom (6-DOF) robot, as shown in Fig. 2. It is controlled when the current value of the base $\boldsymbol{X}_R$ moves to the following target value $\bar{\boldsymbol{X}}_T$ in the robot coordinate.

$$\bar{\boldsymbol{X}}_T = \boldsymbol{X}_R + \boldsymbol{X}_P + \boldsymbol{X}_d = \boldsymbol{X}_H + \boldsymbol{X}_d. \qquad (1)$$

$\boldsymbol{X}_P$ is the relative position and orientation between the base and the work tool. $\boldsymbol{X}_H$ is the position and orientation of the work tool. It is calculated from $\boldsymbol{X}_R$ and $\boldsymbol{X}_P$, and this value is saved as teaching information. $\boldsymbol{X}_d$ is the offset from the work tool, and it is a constant value as the distance and direction to the base. By solving the inverse kinematics problem, each joint angle is determined with respect to each matrix $\bar{\boldsymbol{X}}_T$ during operation. When the work tool stops, $\boldsymbol{X}_P = -\boldsymbol{X}_d$, and, as a result, the following target value becomes $\bar{\boldsymbol{X}}_T = \boldsymbol{X}_R$. Therefore, the base stops with certain offset from the work tool.

3. Emergency-Stopping Avoidance Control

3.1. *Emergency-Stopping State of Robot*

As shown in Fig. 3(a), in the shoulder singular configuration, the origin Σ_5 frame, O_5, locates on the Z_1 axis of Σ_1 frame. The position of the work tool between teaching and playback are the same, but the robot configurations are different. The manipulator can fall into singular configurations during teaching, even if it does not fall into singular configurations during reproduction. If the robot approaches a singular configuration during teaching, the robot cannot be controlled according to expectation. Consequently, the robot operating always ceases in such situations.

3.2. *Avoidance Control Method*

When the robot arm falls into shoulder singular configurations, the position of O_5 with respect to Σ_0 is given as (x_0, y_0, z_0)=($0, 0, z_0$). This point, called

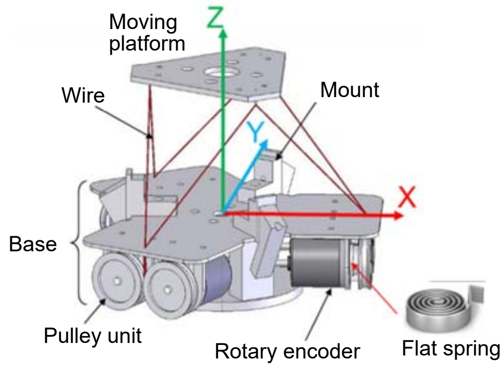
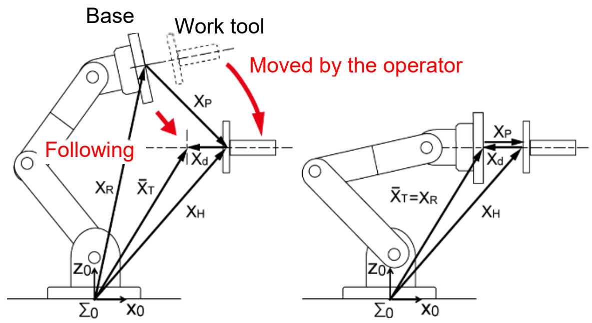
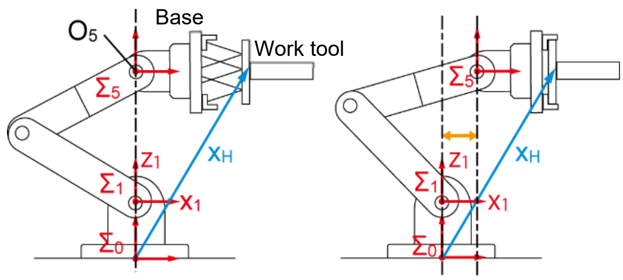


Fig. 1. Parallel wire-type teaching device.



(a) Moving work tool (b) Suspended work tool

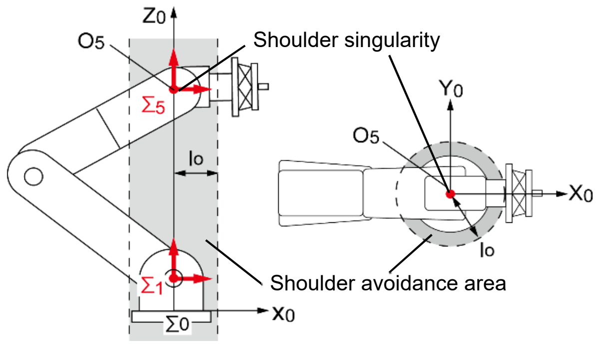
Fig. 2. Coordinate system during teaching operation.



(a) Teaching (b) Play back

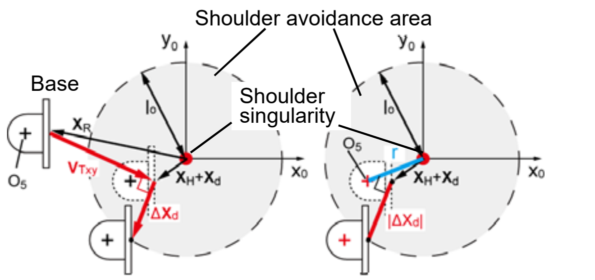
Fig. 3. Shoulder singular configuration in teaching operation.

"shoulder singularity," is shown in Fig. 4. It is defined that whole area inside the cylinder with radius l_o around the axis Z is shoulder avoidance area. This is illustrated in Fig. 5. O_5 passes through the shoulder avoidance area during teaching, and the aforementioned target value is corrected to keep O_5 away from that area. $\bar{\boldsymbol{X}}_T$ is changed by supplementing an avoidance vector expressed on the x_0y_0 plane, $\Delta\boldsymbol{X}_d$, as in Eq. (2). Because the system of wires can be drawn freely in working space, changing matrix $\bar{\boldsymbol{X}}_T$ does not alter the position and orientation of the work tool.



(a) Side view (b) Top view

Fig. 4. Definition of shoulder avoidance area.



(a) Avoidance vector (b) Avoidance amount

Fig. 5. Coordinate system of shoulder singularity avoidance.

$$\overline{\boldsymbol{X}}_T = \boldsymbol{X}_H + \boldsymbol{X}_d + \Delta\boldsymbol{X}_d. \quad (2)$$

As shown in Fig. 5, the direction of the avoidance vector $\Delta\boldsymbol{X}_d$ is an orthogonal direction to the vector $\boldsymbol{V}_{Txy}$ projected from the current value of the base $\boldsymbol{X}_R$ to the following target value $\boldsymbol{X}_H + \boldsymbol{X}_d$ on the x_0y_0 plane. Furthermore, this vector is a direction away from the shoulder singularity. $\Delta\boldsymbol{X}_d$ is expressed as $\Delta\boldsymbol{X}_d$=$[x_d, y_d, 0, 0, 0, 0]^T$. $\Delta\boldsymbol{X}_{d2} = [x_d,\ y_d]^T$ is the x and y component of the vector $\Delta\boldsymbol{X}_d$. This avoidance amount $|\Delta\boldsymbol{X}_{d2}|$ is given by

$$|\Delta\boldsymbol{X}_{d2}| = \sqrt{{x_d}^2 + {y_d}^2} = \begin{cases} \dfrac{c_o}{l_o}(l_o - r), & (0 < r < l_o) \\ 0, & (r \geq l_o) \end{cases} \quad (3)$$

where r is the distance between O_5 and the shoulder singularity. $|\Delta\boldsymbol{X}_{d2}|$ is the proportional coefficient depending on the distance r. c_o reaches the maximal value when $r = 0$.

4. Emergency-Stopping Avoidance Simulation

4.1. *Evaluation by Simulator*

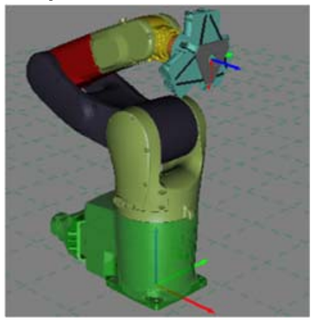
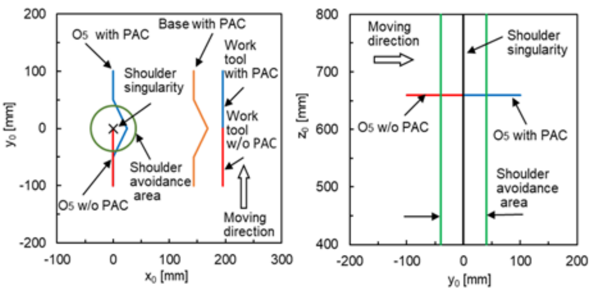


Fig. 6. Graphical manipulator in the simulator.

The developed simulator shown in Fig.6[4] was utilized for the evaluation of the proposed emergency stopping avoidance method. The virtual robot in the simulator is the emulation of the real robot manufactured by DENSO WAVE Co., model VS-060. For validation, the teaching trajectory in the robot coordinate in which robot passes through the singular configuration was prepared. To make it easier to understand the manipulator's behavior, the posture of the work tool was kept constant at (roll, pitch, yaw) = (45.0˚, 90.0˚, –45.0˚). The position of the work tool was kept constant at x_0= 195 mm and z_0= 660 mm. Then, the work tool was translated from y_0= –100 mm to y_0= 100 mm. In this situation, the fifth joint was kept constant at x_0= 0 mm and z_0= 660 mm. This joint passes through the shoulder singularity. Avoiding an emergency stop of the robot depends also on the teaching speed. To conform to work safety standards, the speed of the work tool was set to 50 mm/s. The proposed control parameters are the shoulder avoidance area radius l_o and the maximum avoidance amount c_o. The parameters were set as l_o= 50 mm and c_o= 25 mm. The movement of the robot was evaluated with the developed simulator.

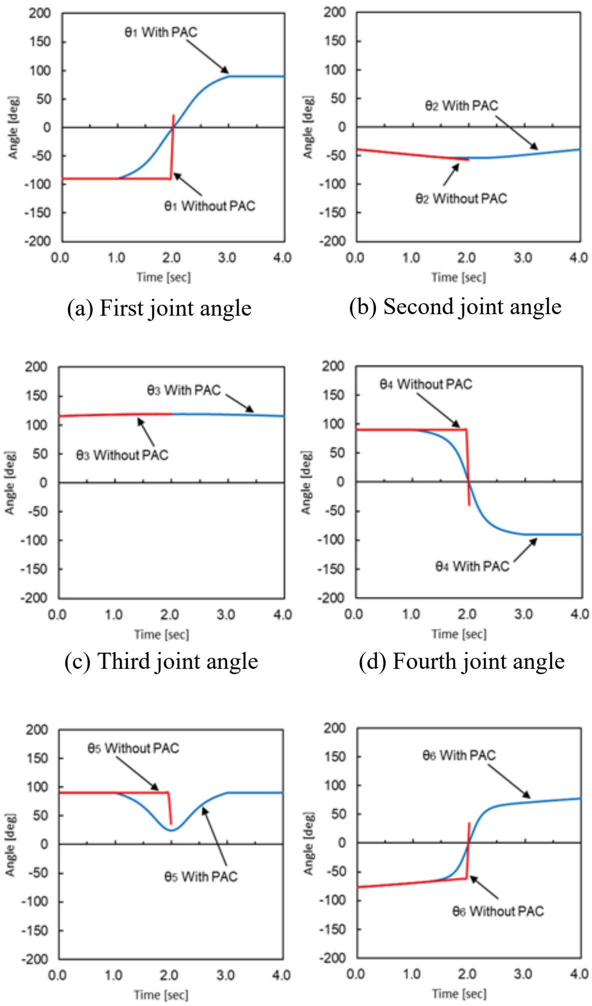
4.2. *Results and Discussion*

Figs. 7(a) and (b) show the trajectory of the work tool, O_5, and PAWTED's base with respect to Σ_0 when the work tool was moved to the prepared trajectory. If emergency-stopping avoidance control is not applied, the robot's following is stopped near the shoulder singularity. For this case, Figs. 8(a), (d), (e), and (f) show the respective joint angles changing rapidly. Because the first joint angle was changed rapidly, the fourth, fifth, and sixth joint angles were also changed rapidly. However, emergency-stopping avoidance control was applied.



(a) $x_0 - y_0$ view (b) $y_0 - z_0$ view

Fig. 7. Trajectories of O_5, base, and work tool points. (PAC: proposed avoidance control)



(a) First joint angle (b) Second joint angle

(c) Third joint angle (d) Fourth joint angle

(e) Fifth joint angle (f) Sixth joint angle

Fig. 8. Each joint angle of the robot. (PAC: proposed avoidance control)

When O_5 was in the shoulder avoidance area, the base was corrected in the $+x_0$ direction orthogonal to the $+y_0$ direction, which is the base progressing direction. Thus, O_5 was also corrected in the same way. The closer O_5 was to the shoulder singularity, the larger the avoidance value was. The farther O_5 was from the shoulder singularity, the smaller changing value was. After leaving the avoidance area, the base returned to the original tracking trajectory. By applying the corrected target value to calculate inverse kinematics, the correcting value of the first, fourth, fifth, and sixth joint angles was suppressed. As a result, the robot avoided the shoulder singular configuration, so that the robot could avoid emergency stopping. Therefore, the proposed control method is effective.

5. Conclusion

In this study, a novel control method combined with PAWTED was proposed in order to resolve the emergency stopping of a robot caused by a shoulder singular configuration throughout teaching. Through simulation results, it was confirmed that the proposed method can avoid the above problem. Experiments to demonstrate the solution are planned.

Acknowledgments

This research was supported by Aichi Prefectural Government through “Knowledge Hub Aichi,” Priority Research Project (II).

References

1. T. Tsuji, T. Shimazaki and M. Kaneko, Analysis of human perception ability for robot impedance, *J. Robotics Society of Japan*, **20**(2) 2002, pp.180-186.
2. M. Sakai, N. Sato, and Y. Morita, Development of parallel wire mechanism for robot teaching, *J. Japan Society of Mecha. Engineers, Series C*, **79**(807) 2013, pp.320-328.
3. M. Sakai, N. Sato, and Y. Morita, Performance Evaluation of Robot Teaching Method with Parallel Wire Mechanism, *J. Society of Signal Processing Applications and Technology of Japan*, **16**(1), 2013, pp. 72-78.
4. N. L. Minh, H. Inuzuka, T. Nakanishi, N. V. Quang, Y. Morita, M. Sakai, K. Hashimoto, Y. Suzuki, Y. Kabeya, N. Suzuki, Y. Sakurai, and T. Fujiwara, A graphical simulator for teaching / playback motion of a serial robot with Parallel Wire Type Teaching Device, *Proceeding of the 2017 International Conference on Mechanical, Electrical and Medical Intelligent System*, Kiryu, Gunma, Japan, Nov. 2017, GS-07.

Design of Humanoid Soccer Robot Based on STM32

Yuheng Zhang[1], Yulong Peng[2], Yizhun Peng [1], Lianchen Zhao[1], Zhou Zhang[1], Wanlong Peng[1]
[1]Tianjin University of Science and Technology, China
[2]Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China
*E-mail: * pengyizhun@tust.edu.cn*

Abstract

Humanoid robot is an important branch in the field of robotics, because its shape and movements are similar to human beings, it provides a good carrier for the related research in the field of artificial intelligence. Humanoid soccer robot of this paper is designed based on STM32 processor. Through theoretical analysis and experimental verification, the stability, working parameters and the division and cooperation of STM32 and 51 series single-chip computers of humanoid soccer robot system are analyzed, and determine the final control scheme. STM32 drives the camera to collect information, process images and make decisions. STC12C5A60S2 controls the robot steering gear to complete the corresponding action. Serial communication is used between the two controllers. The peripheral circuit mainly includes OV7725 camera module and TFT-LCD LCD display module. In order to recognize and approach soccer by humanoid soccer robot and kick soccer into the goal, we should design the action programs such as straight walking, left turning, right turning, head scanning course, left translating, right translating and kicking. Reasonable images processing and decision algorithms are also designed. Finally, the control program is developed on Keil software platform for debugging. The experimental results show that the humanoid soccer robot designed in this paper can accurately identify the soccer ball and the goal, and accurately complete the kicking actions, which meets the relevant requirements of the expected design and scheme.

Keywords: Humanoid robot;Dual controller; erial communication; Images processing

1. Introduction

As a research product of artificial intelligence in the field of robotics, humanoid robots are the frontier and hotspot in the field of robotics. As a member of the humanoid robot family, soccer robot is a dynamic experimental platform with high degree of confrontation[1]. Humanoid robots can imitate human thinking mode to perceive the characteristics of the external environment and make judgments, and make corresponding actions to complete a task[2]. The soccer robot designed in this paper is designed according to the human body shape. It has a head, two arms and two legs, and can stand on both feet and walk on both feet. Robots can perceive the external environment through sensors equivalent to human perception, and sensors must be installed in a position similar to human perception. At the same time, the robot can recognize orange, white and green targets with different colors, and can track targets with single color and regular shape. According to the design requirements, the soccer robot must be able to complete the basic movements such as straight line, left turn, right turn, kicking the ball, and can correctly recognize the soccer, goalkeeper and goal through machine vision, and the distance between the robot and the ball is 0.5 meters. Robots need to switch approaching football and shooting programs independently, and according to the requirements of the field, attack and defense processes must use a set of procedures, no midway switching procedures are allowed.

2. System scheme design

This design adopts the control scheme of double controller, and the hardware mainly includes the body, steering gear, controller, camera and display screen of the humanoid soccer robot. STM32 is used as the main controller to realize the image acquisition and processing function. STC12C5A60S2 is used as the slave processor to control the robot to complete the corresponding actions. The serial port is used for communication between STM32 and STC12C5A60S2. The flow chart of STM32 and STC12C5A60S2 is shown in Figure. 1.

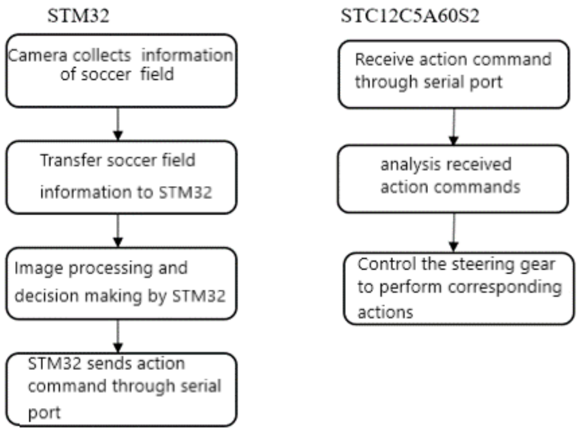


Fig.1. The flow chart of STM32 and STC12C5A60S2

3. Design of robot hardware system

3.1. Robot structure design

This design uses the SHR-8S humanoid robot of Beijing Senhan Technology Co., Ltd. According to the relevant requirements of hardware, the main size of the robot is designed to be 360 mm high. Shoulder width 165 mm, chest thickness 115 mm, mass 1.5 kg. Components include head, arm, hip joint, thigh and so on, belonging to small humanoid robot. In this design, another steering gear is installed on the head, which combines the two degrees of freedom of the head itself, so that the information of the soccer field can be collected quickly. The position of 17 + 1 steering gear is 2 at the head, 2 at the left and right arms, 1 at the left and right shoulders, 1 at the left and right hip joints and 4 at the left and right legs respectively[3] .The design of the structure is shown in Figure.2.

Fig.2. Robot picture

3.2. Introduction of controllers and peripheral modules

The main controller used in this design is STM32F103, 32-bit microcontroller, based on ARM Cortex-M3 core, LQFP-144 package. Compared with similar controllers, CPU can process data quickly and support multi-path execution of data instructions. The CPU operating voltage range is 2.0-3.6V. With external button battery, it can realize power-off data storage. This chip has very high performance, excellent real-time performance, reasonable and rich peripherals, low power consumption, low price, rich software support package[4], rich and comprehensive technical information, easy to develop, which is very suitable for this design.

The slave controller selected in this design is the control board with STC12C5A60S2 single chip microcomputer as the core. Most of the components adopt the full surface mounting technology. The design of the control board is reasonable and powerful[5]. All the external interfaces of STC12C5A60S2 series single chip microcomputer are expanded for convenient use.

Fig.3. STC12C5A60S2

The camera selected in this design is OV7725 camera, 300000 pixels, high frame rate, up to 120 frames per second [6], high-speed dynamic, can quickly capture position information, suitable for robot to capture the position of football. The data format of the output image supports YUV (422 / 420), yCbCr422 and RGB565. The output data format identified in this design is RGB format, which is convenient for feature extraction in the future[7]. The data collected by the camera is cached in the FIFO cache on the back of the camera, and then the external devices obtain the captured image data through the signal pin below.

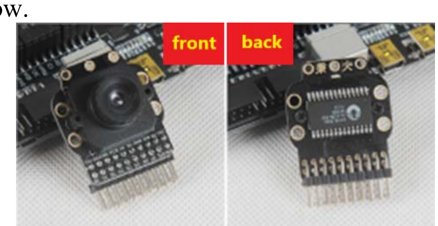


Fig.4. OV7725 camera

The display screen is used to display the information collected by the camera, determine the effective collection range of the camera, facilitate the programming and debugging, and use the display screen as the calibration reference medium to determine the action that the robot will perform according to the different display positions of the football on the display screen, so as to complete the task of tracking the football and shooting. The display screen selected in this design is the 3.2-inch TFT LCD screen .This screen has the advantages of light weight, high reliability, low driving voltage and much less power consumption than CRT screen, which saves a lot of energy [8].

Fig.5. LCD screen

4. Analysis of football recognition and control program

4.1. Soccer recognition

RGB, YUV and HSV are commonly used color spaces for robots [9]. RGB is the physical primary color, Y is the brightness signal, U and V are the chroma signal, H is the hue, S is the saturation and V is the brightness, the conversion formula is as follows:

$$\begin{cases} Y = 0.299R + 0.587G + 0.114B \\ U = -0.147R - 0.289G + 0.436B \\ V = 0.615R - 0.515G - 0.1B \end{cases} \quad (1)$$

$$V = max(R,G,B) \quad (2)$$

$$S = \begin{cases} \frac{V - \min(R,G,B)}{V} & if\ V \neq 0 \\ 0 & if\ V = 0 \end{cases} \quad (3)$$

$$H = \begin{cases} \frac{60(G-B)}{S} & if\ V = R \\ 120 + \frac{60(B-R)}{S} & if\ V = G \\ 240 + \frac{60(R-G)}{S} & if\ V = B \end{cases} \quad (4)$$

RGB model is seldom used in artificial intelligence because it is easy to be affected by the light. However, the conversion from RGB space model to YUV and HSV model involves a lot of calculation, which will be a great burden on the controller of soccer robot and seriously affect the real-time performance of robot. For the robot recognition, there are only three colors: green, orange and white. The color is less and the shape of the recognition object is regular. In RGB color space, the recognition effect is not greatly interfered by experiments under different light intensity. Considering the relatively simple calculation of RGB model, the RGB color space is selected in the design.

The image collected by the camera is an RGB color image, which can be recognized from the image by STM32 image binarization method, and then the subsequent decision-making tasks can be completed. Through repeated experiments under different light conditions, the threshold value of image segmentation is determined, and then the image is segmented into background and target, achieving good experimental results. The comparison before and after processing is shown in Figure 6.

Fig.6. Original graph and binary partition graph

4.2. The design of robot decision

In this design, LCD screen is used to display the image after binary image processing, and LCD screen is used as calibration reference, that is to say, it can display 24 lines and 240 columns of pixels in the middle part collected by the camera. According to the position of the target displayed on the LCD screen to determine the next action to be performed by the soccer robot, figure 7 is a schematic diagram of the LCD screen area with 24 lines and 240 columns of pixels, and the number at the bottom of the figure is the number of columns of pixel points.

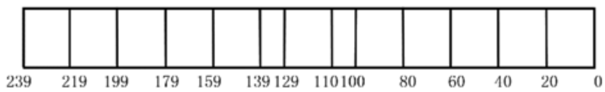


Fig.7. Schematic diagram of the LCD screen

In the process of approaching a soccer, in a line of pixels: if the pixels on each column marked with a number are white, that is, the robot does not see the soccer, the robot heads up; if the pixels on each column marked with a number are black, the robot sees the soccer, and the soccer is in front of the robot, the robot lowers its head and goes straight for one step; if the pixels in the 129 and 110 columns are all Black, no matter whether the other columns are black or white, it is considered that the soccer is in front of the robot, and the robot lowers its head and goes straight for one step; if only the pixels in column 0 are black or the pixels in column 0 and column 20 are black, and the soccer is on the left side of the robot, the robot turns left; if only the pixels in column 239 are black or the pixels in column 239 and column 219 are black, and the soccer is on the right side of the robot, the robot turns right.

From approaching football decision-making program to shooting decision-making program: the robot camera rotates upward along the x-axis from the initial position until the football is detected, and the rotation angle is set as α. As shown in Figure 8, when α ≥ 2.22 °, the approaching football decision-making program is executed; when α < 2.22 °, the foot ball has reached the foot of the robot, and the shooting decision-making program is executed.

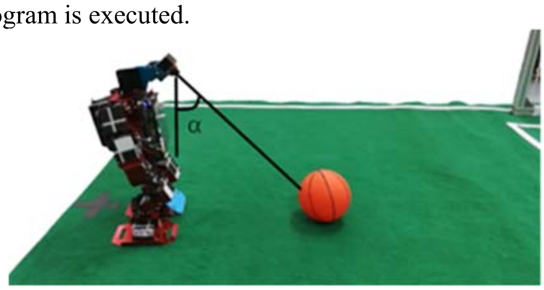


Fig.8. Definition of α

In the shooting decision-making process, because the goalkeeper robot is generally standing in the middle of the goal, the robot shooting idea of this design is to shoot left or right to avoid the goalkeeper. The robot heads up and turns left to detect the left edge of the goal line, as shown in Figure 9. When more than nine columns of pixels marked with numbers on the LCD screen are black, it is deemed that the goalkeeper is not in front of the left edge of the goal line, and the robot shoots at the left edge, otherwise, it shoots at the right edge of the goal line.

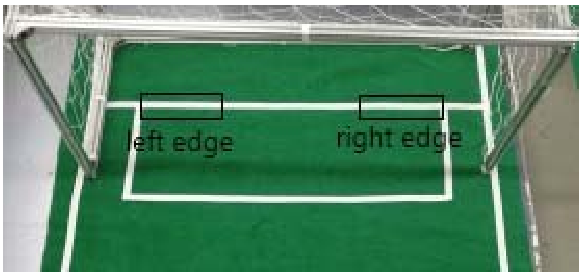


Fig.9. Goal edge

5. Conclusion and future work

Based on STM32F103 single chip microcomputer, this paper studies the humanoid soccer robot system, designs a threshold binary image processing algorithm and soccer tracking and shooting program, and applies the algorithm and program to build the hardware system of humanoid soccer robot, which can recognize and track the objects with single color and regular shape. And completed the football tracking and shooting tasks well. The hardware of the whole system is simple and clear, the execution efficiency of the software is high and the logic is strong.

Although the design meets the design requirements, due to the limited level of the author and the limitation of research time and test conditions, the following improvements are needed for the design and research of humanoid soccer robot:

The height of the soccer robot must be increased, and the shooting speed must be accelerated to meet the requirements of the next stage of the game, that is, to complete the air ball shooting, so the driving source of each joint must have greater torque, and the robot system must make some improvements in the control of the steering gear with large torque. At present, the soccer robot is not good at shooting angle control. The solution is to add gyroscope to the robot body, which makes the robot control system add gyroscope interface. This design only completes the design requirements in the simple competition environment, but also needs to study the soccer recognition, tracking, shooting and other issues in the more complex environment.

References

1. Seekircher A, Laue T, Rofer T, et al. RoboCup 2010: Robot soccer world cup XIV[C]//Springer-Verlag,2011:637-644.
2. Nagatani K,Kiribayashi S,Okada Y,et al.Emergency response to the nuclear accident at the Fukushima Daiichi Nuclear Power Plants using mobile rescue[J].Journal of Field Robotics,2013,30(1):44-63.
3. Ke Liu. Research and implementation of target recognition algorithm and action planning in humanoid robot [D]. Hebei: Yanshan University, 2011: 7-8.
4. Tian Deyong. PWM experiment teaching based on STM32 timer [J]. Electronic technology and software engineering, 2019 (20): 102-103.
5. Yuan Yongchao, Deng jiongfeng. Research on automobile anti-theft device based on STC12C5A60S2 [J]. Computer products and circulation, 2019 (07): 178.
6. Zhimin Ren. Design of an image real-time display system based on OV7725 [J]. Fujian computer, 2013,29 (01): 128-130.
7. Caixia Zhao. Research on target dynamic tracking based on STM32 mobile robot [D]. Inner Mongolia: Inner Mongolia University, 2017.

8. Xiaoliang Chu. Analysis of TFT-LCD technology development trend [J]. Electronic world, 2018 (20): 73-74.
9. Wenping Xi, Jian Zhang. Design of vision recognition system for humanoid soccer robot based ARM [J]. Microcomputer and application, 2015,34 (09): 51-53.

A Design and Implementation of Intelligent Cradle

Ting Zhao[1], Qing He[1*],Yulong Peng[2], Zhou Yang[1],Zhenjiang Chen[1],Shuo Jiang[1]
1 Tianjin University of Science and Technology, Tianjin, 300222,China;
2 Fenyang College of Shanxi Medical Univercity,Fenyang Shanxi, 032200 , China
*E-mail: * heqing@tust.edu.cn*
www.tust.edu.cn

Abstract

The equipment is designed for families to raise children, the device is a cradle of artificial intelligence technology. The data of baby shaking are collected and analyzed by the sensor, so that the cradle bed can bionic cradle shaking. The SVM database training of infant crying can realize the recognition of infant crying and determine the specific meaning of crying, such as hunger, excretion, pain and other factors, and timely inform parents; Based on ROS robot operating system and iflytek platform, it can carry out natural language interaction and autonomous navigation. Parents can call the baby cradle and let it reach the designated position automatically to achieve autonomous obstacle avoidance and path planning without manual interference. At the same time, the camera can view the baby's status and transmit the real-time picture to the mobile phone APP.

Keywords: bionic cradle, infant crying, ROS, iflytek, platform, natural language interaction, autonomous navigation

1. Introduction

Since entering the 21st century, the development of robots has become faster and faster. With the development of science and technology, the integration degree of robot is higher and smaller. The diversified functions of service robots are gradually recognized by the market. According to China's industrial information network, the sales volume of service robots in 2016 was 6.8 million, an increase of 25.93% compared with 2015. From 2011 to 2014, the global average compound growth rate of home service robots is about 22%, and the sales volume is expected to increase to 31 million in 2019. Service robots will affect all aspects of human production and life.

In today's information age, people are busy in various activities and affairs, such as work, social interaction and so on[1]. According to the statistics of the United Nations, 140 million newborns were born in the world in 2018, with 15.23 million in China alone. With the opening of the two child policy, the number of newborns is increasing year by year[2]. A large group of novice parents urgently need a more labor-saving care service robot[3]. In order to solve the problem of parenting, we have developed an intelligent cradle, which can imitate the cradle frequency of parents, understand the crying of babies, and move autonomously, so that novice parents are more relaxed.

2. System overview

This equipment is a cradle bed combining artificial intelligence technology. The cradle bed can realize many functions such as automatic follow, intelligent cradle, natural language interaction, early childhood education, cry recognition, etc. Cradle bed integrates speech recognition and processing and big data processing technology. By installing ROS platform in raspberry pi, its distributed framework can build voice interaction module, bottom motion module, Bluetooth communication module, etc. Three special functions can be realized: first, imitating the swing frequency when human coax the baby, automatically rocking the cradle with different degrees of bionic frequency, imitating grandma gently rocking, imitating mom to coax the baby to sleep rocking, imitating dad to play rocking, and at the same time, timing the cradle time;Second, autonomous navigation, based on slam technology, allows the bassinet to move in a strange environment, quickly build the layout of the environment, provide accurate positioning and navigation reference for the bassinet autonomous movement, as well as human-computer interaction reference. With the functions of real-time positioning, intelligent obstacle avoidance, path planning, etc., the device can move to any indoor and reachable place autonomously, realize fixed-point movement, and be autonomous in the whole process without any human intervention. Third, it is equipped with a voice interaction system, based on iFLYTEK platform, and at the same time collects big data of baby crying, which can communicate with parents voice, recognize baby crying, and intelligently judge baby crying due to hunger, excretion or pain and other factors, so as to inform parents in time; voice interaction can realize parents calling cradle car to self navigate to parents.

3. The hardware structure design

The core platform of the device is the robot operating system (ROS). Its distributed network uses the communication mode of TCP / IP to realize the point-to-point loose coupling connection between modules. The hardware consists of raspberry pi, camera, lidar, ultrasonic ranging sensor, motor, omni-directional wheel, encoder and drive board. The encoder is installed on the cradle motor to calculate the rotation angle of the cradle and fix the rotation angle from the mechanical structure and the upper limit of the software program; the drive board controls the movement of the baby carriage chassis, the chassis adopts three omni-directional wheels, which can realize the omni-directional movement ability of the baby carriage; the microphone recognizes the baby crying for data collection and processing, and uses iFLYTEK platform for natural language Interaction; the camera is used to observe the situation in the stroller and transmit the real-time picture to the mobile app; the ultrasonic ranging sensor is installed around the stroller chassis to avoid obstacles; the lidar is used to collect the surrounding environment information, and the collected data is transmitted to the raspberry pi for the slam drawing, so as to realize the indoor autonomous navigation.

3.1. Chassis motion control structure

The bassinet chassis is made of aluminum alloy. Because the bassinet needs to have safe omni-directional movement ability in the room, the omni wheel ql-10 (Omni wheels is a wheel that can move in many different directions) is selected as its moving tool. It includes a wheel hub and a plurality of driven wheels. Each driven wheel is embedded with two rolling bearings. Compared with the mcnamm wheel, the omni-directional wheel has faster speed, more flexible movement mode and more stable movement system. When three omni-directional wheels are used at the same time, the translation in any direction can be realized, and any complicated arc motion can be performed. The driving mode is three-wheel independent driving, three hollow cup DC servo motors control three omni-directional wheels respectively, and use imdr4e servo driver to drive the motor via CAN bus and RS232 communication interface. The maximum moving speed of bassinet is 0.65m/s, the maximum rotation speed is 200 ° / s, and the minimum turning radius is 0m. The chassis is slightly lower, so it is suitable for indoor and flat road operation. For the safety of the baby, crash strips are installed around the chassis. The chassis structure is shown in Fig.1.

Fig.1. Chassis structure

3.2. Cradle mechanical structure

The bassinet structure is arranged on the bassinet chassis, which is composed of four brackets, an electric motor, a rocker and a bassinet. Two brackets are set up on both sides to fix on the chassis, one side of which is fixed to install the motor, the motor is connected with the rocker, the other side is directly connected with the rocker, and the two rocker are connected with the cradle. Through the motor to give the rocker force, through the rocker swing to drive the cradle swing. The mechanical structure of cradle is shown in Fig.2.

Fig.2. Cradle rocker structure

The motor is connected by the rocker. The motor rotates to drive the rocker to swing back and forth. The rotation angle and speed of the motor are calculated by the encoder. The control board yf00702ea controls the maximum rotation angle of the motor to prevent the shaking from being too intense. At the same time, in the mechanical structure, the shaking limit position is designed to prevent the parents from throwing the baby out by mistake. The cradle motor and encoder are shown in Fig.3.and the control board YF00702EA is shown in Fig.4.

Fig.3. Motor and encoder

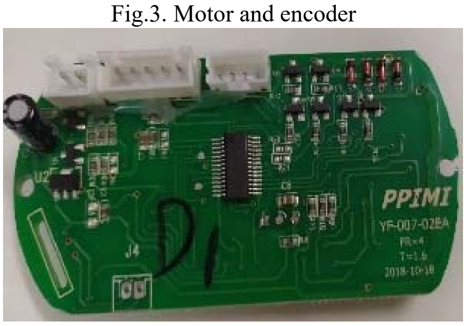


Fig.4. Control board YF00702

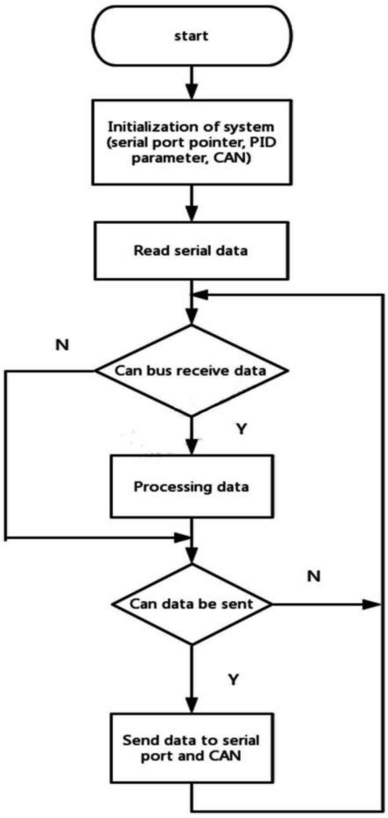


Fig.5. Servo driver workflow

3. 3. Ultrasonic obstacle avoidance sensor

The model of ultrasonic obstacle avoidance sensor used in this bassinet is ULB-1 ultrasonic distance sensor, which has the characteristics of high resolution, high precision and low consumption. Not only in the design, but also in the interference noise processing, with anti noise interference ability. And for the different size of the target, and the change of the supply voltage, do the sensitivity compensation. In addition, it also has standard internal temperature compensation, which makes the measured distance data more accurate[4]. ULB-1 ultrasonic ranging sensor is shown in Fig.6.

Fig.6. ULB-1 Ultrasonic ranging sensor

3. 4. Laser radar

Lidar systems are often based on pulsed laser diodes and silicon avalanche photodiode (APD) arrays, with exploiting 905nm wavelength light[5].It is used to collect the information of the surrounding environment, transmit the collected data to raspberry pi, and use it for SLAM mapping to realize the indoor autonomous navigation, as shown in Fig.7.

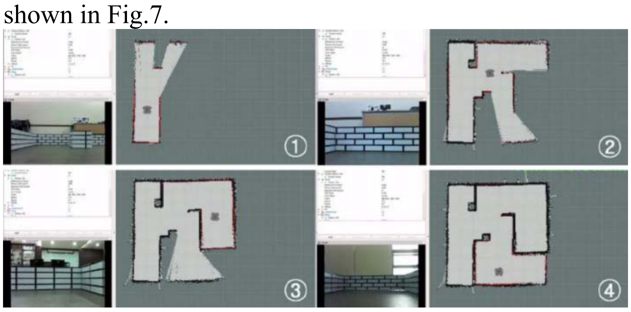


Fig.7. Experimental construction

4. System circuit module design

4.1. Research and analysis of CAN bus

The Controller Area Network (CAN bus) is a bus based on differential signaling originally developed for automotive industry[6].CAN bus topology and physical layer, can bus bit value representation is shown in Fig.8.

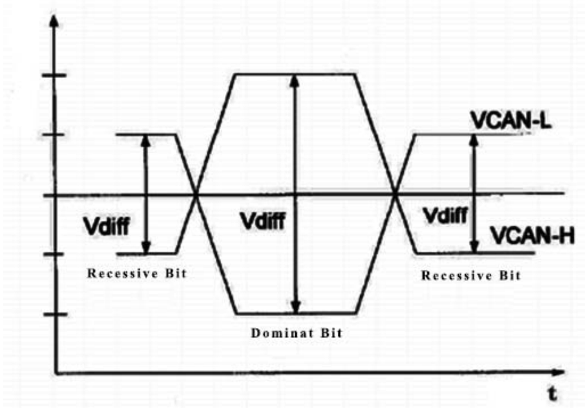


Fig.8. CAN bus bit value characterization

Compared with other network protocols, CAN has the following two outstanding advantages:

- high reliability. The short frame mode is selected for data transmission, and the transmission medium is double insulated wire. CAN can effectively shield all kinds of electromagnetic interference outside. At the same time, CAN is defined with a variety of fault diagnosis mechanisms, which makes it highly reliable and very suitable for use in the controller subnet[7].
- good expansibility.CAN adopts multi main mode communication, and the identifier is defined in the frame structure. The access of controller in the network does not require any change of software or hardware of all controller application layers.

4.2. Crying recognition

From the "donate a cry" project shared on GitHub, all the infant crying data recorded and uploaded by the user's mobile phone were screened to obtain 450 clear infant crying data with a duration of about 5 s. The data has been converted to CAF or 3gp with a uniform bit format of 128kbps, and all the data used has been converted to WAV format with a sampling rate of 8 kHz.

The feature extraction method of baby crying: when baby crying, the detection device (microphone) will receive the signal, and then through filtering processing, and then through the basic acoustic features for feature extraction, get the feature data[8]. Infant cry can be divided into several crying units, which have different acoustic characteristics. The flow chart of cry recognition system is shown in Fig.9.

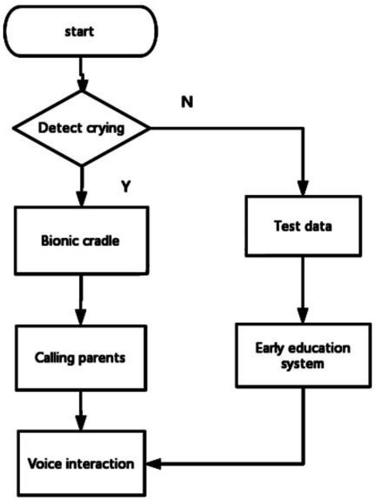


Fig.9. Flow chart of cry recognition system

As shown in Fig.10., there are three expiratory crying units (segments 1, 3 and 5) with longer duration and lower fundamental frequency, and two inspiratory crying units (segments 2 and 4) with shorter duration, dullness and higher fundamental frequency The duration of inspiratory crying unit is short, and sometimes it is pure sound, sometimes it is dullness, and its acoustic characteristics are not stable, so breath crying unit is used for analysis.

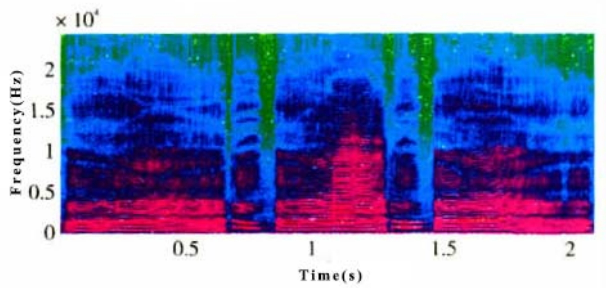


Fig.10. Two microphone array for corpus sample analysis

The crying type expiratory crying unit calculates the fundamental frequency (F0) and the first (F1), second (F2) and third (F3) formants of the crying signal frame by frame The fundamental frequency corresponds to the frequency of the glottal pulse excitation signal, and the resonance peak corresponds to the resonance frequency of the channel.

The device uses SVM cry recognition model, uses tensorflow in-depth learning keras tool kit to realize SVM model, Support Vector Machine (SVM) is a supervised machine learning algorithm. Firstly, preprocessing and feature extraction are carried out, including removing silent segment, adding window, framing and amplitude normalization. The feature vector extracted by SVM experiment is the statistical average feature of all frames in each corpus.

For the identification of this product, we use the kernel function with the highest recognition rate in the laboratory process, which is Gaussian kernel function, as the SVM baseline model.

5. Testing and conclusion

5.1. Test plan

The test scheme of omni-directional wheel is as follows:

- Write the control program of the omni-directional wheel, and then place it close to the longer wall and move it freely, so that it can directly observe the change of the gap between the intelligent bed and the wall, and record the error.
- Place the intelligent bed on the flat and easy to observe ground, determine its geometric center and mark it on the ground, make it rotate around the center point, observe the bed offset after a period of time, and record.
- Place the intelligent bed at the right angle bend, make sure the ground is flat and smooth, make it pass the right angle bend continuously, observe the friction and deviation, and record.

The test scheme of the shaker is as follows:

- Set the intelligent bed to a fixed gear, record the included angle and time, calculate the shaking frequency, record a long time, and calculate the error.
- Put the intelligent bed in different gears, and compare the changes of amplitude and frequency.

The scheme of ultrasonic obstacle avoidance test is as follows:

- Place the intelligent bed in the open and flat space, set obstacles of the same specification in the surrounding fixed area, so that the bed can only move towards the direction of the obstacles, return to the original place after the successful detection of the obstacle avoidance, and then move to other directions to judge the success probability and efficiency of the obstacle avoidance.
- Change the shape of obstacles, and then conduct obstacle avoidance test. Record data and analyze.

Bluetooth Test: transfer data of different types and sizes into Bluetooth at one time, and judge whether Bluetooth is in fault by many experiments.

The cry recognition test scheme is as follows:

- Put the cry data collected from the same infant with different purposes into the in-depth learning module, and compare the results of learning prediction with the actual situation to get the probability of success.
- Put the cry of different infants into the deep learning module to detect the success rate of deep learning.

5.2. Test results

After repeated tests, the crib can meet the needs of users for safety and intelligence, achieve the real sense of "people-oriented"[9], and protect the baby from harm in the daily life environment, with stable hardware and software. Omnidirectional wheel works safely and normally, avoiding obstacles accurately and timely. Bluetooth receives all kinds of data and works normally. It also has a certain recognition success rate for children's crying, which can help parents understand their children's physiological needs.

References

1. Aiyan Wu, Xuewei Cui. Research progress of home robots [J]. Journal of Tangshan University, 2009,22 ，3，P69-P71
2. Binkun Liu,,Haitao Liu,Design of Intelligent Security stroller for data Communication based on STM32,Science and Technology & Innovation,2018,16,P62-P63.
3. Yaoyao Zhang, Shuwan Lei ,Jiayi Zhou ,Gejile Hu . Design and exploration of automatic swing baby cradle [J]. Electronic production, 2013 ,9,P26
4. Razvan Luca, Fritz Troester, Robert Gall and Carmen Simion,Autonomous parking procedures using ultrasonic sensors,Annals of DAAAM & Proceedings,Jan. 1, 2010 ,P691
5. Extance, Andy. "Powering cars of the future: Andy Extance asks: which laser wavelengths, detector architectures and ranging techniques are required to deliver competitive and cost-effective automotive lidar performance?" Electro Optics, no. 296, 2019, p. 14+. Gale Academic Onefile
6. Dosek, Roman, et al. "Secure high level communication protocol for CAN bus." Annals of DAAAM & Proceedings, 2015, p. 1009+. Gale Academic Onefile.
7. Yuan Gao, Xiu Wang,Shuo Yang,Yuanyuan Zhai, et al. Study on monitoring and evaluation system of sowing depth based on CAN bus [J / OL]. Journal of agricultural machinery: 1-12 [2019-11-21]
8. Alpika Tripathi, Geetika Srivastava, K.K. Singh and P.K. Maurya,Epileptic Seizure Data Classification Using RBAs and Linear SVM.Biomedical and Pharmacology Journal(Vol. 12, Issue 2.)June 1,2019.
9. Guangtao Ma,Shiming Zheng,Safety Design research of stroller,China Academic Journal Electronic Publishing House,2016,11,P126-P127.

A Design and Implementation of Intelligent Networking Bookcase

Zhou Zhang[1],Yajun Li[1*],Yizhun Peng [1,2], Yuqi Zhao[1], Hucheng Wang[1]

[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: lyj _2872@sohu.cn

Abstract

The device is designed for the effective use of book resources. Users can share books in their hands through the device. Multiple devices connected together through wireless network can form a book storage network to realize the flow of paper books. The camera is used for the ISBN code identification of books, the WiFi module is used for the wireless network communication between the device and the background server, the bluetooth module is used for the near distance communication between the device and the mobile terminal, and the infrared module is used for detecting the storage state of the device. Equipment ISBN read, storage of equipment state inspection, such as simple data processing done by the device with a built-in MCU, user registration, user reading habits, such as book search more complex data calculated by the background server processing,.By borrowing the concept of distributed computing, implements the equipment, and efficient use of server resources, reduce the workload of the entire system at the same time.

Keywords: book, wireless network, WiFi, bluetooth, ISBN, MCU, background server

1. Introduction

With the continuous popularization of Internet technology, the exchange and transmission of information is more convenient and fast, and electronic reading develops accordingly. Compared with the traditional paper reading electronic reading is convenient and quick, and easy to obtain, and the advantages of easy to carry, but electronic reading is more fragmented reading, digital media in order to gain the reader's preferences, with simple, rapid, shallow and other information to bo person eyeball, ideological content of the article without giving carefully chosen, causes the reader too pursuit of reading speed and the number, the lack of deep thinking[1]. Compared with electronic reading, paper reading can stimulate readers to explore in a deeper level. At the same time, paper reading has many advantages, such as less harm to eyesight, greater appreciation space and lower error rate. So in today's advanced science and technology, we still need to promote the traditional way of reading - paper reading.

However, while promoting paper reading, some problems existing in paper books should not be ignored. These problems mainly include high price of paper books, taking up more storage space, and not easy to carry. Therefore, this paper designed a kind of Shared bookcase that can be placed in the community, school or shopping mall and other public areas, so that users can not only share books they no longer use but also borrow books for reading. Thanks to the rapidly developing communication technology, multiple Shared bookcases can be connected through the network to form an extensive book network.

2. Embedded hardware system design

The embedded hardware system adopts STM32F103 series single-chip microcomputer produced by stmicroelectronics as the main control chip. In addition, bluetooth communication module, DTU module, photoelectric switch module and beidou positioning module are integrated around the main control chip. Among them, bluetooth module is used for communication between embedded device and mobile phone client, DTU module is used for communication between embedded device and background server, photoelectric switch module is used for detecting the current storage state of storage unit, and beidou positioning module is used for book case positioning.

2.1 main control chip

The core microcontroller of embedded hardware system is a STM32F103ZET6 chip with the kernel of ARM cortex-m3, introduced by stmicroelectronics, whose working frequency is up to 72MHz. On-chip integration up to 512KB Flash memory and 64KB SRAM memory; up to 9 communication interfaces, including 2 I2C interfaces, 3 USART interfaces (support ISO7816 interface, LIN,IrDA interface and modem control), 2 SPI interfaces, USB 2.0 full speed interface; up to three synchronized 16-bit timers, each with up to four channels for input capture/output comparison /PWM or pulse counting, and up to 140 I/O ports with interrupt capabilities. Compared with other MCU has high performance, low cost, low power consumption, can fully meet the system requirements[2].

Fig.1. STM32F103ZET6 chip

2.2 Bluetooth communication module

Bluetooth communication is a short-range wireless communication technology. Its initial goal is to replace cable connections on existing digital devices, such as handheld computers and mobile phones. Bluetooth technology is characterized by small volume of communication module, low energy consumption, especially suitable for remote, high speed transmission of single-chip microcomputer system equipment. Bluetooth communication transmission distance is usually less than 10 meters, the use of bluetooth technology can be realized by short-range wireless link instead of communication cable, not only to avoid the trouble of wiring, but also to facilitate wireless control of equipment, there is a very wide application space. Bluetooth receiver integrated circuit consists of rf, baseband and link management. In the terminal of MCU, there are four terminals, two ports are connected with power supply, RxD in communication port is connected with TxD in MCU, TxD in communication port is connected with RxD in MCU, that is, reading and writing terminals should be cross-connected[3].

Fig.2. Bluetooth communication module

2.3 DTU module

DTU(Data Transfer unit) Data transmission unit is a wireless terminal device specially used to convert serial port Data into IP Data or IP Data into serial port Data through wireless communication network for transmission[4].

Fig.3. DTU module

2.4 Photoelectric switch

The principle of photoelectric switch is the same as that of infrared detection, which belongs to active infrared detection. The photoelectric switch module emits pulse infrared light at a specific frequency, and the infrared photosensitive receiving tube is used to receive the reflected light ahead, and the distance of the obstacle is judged according to the intensity of the received light. Because of its simple structure and low cost, it is often used for non-contact distance measurement, obstacle detection, obstacle color detection and other devices[5].

Fig.4. Photoelectric switch

2.5 Beidou positioning module

Beidou navigation system is an active bidirectional ranging 2d navigation system. The ground center control system is solved to provide users with 3d positioning data. Besides the function of GPS satellite positioning, it also adds the communication function, which can play a great role in some fields such as ocean monitoring and military communication. Beidou is the world's first satellite navigation system to provide three-frequency signal services. The three-frequency signal can better eliminate the influence of high-order ionospheric delay and improve the positioning reliability. Besides, beidou also has its own special short message communication function[6].

Fig.5. Beidou positioning module

3. System software design

The system software layer design is divided into three parts: client end, embedded end and background server end. The communication between the client and the embedded end is conducted through bluetooth, and the communication between the client and the background server and between the embedded end and the background server is conducted through the public mobile communication network.

3.1 Client software design

The mobile client takes APP as the carrier to provide the operation of saving and borrowing books, Its program flow chart is shown below.

When using the APP, users must scan the qr code pasted on the box to confirm the current book case number. The content of the qr code is the 48-bit MAC Address (Media Access Control Address) of the bluetooth module of the book case, which is globally unique and can be used as the number of different book cases[7]. After obtaining the case number, the user can store or borrow books according to the instructions of the software interface.

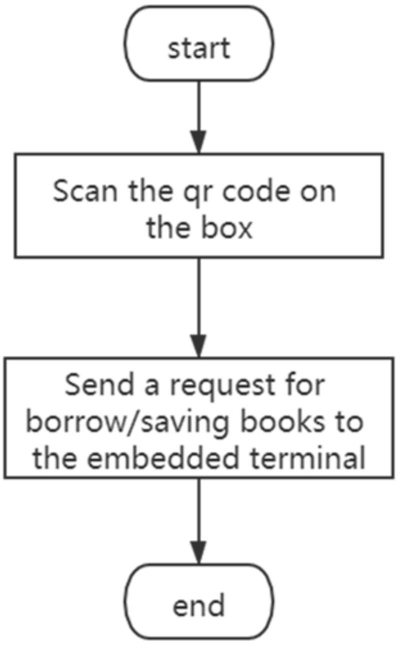


Fig.6. Client program design flow chart

3.2 Embedded end software design

The software design of embedded device includes the communication between the embedded device and the mobile phone client, the control of embedded device on the bookcase storage unit, and the detection of embedded device on the storage unit when storing books.

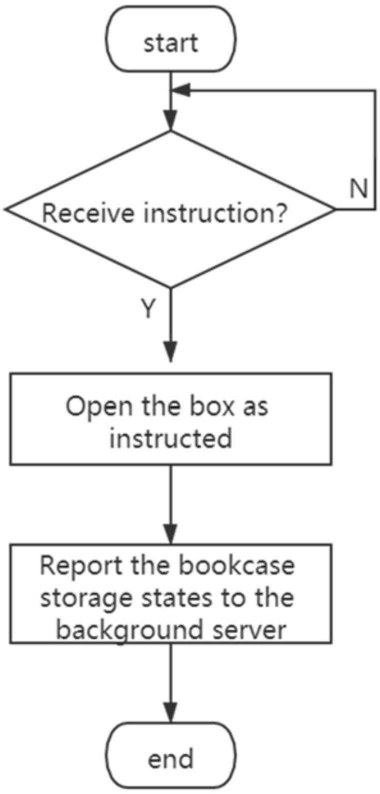


Fig.7. Embedded programming flowchart

3.3 Background server software design

The background server includes two parts, the server and the database, which are respectively used to realize the data interaction with the client and record the information such as the bookcase storage. Its program flow chart is shown in the figure. Whenever the storage status in the bookcase changes, the information in the database can be updated in real time.

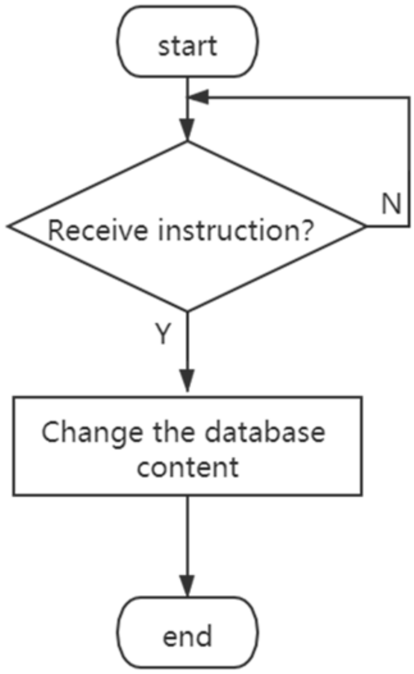


Fig.8. Background server programming flow chart

4. Conclusion

With the development of modern lifestyle, people's pace of life is speeding up, so it is particularly necessary to set aside more space for bookstores in the city's diversified space. Need to see, the flow of people in business circles, but the noisy environment is not suitable for deep reading after all. And schools , parks, hotels, and other places have leisure, knowledge features, and reading is very matching. Break down the mechanism barrier, update the business form, let the books naturally into these places, holding books quiet cultural street scene will be more and more. In the case of Beijing, the Palace Museum has added 6 exquisite bookstores in recent years, with flowers flourishing and a room full of books. Some hotels have also made the switch from "walking in" to "staying" and "reading" by opening book corners. Facts show that as long as a lot of support, do the characteristics, reading can completely in the continuous "break the circle" to achieve diverse development[8].

The mobile book network designed by this paper is based on the above reasons. Through the application of Internet of things technology, books can be Shared among readers, so that paper books can give full play to their value and readers can read paper books more conveniently and quickly.

Acknowledgements

This research was supported by Student's Platform for Innovation and Entrepreneur Training Program, the Ministry of education of China(201910057158).

References

1. Guo xiaona, huo xiaoqin. Deep reading: knowledge creation in the era of artificial intelligence [J]. People's education,2019(18):62-65.
2. Chen xinfen, peng yinqiao. Intelligent home control system based on STM32 and ZigBee [J]. Electronic technology and software ngineering,2019(21):16-18.
3. Bo likang, zhang haiyan, duan weiyang. Application of wireless communication technology in MCU ommunication system [J]. Digital communication world,2019(09):158.
4. Deng lei. Application of wireless system based on DTU and GPRS technology in ceramic filter [J]. World non-ferrous metals,2019(15):294-296.
5. Qin qinglei. Design of an infrared obstacle detection module with ultra-low power consumption [J]. Electronic fabrication, 2017, (7) : 15-16.
6. Shen lei-xian, pang jia-yi, zhang qingyun, xu he. Design and implementation of beidou-based positioning service platform [J]. Computer technology and development, 2019, (01):1-4.
7. Tian jie. Research on bluetooth-based traffic perception system detection device [J]. Heilongjiang traffic science and technology, 2008,41(05):216-217.
8. Bao nan. Urban culture cannot be separated from the cultivation of "bookiness" [N]. Beijing daily,2019-11-27(003).

A Design and Implementation of Quad-rotor UAV

Junjie Lin [1], Chunxia Zhang [1*], Yizhun Peng [1,2], Ting Zhao [1], Zhengke Xu [1]

[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: * zhangchx@tust.edu.cn*
www.tust.edu.cn

Abstract

In the past 5 years, considerable attention has been paid to unmanned intelligent devices. The Quad-rotor UAV(Unmanned Aerial Vehicle) is an unmanned aircraft controlled by radio remote control equipment and self-contained program control device. Our UAV is based on TI MSP432, which can be connected to PC via USB port. This printed circuit board is used to transmit the program for the flight control of UAV. The OpenMV module serves as the data source of the line patrol controller, which is the top-level controller, the same level as the remote controller. The IMU unit calculates information to attitude controller, in order to keep the flight of UAV stable. After the IMU solution information is fused with the optical flow sensor, the information is sent to the horizontal controller to control the flight of the UAV in the horizontal direction. Similarly, After the IMU information is fused with the laser height information, the height controller is applied to control the flight height of UAV.

Keywords: UAV, OpenMV module, IMU unit

1. Introduction

With the rapid development of the times and the rise of the intelligence of unmanned devices, the unmanned intelligent devices have received extensive attention in the past 5 years. Four rotor UAV is a kind of UAV controlled by radio remote control equipment and independent program control device[1]. At present, the application of UAV in aerial photography, agricultural plant protection, mapping and other fields has greatly expanded the use of UAV itself [2][3].

Compared with general aircrafts, UAVs are widely used in both military and civil fields. The outstanding advantage of the UAV is that people can control the aircraft remotely, or even let the aircraft perform tasks according to the designed procedures to liberate human resources.

Based on the above discussion, this design focuses on the basic function realization of UAV and the realization of camera and line patrol function based on OpenMV. This design system uses MSP432 microcontroller of TI company for flight control, OpenMV3 camera for machine vision processing[4], optical flow sensor and other technologies, through program design and hardware production, the UAV can complete various tasks. In various fields of military and residential life, the design has a very high practical value.

2. Main control device

2.1. Selection of controller

With the development of electronic technology, we have more possibilities in the selection of main controller. Our team had planned to use STM32 single-chip computer and TI series single-chip computer respectively to carry out the flight control experiment of the line patrol robot.

Compared with 51 single-chip microcomputer, STM32 single-chip microcomputer has the advantages of faster processing speed and more peripheral serial ports. Besides, our team members are familiar with the application of STM32 single-chip microcomputer. Therefore, we choose this single-chip microcomputer in the beginning. We read the data collected by MPU6050 through the programming code, determined the state of the aircraft, and then carried out the corresponding data processing.

We find that in the flight control of Quad-rotor UAV, compared with TI series single-chip microcomputer, STM32 single-chip microcomputer has some disadvantages in program processing speed and power consumption. When the main controller of Quad-rotor UAV has a large amount of computation, STM32 cannot meet our requirements.

TI uses the Cortex-M kernel, which contains the access to the complete arm instruction set. In addition. It also includes the DSP extension instruction and a floating-point FPU module, which improves the performance compared with STM32. Therefore, after the discussion with team members, we chose TI series MSP432 single-chip microcomputer.

2.2. Control system scheme

We adopt the scheme of minimum system of single-chip microcomputer. The minimum system of single chip microcomputer has high-performance simulation technology and abundant on-chip peripherals, which can significantly reduce the design of peripheral circuits and reduce the difficulty of system design. It is very suitable for our system design, and this scheme has an efficient and flexible development environment.

3. Machine vision tools

Optical flow is the apparent movement of the image brightness mode. Under certain conditions, the motion information of an object can be obtained based on the optical flow[5]. We install the optical flow sensor at the bottom of the drone, use the optical flow method to analyze the ground feature information collected by the camera, calculate the speed of the aircraft relative to the ground, and combine the speed obtained by the optical flow method with the speed obtained by the inertial element to obtain more accurate Data, and then obtain the relative position information of the aircraft through integration to achieve positioning. Optical flow methods are divided into two categories, dense and sparse. Among them, dense optical flow calculation is complex and requires a large amount of resources. It requires the processor to have high computing power, so it cannot be applied to embedded platforms[6].

The maximum output of image information collected by Ov7725 camera is 300000-pixel image, which is smaller than the maximum output pixel of OpenMV. With active crystal oscillator and voltage regulator chip, and FIFO frame buffer chip - the chip contains 384K FLASH, which can cache two frames of QVGA image data, but the stability is relatively poor.

OpenMV is an embedded camera with STM32 as the processing core. It is equipped with micro Python interpreter and supports Python Programming on the embedded. This system uses OpenMV instead of computer to carry out a series of image acquisition and processing tasks, and can directly control the pan tilt system, simplifying the structure of the system, making the whole system less cost, smaller volume, and easier to apply to real life scenarios. When the whole moving target tracking system works offline, the system only takes up 10 cm *10 cm * 10 cm space, which is incomparable with another scheme. In addition, the anti-interference ability of OpenMV is stronger than that of Ov7725 camera.

To sum up, our team decided to choose the OpenMV3 camera and use Python language for machine vision processing. OpenMV is shown in the Fig.1.

Fig.1. OpenMV

4. Microcontroller

MSP430 is a 16-bit reduced instruction set microcontroller. Developers can write commonly used systems flexibly, including timer, input / output expander, system reset controller, electrically erasable programmable read-only memory (EEPROM), etc.

32 in MSP432 represents that the MCU is 32-bit. Compared with 16-bit RISC (reduced instruction set) MSP430, MSP432 adopts 32-bit RISC, which greatly improves the performance. While optimizing the performance, the power loss is reduced, and its effective power consumption and standby power consumption are only 95μA/MHz and 850nA/MHz respectively. After comprehensive consideration, we choose MSP432 controller. MSP432 board is designed as shown in the Fig.2.

Fig.2. MSP432 board

5. Mpu6050 sensor

Mpu-6050 is the first integrated 6-axis motion processing component in the world. Compared with the multi-component scheme, Mpu-6050 avoids the problem of the difference between the time axis of the combined gyroscope and the accelerator, and reduces a lot of packaging space. Mpu-6050 has a full sensing range of ± 250, ± 500, ±1000 and ±2000°/sec(DPS), which can accurately track fast and slow movements.

5.1. Calculation

The mpu6050 sensor has an internal digital LPF, which can be modified by setting the value of the config register.

This is the initialization program of mpu6050 sensor written by our team during the calculation.

```
I2C_Write_Byte(MPU6050_ADDRESS,
PWR_MGMT_1,      0x80);
I2C_Write_Byte(MPU6050_ADDRESS,
SMPLRT_DIV,     0x00);     0x00(1000Hz)
I2C_Write_Byte(MPU6050_ADDRESS,
PWR_MGMT_1,     0x03);
I2C_Write_Byte(MPU6050_ADDRESS,
CONFIGL,           0x04);       0x04(20Hz)
I2C_Write_Byte(MPU6050_ADDRESS,
GYRO_CONFIG,    0x18);
I2C_Write_Byte(MPU6050_ADDRESS,
ACCEL_CONFIG,   0x18);
```

5.2. Automatic correction (triaxial acceleration sensor)

$${}^{b}a_m = T_a K_a ({}^{b}a'_m + b'_a)$$

6. UAV attitude control

We use PID to control the position of the aircraft to obtain the desired speed, and obtain the actual speed based on the sensor fusion. The PID is used to control the speed of the aircraft, and then the closed-loop PID control algorithm is used to control the attitude of the aircraft [7]. In attitude control, we need to know three values: quaternion, rotation matrix and Euler angle.

If you need to transform vectors between coordinate systems, you can choose matrix form; another method is to use Euler angle as the “master copy” of azimuth, but maintain a rotation matrix at the same time. When the Euler angle changes, the matrix also needs to be updated at the same time. When large amounts of data are saved (such as animation), Euler angles or four yuan numbers are used, while Euler angles occupy less than 25% of the memory, but the conversion to matrix is slow. If the animation data needs to nest the connection between coordinate systems, quaternions may be the best choice. Smooth interpolation can only use quaternions. In other forms, we must turn to quaternion, then turn back after interpolation, and then get the attitude solution [8] [9].

7. Hardware circuit design

7.1. Overall system diagram

The overall diagram of the system is shown in the Fig.3.

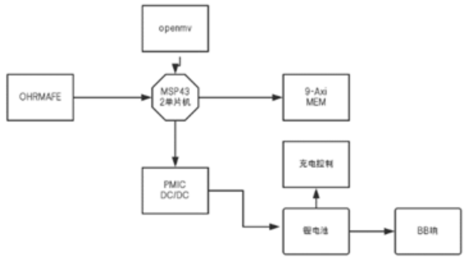
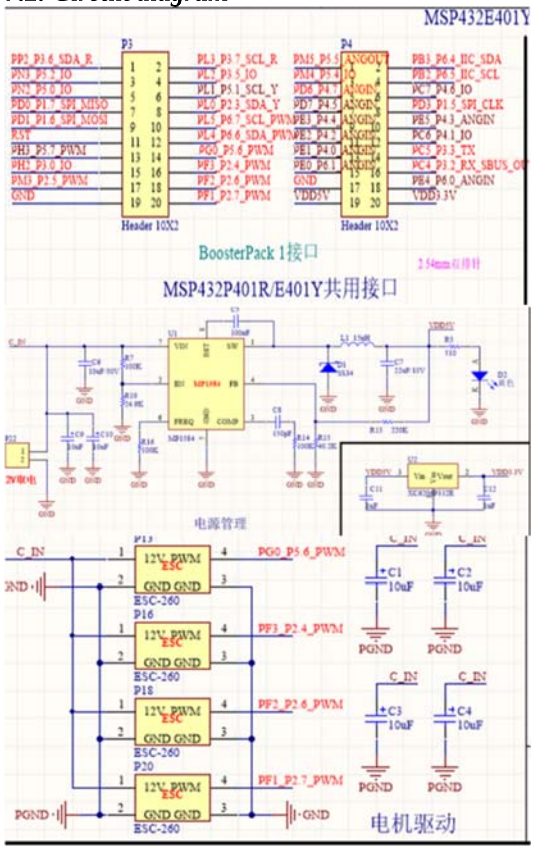


Fig.3. Overall system diagram

7.2. Circuit diagram



7.3. Power supply

The power supply consists of transformer, filter and regulator. Provide 5V or 12V voltage for the whole system to ensure the normal and stable operation of the circuit. This part of the circuit is realized by 3 terminal voltage-regulator tube.

8. Program design

The program design is divided into two parts, which are written by different members. The first part is the control of the aircraft, including take-off, hover height determination, rotation, horizontal flight and landing, etc. The second part is the OpenMV self-process, including color and line recognition, shooting, storage and so on.

The main idea of our programming is to first use the C language to write the one button take-off procedure, and then use the keil5 software and optical flow sensor to realize the four-rotor aircraft's height and stability.

After that, the joint debugging of OpenMV and flight control is carried out to realize the cable seeking flight of the aircraft. At the same time, the self-program of OpenMV realizes the function of taking photos and storing in the flight process.

9. Test plan and result analysis

According to the requirements of UAV safety performance, we set up the experimental site in our laboratory and carried out many tests. The main testing instruments and instruments we use are: MSP432 microcontroller, optical flow sensor, OpenMV3cam7 camera, HIGHDISCHARGELI-POBATTERY lithium battery and sw-lds50a laser module.

Through the above module and the corresponding program, we have been able to achieve the line finding and basic flight and can achieve the function of fixed height and camera shooting storage.

Acknowledgements

This research was supported by Student's Platform for Innovation and Entrepreneur Training Program, Tianjin University of Science and Technology(201910057245).

References

1. Tang Qiang, Zhu Zhiqiang, Wang Jianyuan.Research on Autonomous Flight Control of Foreign UAVs [J]. System Engineering and Electronics, 2004, 26 (3)
2. Tang Shufang. Research on four rotor aircraft technology based on optical flow positioning [D]. Nanjing. Nanjing University of Aeronautics and Astronautics
3. Chao Hai Yang, Cao Yong Can, Chen Yang Quan.Autopilots for small unmanned aerial vehicles:A survey[J].International journal of control, automation, and systems, 2010, 8 (1)
4. Chen Kaiyue. Design and implementation of machine vision test system based on image processing [D]. Hebei University of science and technology. 2017
5. Jiu Ruibo. Airborne video stabilization algorithm based on FAST-9 corner and optical flow method [J]. Science and Technology Information. 2019 (01)
6. Fankhauscr P, Huttcr M. ANYmaI: A Unique Quadruped Robot Conquering Harsh Environments[J]. Rcsearch Features, 2018(126):54-57.
7. Gu Zhixin, Yao Bo, Xu Kaihong, et al. Research on rock sheep limbs dynamics analysis and related technologies. J Xi. Development and innovation of mechanical and electrical products, 2010, 23 (J).
8. Song Yu, song Junwei. Design simulation of attitude stability controller for four rotor UAV [J]. Computer simulation. 2017 (08)
9. Huang Jiaqi, Wu Xiaoyan, Zhang Rui, Ma Zhen. Research on optimal control of attitude stability of four rotor aircraft [J]. Computer simulation. 2016 (07)

Characteristic Analysis and Synchronization Control of a Non-equilibrium System

Lianchen Zhao, Xinyu Zhang, Yizhun Peng*, Yuheng Zhang, Hongyan Jia
Tianjin University of Science and Technology, China
* *pengyizhun@tust.edu.cn*

Abstract

In this paper, the dynamic characteristics of a chaotic system without equilibrium point are studied. Through numerical simulation and theoretical analysis, the chaos characteristics of the system without equilibrium point are studied. In this paper, the nonlinear feedback synchronization control method is used to synchronize the system, and through software simulation, the results are analyzed to determine whether the drive system and the response system are synchronized. At the end of this paper, we give the advantages and disadvantages of this synchronization method.
Keywords: Synchronous control Nonlinear feedback synchronization State Observer synchronization

1. Introduction

Chaos is a more complex and special motion in nonlinear dynamics system than other forms, which can be found everywhere in nature. With the development of the research on chaos system, the emergence of the system without equilibrium point and the analysis of its characteristics are becoming more and more popular[1].
Chaos synchronization usually means that there are at least two vibration systems and their phases can be coordinated. Although chaos synchronization and chaos control are independent in definition, it is generally considered as a special form of chaos control in academia. Before the chaos synchronization is proposed, the traditional chaos control is to stabilize the system on the track of unstable period, and the chaos synchronization is to reconstruct the two systems. Many researchers put forward many concepts of synchronization, such as complete synchronization, same step, generalized synchronization, lag synchronization and so on[2]. In addition to the rich research results in theory, there are also many synchronization schemes in experiments, such as Guan Xin equal based on neural network, studied the synchronization problem of two chaotic systems in the presence of disturbance, Lu Xiang and so on designed a delay synchronization controller to realize the synchronization control between spatiotemporal chaotic systems[3]. At present, scholars are keen to combine electronic circuit, laser system, neural network and computer system with chaos synchronization, so as to realize some application schemes of chaos control and synchronization. Therefore, chaos synchronization is one of the hottest research directions in the current chaos, and

also one of the most widely used topics. Its future prospects are very broad.

2. Theoretical basis of chaotic system

Chaos can be seen everywhere in our life, such as the reproduction of biology, the fluctuation of stock market, the sudden change of geomagnetic field and so on. Chaos is a kind of seemingly random but actually deterministic motion. Compared with the mechanical motion in classical mechanics, it can predict the motion state of the system at a certain time in the future through calculation and analysis according to the system state at the initial time and the deterministic law. In chaotic system, the state of the system will not be accurately predicted, and only the future motion trend can be determined.
Li Tianyan, an American Chinese scholar, and his thesis tutor York first proposed the concept of chaos, and gave the mathematical definition of chaos.

Li Yorke definition[4]:
The mapping f contains all periodic points. There is an uncountable subset $S \in \mathrm{I}$ without periodic points, and the following conditions are satisfied:

① $\forall x, y \in S, x \neq y \lim_{n\to\infty} sup\left|f^{(n)}(x) - f^{(n)}(y)\right| > 0$

② $\forall x, y \in S, \lim_{n\to 0} inf\left|f^{(n)}(x) - f^{(n)}(y)\right| = 0$

③ $\forall x \in S, \lim_{n\to 0} sup\left|f^{(n)}(x) - f^{(n)}(y)\right| > 0$

f which satisfies the above conditions, is said to be chaotic on set s. From the mathematical definition put forward by two scholars, we can get the essential characteristics of a chaotic system, which are boundedness, aperiodicity and initial value sensitivity.

In 1989, there was a new definition of Devaney[5]: Let X be a metric space. A continuous mapping: $f: X \to X$ is called chaos on X,and f satisfies the following conditions:

(1)f is topologically transitive;

(2)The periodic points of F are dense in X

F is highly sensitive to the initial conditions.

2.1. Research methods of chaos

In order to understand the chaotic nature of the system, we usually simulate the bifurcation diagram, time sequence diagram, phase trajectory diagram and Lyapunov exponent spectrum of the system to analyze the system.

Select the time parameter as the ordinate and take a separate state quantity as the abscissa. The graph drawn in this way is called the sequence diagram. In this way, we can draw all the state variables into one graph, so we can get the phase trace graph of the system. From the time sequence diagram, we can clearly know whether the current state of the system is in chaos or periodic state through numerical changes. From the phase trajectory, we can clearly and accurately judge the state of the system from the shape of the attractor.

In a general nonlinear system, we assume that there is a parameter variable, when the parameter variable changes very little near a specific value, then the small change will cause a huge change in the topological properties of the phase space. This phenomenon is called bifurcation, and the critical point is defined as the bifurcation value, when the parameter is on the coordinate axis In the abscissa of, we call the point of bifurcation and define the point that does not cause bifurcation as the constant point. As a common phenomenon, bifurcation is very important in flight linear system. In order to get the bifurcation diagram, we need to get all the values of the parameters in a certain range, so that the topological properties of the trajectory lines in the phase space will be shown, and the properties shown in the diagram are the bifurcation diagram we want. When the parameters are taken to different values, the system will also show different states. After analysis, we can see that the occurrence of bifurcation points means that the system will be unstable, so the bifurcation points are sometimes called unstable points.

Lyapunov index is a very important reference. It is an index law to measure the attraction or divergence of adjacent orbits in phase space, and judge whether the adjacent orbits in phase space attract or diverge according to its positive and negative energy. When one of the Lyapunov exponents of the system is less than zero, it means that the motion of the system in this direction will be a stable state[6]; otherwise, when one of the Lyapunov exponents of the system is greater than zero, it means that the motion of the system in this direction is unstable, or the phase space of the system will be an expanded state, that is, the adjacent trajectory will change more and more Divergence, because of this nature, we can see that the long-term behavior of the system in this state is unpredictable. Therefore, it can be realized to judge the state of the system according to its Lyapunov index.

Here we take a three-dimensional chaotic system as an example to illustrate the relationship between the Lyapunov exponent and the state of the system[7]. Because the system is a three-dimensional system, it is not difficult to see that the system has three Lyapunov exponents. The three of them together are called the Lyapunov exponent spectrum. After calculation, when all of them are negative, the attractor is called fixed point and the system is in steady state. When one of them is zero and the other two are negative, the attractor is periodic and the system is in periodic state. The difference between stationary state and periodic state is not obvious, so they can also be regarded as the same state. If one of them is zero and the other two are not zero, and the attractor is a chaotic attractor, then the system is in a chaotic state[8].

2.2. Mathematical model of three-dimensional chaotic system without equilibrium point

The state equations of the three-dimensional chaotic system studied in this paper are as follows：

$$\begin{cases} \dot{x} = ay + xz \\ \dot{y} = -bx + yz \\ \dot{z} = 1 - x^2 - y^2 \end{cases}$$

Among them, $x, y, z \in R$ is the state variable of the system, and the parameters a = 0.05, b = 1 are selected. The phase trajectory of the system and the projection of chaotic attractor on each phase plane are simulated by

MATLAB software. Phase trace diagram is shown in the Fig.1.

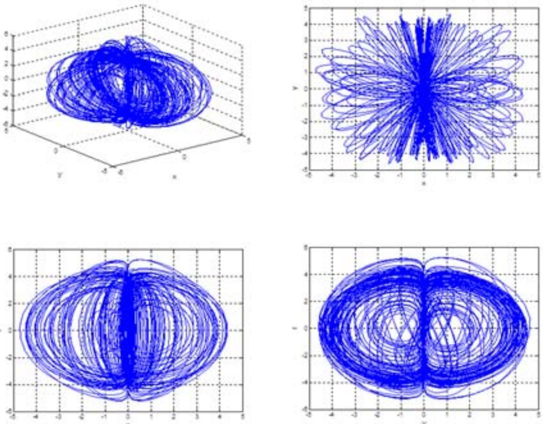


Fig.1. Phase trace diagram of the system in each plane

When the given initial value is (x (0), y (0), Z (0)) = (- 1, - 1,4), the three Lyapunov exponents of the system are λ 1 "= 0.0386" λ 2 "= - 0.1232" λ 3 "= - 0.0968" by software simulation, and the formula can be used at the same time:

$$D_L = j + \frac{1}{\left|\lambda_{L_{j+1}}\right|}\sum_{i+1}^{j}\lambda_{L_i}$$

The dimension of the attractor is calculated as D_L = 2.18, which shows that the system appears chaos.

2.3. Characteristic analysis of system without balance point

(1)Symmetry

We define a new set of coordinates

$$\begin{cases} z_1 = -x \\ z_2 = -y \\ z_3 = z \end{cases} \qquad \text{(a)}$$

We found

$$\begin{cases} \dot{z}_1 = -ay - xz = az_2 + z_1 z_3 \\ \dot{z}_2 = bx - yz = -bz_1 + z_2 z_3 \\ \dot{z}_3 = 1 - x^2 - y^2 = 1 - z_1^{\,2} - z_2^{\,2} \end{cases} \qquad \text{(b)}$$

This means that the coordinate value (x, y, z) of the differential equation of the system is transformed into (- x, - y, z) without change, that is, the system has symmetry about the Z axis.

(2)dissipation

$$\nabla \dot{V} = \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z} = 2z$$

When $\nabla \dot{V} < 0$, the system is always dissipative and converges with the index DV / dt = e ^ (- t). In this system, if $\nabla \dot{V} < 0$ is to be determined first, then with t → ∞, each volume element including the system trajectory will shrink to 0. The asymptotic dynamic behavior of all system trajectories will be fixed in one attractor and limited to a limit subset of volume 0, which further proves the existence of attractors.

Based on the simulation of the system without balance point with different parameter values and the same initial value of [- 1-14], Fig.2. is obtained. These two figures are the sum of the three Lyapunov exponents of the system from - 5 to 5 after fixing one parameter. It can be seen that no matter which parameter is changed by the system, the sum cannot be stable at zero, so that It can be said that the system is not a conservative system.

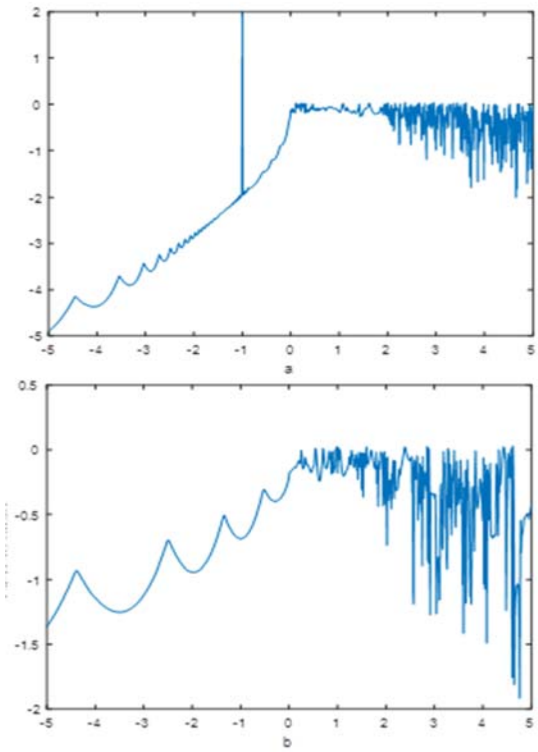


Fig.2. The sum of Lyapunov exponents with parameters a and b

(2) Equilibrium analysis

$$0 = ay + xy \qquad \text{（a）}$$
$$0 = -bx + yz \qquad \text{（b）}$$
$$0 = 1 - x^2 - y^2 \qquad \text{（c）}$$

for（a）and（b）:

$$xyz = -ay^2 = bx^2 \qquad \text{(d)}$$

$$\text{If} \quad bx^2 + ay^2 = 0 \qquad \text{(e)}$$

When a>0,b>0, $x = 0, y = 0$

Contradictory, that is, the system has no equilibrium point.

2.4. Initial value sensitivity

There are two parameters in the system, a = 0.05 is determined first, and the characteristic of parameter B is analyzed. The bifurcation diagram and Lyapunov exponent spectrum of the system are drawn with MATLAB software.

The initial values are [-1 -1 4] and [-2 -2 4], a = 0.05. Through software simulation, the following figures are obtained. From these two figures, it can be seen that the three-dimensional chaotic system without equilibrium point is in periodic state at b ∈ (- 2,0) and in chaotic state at b > 0, so it can be judged that the change of initial values has a very significant impact on the chaotic state of the system.The consequence is shown as Fig.3.

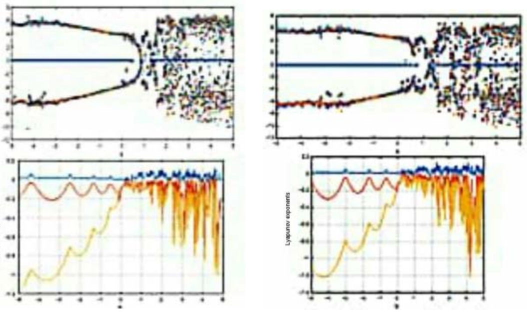


Fig.3. Bifurcation diagrams and Lyapunov exponents with initial values of [-1 -1 4] and [-2 -2 4]

3. Synchronous control analysis of a system without balance point

There is an error between the response system and the drive system, and the feedback control is to have a feedback between them. By controlling the feedback, the error is gradually eliminated, and finally the synchronous control is achieved[9].

First, the three-dimensional chaotic system without equilibrium point is taken as the driving system

$$x = x_1, y = x_2, z = x_3$$

The drive system is:

$$\begin{cases} \dot{x}_1 = ax_2 + x_1x_3 \\ \dot{x}_2 = -bx_1 + x_2x_3 \\ \dot{x}_3 = 1 - x_1^2 - x_2^2 \end{cases}$$

The response system is:

$$\begin{cases} \dot{y}_1 = ay_2 + y_1y_3 + u_1 \\ \dot{y}_2 = -by_1 + y_2y_3 + u_2 \\ \dot{y}_3 = 1 - y_1^2 - y_2^2 + u_3 \end{cases}$$

Where is the system synchronization variable and the parameter is $e_1 = y_1 - x_1, e_2 = y_2 - x_2, e_3 = y_3 - x_3$.

Let the error between the response system and the drive system be the error vector $\boldsymbol{e} = [e_1, e_2, e_3]^T$. By subtracting the two formulas, the error equation is:

$$\begin{cases} \dot{e}_1 = 0.05e_1 + y_1y_3 - x_1x_3 + u_1 \\ \dot{e}_2 = -e_1 + y_2y_3 - x_2x_3 + u_2 \\ \dot{e}_3 = -y_1^2 + x_1^2 - y_2^2 + x_2^2 + u_3 \end{cases}$$

Thus, the synchronization problem of chaotic system can be transformed into the stability problem of error system at the origin (0,0,0). If an appropriate control law $\boldsymbol{u} = [u_1, u_2, u_3]^T$ is selected to stabilize the error system, the response system and the drive system are synchronized. By using the nonlinear feedback control method, for the drive system and the response system, the controller of the system is:

$$\begin{cases} u_1 = x_1x_3 - y_1y_3 + k_1e_1 \\ u_2 = e_1 + x_2x_3 - y_2y_3 + k_2e_2 \\ u_3 = y_1{}^2 + y_2{}^2 - x_1{}^2 - x_2{}^2 + k_3e_3 \end{cases}$$

Put the controller into the error system and sort it out:

$$\begin{cases} \dot{e}_1 = 0.05e_1 + k_1e_1 \\ \dot{e}_2 = k_2e_2 \\ \dot{e}_3 = k_3e_3 \end{cases}$$

In order to obtain the value range of controller parameter K, the Lyapunov function is constructed as follows:

$$V = \frac{1}{2}(e_1{}^2 + e_2{}^2 + e_3{}^2)$$

The formula is derived along the error e , and the comprehensive formula is obtained

$$\dot{V} = \dot{e}_1 e_1 + \dot{e}_2 e_2 + \dot{e}_3 e_3 = (0.05 + k_1)e_1^2 + k_2^2 + k_3^2$$

Obviously, when the controller parameters, there must be response system and drive system synchronization. In this paper, the controller parameters are k1 = - 3, k2 = - 3, k3 = - 3.

Next, the MATLAB software is used to verify whether the nonlinear feedback synchronization method used in this paper can control the system synchronously. First use the Simulink module of MATLAB software to build the drive system and response system, and then simulate. Through simulation, the synchronization curve is obtained as shown in the figure.4. From the figure, it can be seen that the final output of the drive system and the response system is the same after a period of synchronization control under the condition of different initial values, that is to say, the nonlinear feedback synchronization control method applied plays the role of synchronization control.

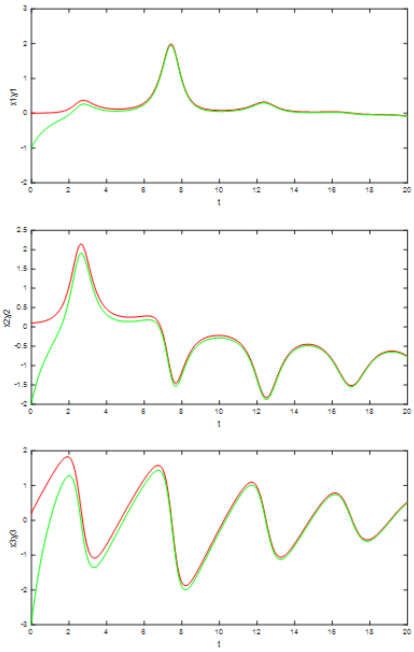


Fig.4. Synchronous curve

4. Conclusion

The full text is summarized as follows:

1. The characteristics of a three-dimensional system without equilibrium point are analyzed. Its dissipativity, symmetry, etc. are analyzed. The relevant characteristics are shown by Lyapunov exponent spectrum, bifurcation diagram, etc.
2. When the parameters are determined, a Lyapunov function is designed by using the nonlinear feedback synchronization method. Finally, the simulation results are obtained through simulation, and the conclusion that the system can be synchronized by this method is drawn after analysis.

In the work of this paper, the basic dynamic characteristics and synchronous control of a system without balance point are studied preliminarily, but the research on the hidden characteristics and coexistence characteristics of the system without balance point is not deep enough, whether there are other hidden characteristics needs further study, and in the synchronous control, there are not enough methods, whether there are better methods Synchronization, which needs to be done in the next phase.

References

[1] Zhang Wenxiao, Jiang Guoping. Control method of unstable equilibrium point of chaotic system with unpredictable state [J]. Computer technology and development, 2015,25 (01): 79-81 + 86

[2] Luo Dongmei. Analysis of chaotic sequence characteristics and Research on chaotic synchronization technology [D]. Chongqing University, 2011

[3] Wang Yu. Chaos synchronization control of nonlinear systems [D]. Northeast Petroleum University, 2018

[4] Shi Qingmei. Analysis and synchronization of a new chaotic system [D]. North China Electric Power University, 2012

[5] Anirban Ray,D. C. Saha,Papri Saha,A. Roy Chowdhury. Generation of amplitude death and rhythmogenesis in coupled hidden attractor system with experimental demonstration[J]. Nonlinear Dynamics,2017,87(2).

[6] Hamid Reza Abdolmohammadi,Abdul Jalil M Khalaf,Shirin Panahi,Karthikeyan Rajagopal,Viet-Thanh Pham,Sajad Jafari. A new 4D chaotic system with hidden attractor and its engineering applications: Analog circuit design and field programmable gate array implementation[J]. Pramana,2018,90(6).

[7] Huitao Zhao,Yiping Lin,Yunxian Dai. Hopf bifurcation and hidden attractor of a modified Chua's equation[J]. Nonlinear Dynamics,2017,90(3).

[8] J. Timmer,H. Rust,W. Horbelt,H.U. Voss. Parametric, nonparametric and parametric modelling of a chaotic circuit time series[J]. Physics Letters A,2000,274(3).

[9] Yu yeqiang. Research on synchronous control strategy of fractional chaotic system [D]. Northeast Petroleum University, 2017

Classification and Recognition of Baby Cry Signal Feature Extraction Based on Improved MFCC

Zhenjiang Chen[1,2], Yizhun Peng[1,2*], Di Li[1,2], Zhou Yang[1,2], Nana Wang[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)*

** pengyizhun@tust.edu.cn*

Abstract

Since MFCC was proposed, it has been widely used in feature extraction of speech signals. However, for some specific sound signals, such as baby crying signal, the direct MFCC feature extraction has a low classification and recognition rate. Through the study of MFCC feature extraction process, it is found that if each filter in the triangle filter bank is shifted upward by an ∂_i ($\partial_i \geq 0$).In addition, in the calculation of single frame MFCC, a continuous segment of sound information is reconstructed. The improved MFCC feature extraction can greatly improve the recognition rate and speed of baby crying recognition.

Keywords: MFCC, feature extraction, triangle band-pass filter bank, after framing, classification and recognition rate

1. Introduction

Voice information, especially infant crying information, contains not only semantic information, but also individual information. The research of feature parameters is the basis of infant crying recognition. Feature parameter selection should be able to fully and accurately express the feature information of baby voice.[1] The recognition of infant voice information needs to go through the process shown in Figure 1.

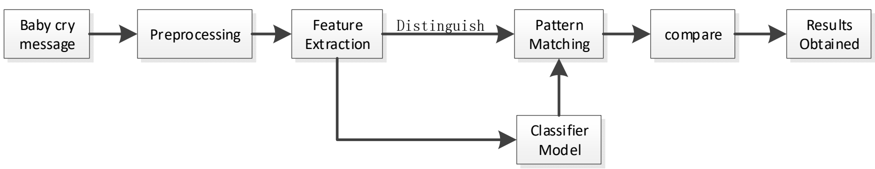


Fig. 1. General process of baby crying information recognition

Feature extraction of audio signal is the first step of any automatic speech recognition system. Therefore, the accuracy of the feature extraction of audio signal directly determines the recognition rate and the recognition effect of the final result. The feature extraction of audio signal is to extract the identified components from the audio signal, and then filter out other non important information, such as background noise, the mood of the speaker and so on.

In speech recognition and speaker recognition, the most commonly used speech feature is mel scale frequency cepstral coefficient (MFCC). Because the human auditory system is an ideal speech recognition system, and MFCC simulates the human auditory characteristics, which is a speech characteristic parameter in line with the human auditory characteristics, and achieves a high recognition rate in practical application.[2]

2. Standard MFCC Feature Extraction Process

2.1. *Mel Frequency*

Mel frequency is proposed based on the auditory characteristics of human ear. It has a nonlinear relationship with Hz frequency. Generally speaking, the bass is easy to mask the treble, while the treble is difficult to mask the bass. The critical bandwidth of sound masking at low frequency is smaller than that at high frequency.[3]

Mel cepstrum coefficient, MFCC is a cepstrum parameter extracted from mel scale frequency domain. Mel scale describes the nonlinear characteristics of human ear frequency. Because of the non-linear relationship between Mel frequency and Hz frequency, the calculation accuracy of MFCC decreases with the increase of frequency.[4] Therefore, in the application, only low frequency MFCC is often used, while medium and high frequency MFCC are discarded. Its relationship with frequency can be approximately expressed as follows:

$$\mathrm{Mel}(f) = 2595 \times \lg(1 + \frac{f}{700}) \qquad (1)$$

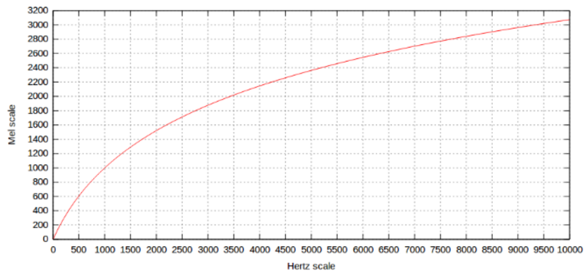


Fig. 2. Mel frequency versus linear frequency

2.2. *General MFCC Extraction Basic Process*

For the continuous audio signal, the first step is to pass through a high pass filter. The main purpose is to improve the high frequency part and flatten the signal spectrum. The next step is framing, in order to avoid excessive changes in the adjacent two frames. Then we add Hamming window to increase the continuity of the left and right ends of each frame. Then FFT is used to transform the time-domain signal characteristics into the energy spectrum in the frequency domain.[5] Moreover, Mel filter bank is also a triangular band-pass filter bank. Its main purpose is to smooth the spectrum, eliminate the harmonic, highlight the formant of the original voice, and reduce the amount of computation. Then the logarithm operation, and finally the discrete cosine transformation. Finally, the first-order and second-order difference parameters of MFCC are obtained.[6]

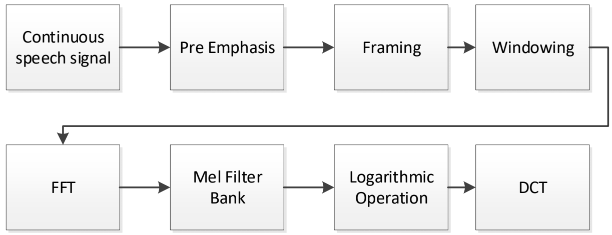


Fig. 3. General basic steps of MFCC extraction of speech feature parameters

The standard MFCC calculation steps are as follows:

(1) Pre Emphasis, Framing, Windowing and other preprocessing of the input sound signal.

(2) FFT of frame signal:

$$\mathrm{Y(k)} = \sum_{n=1}^{N} y\,(\mathrm{n}) e^{\frac{-j2\pi nk}{N}} \qquad (2)$$

In the above expression, *y(n)* is the frame signal and *N* is the frame signal length.

(3) The energy spectrum $|\mathrm{Y(k)}|^2$ of the sound signal is passed through the Mel filter bank to get the output of M filters.

(4) Take the logarithm of the output of each filter, and get the corresponding logarithm energy.

(5) The logarithm energy of Mel filter bank is transformed into discrete cosine, and the MFCC coefficient is obtained

$$\mathrm{F}_n = \sum_{m=0}^{M-1} S(m) cos\left(\frac{\pi n(m-0.5)}{M}\right) \qquad (3)$$

Where $S(m)$ is the logarithm energy output by Mel filter bank, n = *1, 2, …, L, L* denote the dimension of cepstrum coefficient.

2.3. *Analysis of triangular bandpass filter banks*

The energy spectrum of the frequency domain signal is passed through a set of Mel scale triangular filter banks, and a filter bank with m filters (the number of filters is similar to the number of critical bands) is defined.[7] The filter used is a triangular filter with a center frequency of f(m), *M = 1,2,..., M*. The interval between each f(m) decreases with the decrease of m value and widens with the increase of m value, as shown in Figure 4:

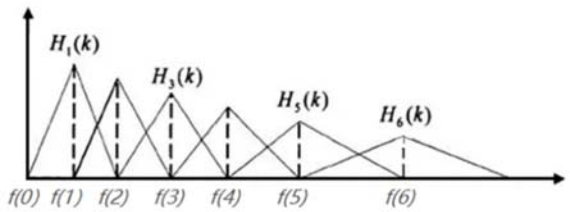


Fig. 4. Delta band-pass filter bank

The frequency response of a standard triangle filter is defined as follows:

$$\mathrm{H}_m(k)=\begin{cases}0 & ,k<f(m-1)\\ \frac{2(k-f(m-1))}{(f(m+1)-f(m-1))(f(m)-f(m-1))}, & f(m-1)\le k\le f(m)\\ \frac{2(f(m+1)-k)}{(f(m+1)-f(m-1))(f(m)-f(m-1))}, & f(m)\le k\le f(m+1)\\ 0 & ,k\ge f(m+1)\end{cases} \tag{4}$$

Because the audio signal to be processed is baby crying, which is an abnormal voice signal, it is easy to lose some important features in the low-frequency stage, so the classification and recognition of MFCC after feature extraction is poor, and the recognition rate is low.

3. Improved MFCC Baby Cry Feature Extraction

3.1. *Improved Triangular Bandpass Filter Bank*

The frequency response formula of the improved triangle filter is as follows:

$$\mathrm{H}_m(k)=\begin{cases}\partial_k & ,k<f(m-1)\\ \frac{2(k-f(m-1))}{(f(m+1)-f(m-1))(f(m)-f(m-1))}+\partial_k, & f(m-1)\le k\le f(m)\\ \frac{2(f(m+1)-k+\varepsilon_k)}{(f(m+1)-f(m-1))(f(m)-f(m-1))}+\partial_k, & f(m)\le k\le f(m+1)\\ \partial_k & ,k\ge f(m+1)\end{cases} \tag{5}$$

In the above expression, $\partial_k=\frac{1}{k}\sum_{m=1}^{k}\frac{f(m-1)}{f(m)}$,

$$\varepsilon_k=\frac{1}{k}\sum_{m=1}^{k}\frac{k-f(m)}{f(m+1)-k},f(m-1)\le 2k\le f(m+1) \tag{6}$$

The schematic diagram of the improved triangular bandpass filter bank is as follows:

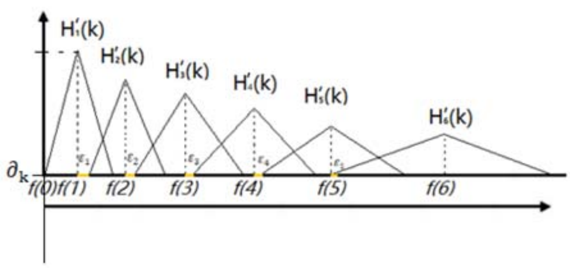


Fig. 5. Delta bandpass filter bank after entering

3.2. MFCC After Framing

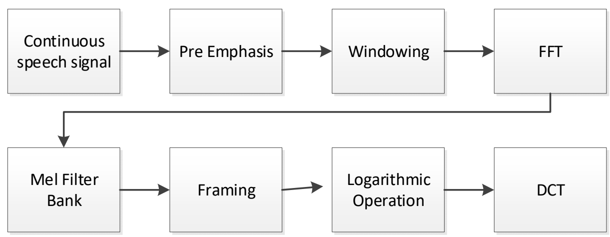
Frame is a collection of N sampling points into an observation unit. Generally, the value of n is 256 or 512 in speech processing, covering about *20-30ms*.[8] In order to avoid the change of two adjacent frames, there will be a section of overlapping area between two adjacent frames. This overlapping area contains M sampling points, usually the value of M is about 1/2 or 1/3 of n. Generally speaking, the sampling frequency of speech signal used in speech recognition is 8kHz or 16KHz. For 8kHz, if the frame length is 256 sampling points, the corresponding time length is $256/8000\times 1000 = 32ms$.[9]

If the original signal is divided into frames and windowed directly, many frames can be obtained. FFT (fast Fourier transform) is used for each frame. The function of Fourier transform is to transform the time-domain signal into the frequency-domain signal. After each frame is FFT, the frequency-domain signal can be stacked in time to get the spectrogram.

After FFT, we can get n frequency points, and the frequency interval (also called frequency resolution) is *Fs/N*. for example, if the sampling frequency is 16Hz and *N* is 1600, then after FFT, we will get 1600 points, and the frequency interval is 10Hz. The module of 1600 values obtained by FFT can represent the amplitude corresponding to 1600 frequency points. However, if we pursue the computing speed, it obviously consumes a lot of time for FFT after framing.[10]

Through analysis, if FFT calculation is carried out first, then the frequency spectrum is divided into frames, which greatly reduces the calculation time and improves the MFCC feature extraction speed.

Fig. 6. MFCC baby cry feature extraction flow chart after framing



4. Conclusions

Through the improved MFCC baby cry feature extraction and SVM classification and recognition, we found that compared with the standard MFCC feature extraction, SVM classification and recognition has a significant improvement in speed and accuracy.

In addition, the above improved strategy is aimed at the essential part of the MFCC feature extraction process, so it has a similar effect on some other sounds in the middle and low frequency band.

References

1. Chen Di, Gong Weiguo, Li Bo. Noise robustness speech high frequency weighted MFCC extraction for speaker recognition [J]. Journal of instrumentation, 2008 (03): 668-672
2. Gao Ming, sun Rencheng. Speaker feature parameter extraction algorithm based on improved MFCC [J]. Journal of Qingdao University (NATURAL SCIENCE EDITION), 2019,32 (01): 61-65 + 73
3. Yuan Gang, Li Tinghua, Jiang Youwen, Jiao Tao. An improved speech recognition algorithm based on MFCC and PCA [J]. Journal of Nanyang Institute of technology, 2015,7 (06): 56-60
4. Chuan Xie,Xiaoli Cao,Lingling He. Algorithm of Abnormal Audio Recognition Based on Improved MFCC[J]. Procedia Engineering,2012,29.
5. Yu Junjun. Application of speech recognition system based on improved MFCC parameters [D]. Jiangxi University of technology, 2012
6. Wang Liang. Research on abnormal sound recognition technology based on MFCC [D]. Harbin Engineering University, 2014
7. Xiong Wei, Luo Yungui. Research on MFCC algorithm of speech recognition [J]. Modern business industry, 2010,22 (03): 291-292
8. Ye Qingyun, Jiang Jia. Improved algorithm based on speech MFCC features [J]. Journal of Wuhan University of technology, 2007 (05): 150-152
9. Zhang rongqiang. Method of feature extraction in speaker recognition [D]. Dalian University of technology, 2005
10. Song Le. Research on improved feature extraction algorithm in speaker recognition [D]. Taiyuan University of technology, 2014

Design of Space Remote Sensing Data Storage Platform Based on Distributed File System

Di Li[1,2], Yizhun Peng[1,2*], Ruixiang Bai[1], Zhenjiang Chen[1,2], Lianchen Zhao[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

Due to the large space remote sensing data, a space remote sensing data has seven or eight hundred megabytes or more, and a large amount of space remote sensing data is generated every day for hundreds of GB, TB or even more, so a large amount of space is needed for storage. Space remote sensing data. In order to solve such problems, this paper prepares for the analysis of the subsequent space remote sensing data and combines the distributed file system to store the space remote sensing data. The installation of the CentOS 6.5 virtual machine through VMware to build an HDFS cluster, through a Namenode, three Datanode nodes to achieve access to space remote sensing data. Through the upload of the server, the space remote sensing data can be uploaded to the client, and the space remote sensing data can be downloaded through the client.

Keywords: Namenode, Datanode, HDFS, Space remote sensing data, Hadoop

1. Introduction

With the continuous development of the information age, data has shown an exponential growth, and the era of big data has emerged as a result. With the continuous progress of the times, the continuous development of science and technology, the rapid development of remote sensing technology is also quite rapid. Multi-sensors, multi-temporals, high spatial resolution, high spectral resolution of remote sensing data are increasing, and the data types are becoming more and more complex. . Among them, satellite remote sensing data is a layered digital image matrix, a data type with spatial characteristics, and an unstructured data, which cannot be represented by key-value pairs. Because traditional relational databases store structured data in the form of tables, they cannot store unstructured data such as remote sensing data[1].

The amount of space remote sensing data is huge, and the amount of data in a certain period of time in a certain area will reach 900 million or more. If comprehensive analysis and processing of space remote sensing data for multiple time periods are performed in a certain area, the amount of data will increase in the form of geometric multiples, and the data will reach hundreds of GB or even more. Capacity cannot afford the storage of massive data. Therefore, in recent years, with the continuous development of big data, distributed storage technology has gradually matured and improved, and distributed systems such as Hadoop and Spark have emerged as the times require. The distributed system uses multiple independent server nodes connected together to form a physically distributed, logically unified computer cluster distributed data storage system under the unified scheduling of the master node server, which can solve the multiple disadvantages of single-machine storage. Provides an effective and secure method for mass data storage.

2. Introduction of related technologies

2.1. Distributed Architecture Hadoop

Distributed architecture Hadoop can be built into clusters with common computer configuration. As the basic platform of cloud computing, Hadoop mainly consists of three modules: HDFS, MapReduce and Yarn. HDFS is a distributed file system, mainly serving as distributed storage in clusters. Function to achieve distributed storage of big data. MapReduce is a distributed computing programming framework whose main function is to implement distributed parallel computing in a cluster. Yarn distributed resource scheduling platform, the main function is to help users call a large number of MapReduce programs, and allocate computing resources reasonably. HDFS provides support for reading and writing files during MapReduce task processing[2]. MapReduce implements task distribution, tracking, execution, and collection of results based on HDFS. The two functions interact with each other to complete the core tasks of Hadoop cluster. Hadoop can freely organize computer resources, build a distributed cloud computing platform, and make full use of the computing and storage capabilities of the cluster to complete the storage of massive data.

The advantages of Hadoop clusters are as follows:

- Hadoop clusters can be scaled horizontally. When the data is too large and too large, the cluster can not bear the pressure. The cluster storage capacity can be directly expanded by adding nodes. Dynamic data movement can reduce the pressure on each node.
- Hadoop cluster adopts master-slave architecture. Nodes are divided into two categories: Namenode is the main node responsible for storing cluster metadata, which plays a role in supervising the execution of MapReduce. Datanode is the child node responsible for storing specific data. Perform specific tasks to keep the heartbeat with the primary node.

2.2. Distributed File System HDFS

The distributed file system HDFS has the same characteristics as the ordinary file system, and (1) has a directory structure. (2) All files stored in the system are files. (3) The system provides functions such as copying, moving, creating, deleting, modifying, and viewing files. The distributed file system and the stand-alone file system are different. The file system stored in a single machine is only in the operating system of one machine, and the distributed file system spans multiple machines. A single file is placed on a single machine's disk, while a distributed file system stores files on multiple machines. The working mechanism of the distributed file system is: when the client stores a file to the distributed file system, the distributed file system cuts and blocks the stored file, and stores the diced in the child nodes in the cluster. On the disk; once the file is cut and stored in the distributed file system, there is a mechanism for recording the dicing information of each file stored by the user, and dicing the specific storage path; in order to ensure the security of the data Sex, to ensure that data will not be lost, the distributed file system will store multiple backups of each file in the cluster to prevent data loss when a server hangs. In general, a distributed file system consists of a primary node server and N child node servers.

3. Space remote sensing data storage

3.1. Characteristics of space remote sensing data

Aerospace remote sensing data is taken by space satellites. Remote sensing satellite data is used by remote sensing satellites to detect the reflection of electromagnetic waves on the Earth 's surface objects in space and the electromagnetic waves emitted by them, so as to extract information about the object, complete the identification of objects at long distances, and convert these electromagnetic waves. The visible image is recognized as the satellite image[3]. Space remote sensing data is difficult to represent with key-value key-value pairs, which is an unstructured data. The following figure shows the composition of space remote sensing data.

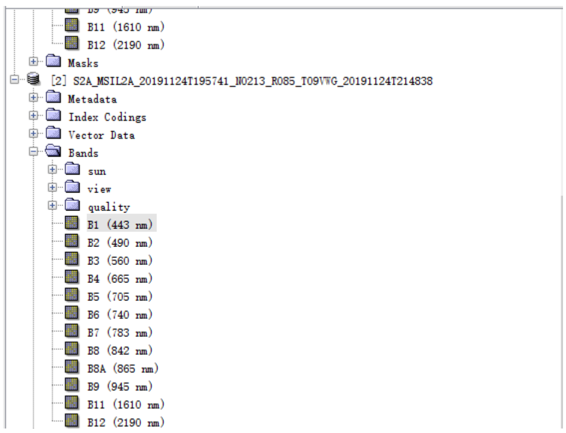


Fig.1. Composition of space remote sensing data

The above figure shows the space remote sensing data at a certain time in a certain area[4]. The specific space remote sensing data is shown below:

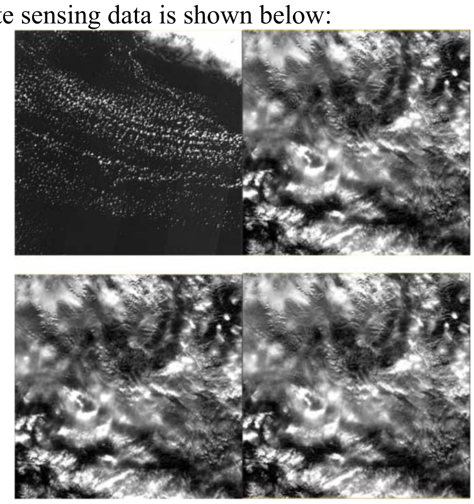


Fig.2. Specific display of space remote sensing data

3.2. Distributed file system setup

According to the characteristics of space remote sensing data, relevant space remote sensing data will be stored in the HDFS distributed file system. The system will build a distributed file system based on Hadoop cluster to store space remote sensing data. The distributed file system has four nodes, which are a master node Namenode node and three child nodes Datanode nodes, among which the master node It is responsible for recording the location of the stored file partition and the node where the chunk backup is located. The main task of the child node is to store the specific block[5].

3.3. Cluster architecture design

Using four ordinary computers, using a local area network to form a Hadoop cluster, you can use a common computer as a client for the client to log in to the client to access the cluster. The specific architecture is shown in the figure:

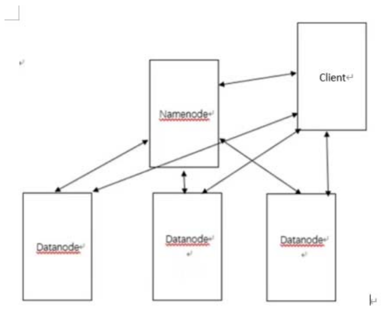


Fig.3. Block diagram of hadoop cluster architecture

The configuration of each computer is the same, the processor AMD Ryzen 3 2200G with Radeon Vega Graphics 3.50Ghz, memory 16g, hard disk 1T.

3.4. Building a clustered software environment

Install VMware Workstation Pro15 on the Windows host and four virtual machines on the VMware Workstation Pro15. The operating system is all Centos6.5, the JDK version is jdk1.8.0_212, and the Hadoop version is hadoop-2.8.5.

In a Hadoop cluster, one virtual machine is used as the primary node Namenode node, and the other three are used as child node Datanode nodes. The Master node IP is 191.168.220.30 and the NameNode and Secondarynamenode are installed. The Slave1 node IP is 191.168.220.31 and the DataNode is installed. The Slave2 node IP is 191.168.220.32. The DataNode is installed. The Slave3 node IP is 191.168.220.33[6].

The main steps in building a distributed file system cluster:

- Modify the machine's host name and specific IP address to configure the machine's host name to the Windows local domain name mapping file.
- Configure the basic software environment of the Linux server. For example, turn off the firewall and disable it. Install the JDK to configure its environment variables and the domain name mapping configuration of the hosts in the cluster.
- Modify the configuration file, specify the default file system as: hdfs, the primary node that specifies hdfs is the machine, the local directory that specifies the namenode storage metadata, and the local directory where the datandoe storage folder is specified.
- Start HDFS. First, you need to initialize the metadata directory of the namenode.

4. Experimental results

By looking at the server root directory to know that there is aerospace remote sensing data, it is uploaded to the distributed file system and stored in the spacedata folder by the instruction hdfs dfs -put /S2A_MSIL2A_20191124T195741_N0213_R085_T09VWG_20191124T214838.zip /spacedata.

Fig.4. Where the remote sensing data is located

Can check the storage location and backup status of the aerospace remote sensing data through the client. Enter the client IP address and log in to the /spacedata under the client's Browse Directory to see the stored space remote sensing data, as shown in the following figure：

Fig.5. Remote sensing data is stored in hdfs

Click on one of the data to observe the size of the data and the location of the server where the backup is located, and the data can be downloaded through this page.

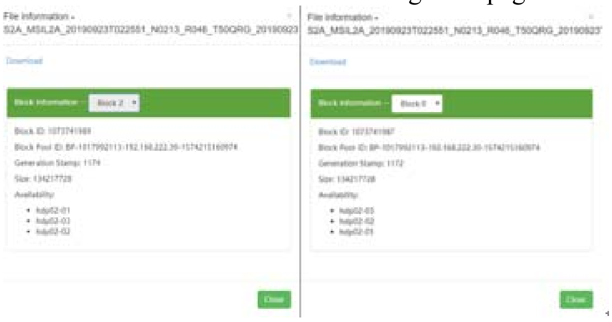


Fig.6. Block information of remote sensing data

Summary

Space remote sensing data is very important data and has important research significance for national geomorphology. Space remote sensing data is unstructured data so traditional relational databases cannot satisfy its storage. The storage of HDFS-based space remote sensing data is good for future data retrieval. The foundation is convenient for future research work.

References

1. Cheng Li, *Research on Distributed Storage of Remote Sensing Data Based on Hadoop*. (Shandong, Shandong Agricultural University, 2018)
2. Dazhi Wang, Research on Cross-Cluster Distributed File System Based on HDFS. *Information Technology and Informatization*, 2019 (8).
3. Zhongyi Chen, Distributed File System Based on Hadoop. *Electronic Technology and Software Engineering*, 2017 (9): 175-175.
4. Huan Yan, *Research and Implementation of Parallel Index Technology for Aerospace Information System*. (Xi'an, Xidian University, 2018)
5. Fanjun Meng, Wei Cao, Zhiqiang Guan, HIS-based Distributed Storage of AIS Data. *Information and Communications*, 2016 (5): 172-174.
6. Lixuan Chen, Shiyu Du, Chenlin Huang, et al. Design and implementation of teaching cloud platform based on distributed file system. *Wireless Internet Technology*, 2019 (9): 94-9

Research on Semantic Map Establishment of Parking Lot Based on Deep Learning and Multi-sensor

Shiqian Zhang[1,2], Yizhun Peng[1,2*], Ruixiang Bai[1], Yuheng Zhang[1,2], Tianye Jian[1,2]
[1]*College of Electronic Information and Automation, Tianjin University of Science and Technology, China*

[2]*Advanced Structural Integrity International Joint Research Centre, Tianjin University of Science and Technology, Tianjin, 300222, China)*

** pengyizhun@tust.edu.cn*

Abstract
With the continuous development of robot technology, SLAM, which is one of the key technologies to realize fully autonomous mobile robots, has become a hot topic in the scientific community. Traditional SLAM technology relies on a single type of sensor to obtain information, and the semantic information of the map is rarely obtained. This paper mainly improves the SLAM mapping technology of the parking lot from the information fusion of sensors and the deep semantics of obtaining map semantic information. Through the lidar and RGB-D camera to obtain multi-modal information, combined with deep learning, extract the semantic information such as the parking lot number and parking position in the image information, help the robot to map and locate faster and better. Finally, simulations were performed using ROS and gazebo to verify the feasibility of the system.
Keywords: SLAM, Lidar, RGB-D, Deep learning, Semantic map, ROS, Gazebo

1. Introduction

With the popularity of robots, drones, unmanned vehicles, and VR / AR in recent years, SLAM technology is also well known and considered to be one of the key technologies in these fields. Making robots more accurate and faster for positioning and navigation in unknown environments has become the main development direction of intelligent mobile robots. In the face of complex and uncertain external environments, robot systems need to integrate environmental sensing, path planning, and motion navigation functions, and continuously explore their environment through external and internal sensors configured by themselves, so as to achieve autonomous movement in unknown environments[1].

SLAM (Simultaneous Localization and Mapping), which was first proposed in the field of robotics, refers to: the robot starts from an unknown place in an unknown environment and locates its own position and attitude by repeatedly observing environmental features during the movement. , And then build an incremental map of the surrounding environment based on its own location, so as to achieve the purpose of simultaneous positioning and map construction[2]. Because of the important academic value and application value of SLAM, it has always been considered as the key technology to achieve fully autonomous mobile robots. However, the current SLAM technology still lacks in mapping. Traditional SLAM maps make no distinction between dynamic and static objects. This leads to a large error when the robot uses the map for positioning and navigation. This article aims to study and improve some disadvantages of traditional SLAM:

(1) Mapping accuracy of traditional SLAM maps is low.

(2) The traditional SLAM collects information singly and relies mainly on lidar.
(3) The information obtained by traditional SLAM mapping is far from enough.

2. System introduction

The system flow of this article is shown in Fig.1:

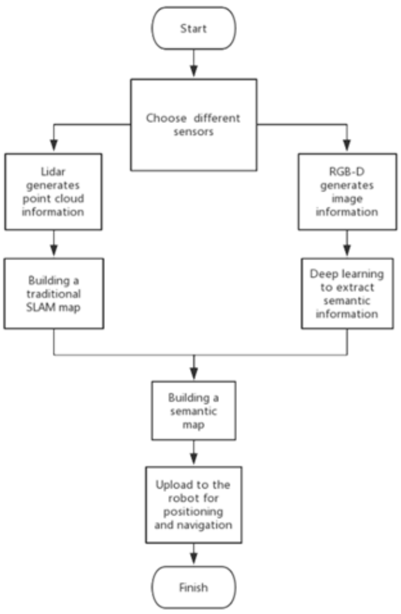


Fig.1 The system flow

2.1. The system of SLAM

The ROS open source community contains many SLAM function packages, that is, some commonly used SLAM algorithms, which can be used directly or secondary development, such as Gmapping, Karto, Hector, Cartographer, ORB-SLAM, etc.
Among them, ORB-SLAM2 is considered to be one of the most complete and stable SLAM systems, which can run monocular, binocular and RGB-D cameras[3].
This article uses the most commonly used and mature Gmapping function package. It integrates the Rao-Blackwellized particle filtering algorithm. Fig.2 shows the overall framework of the Gmapping function package.

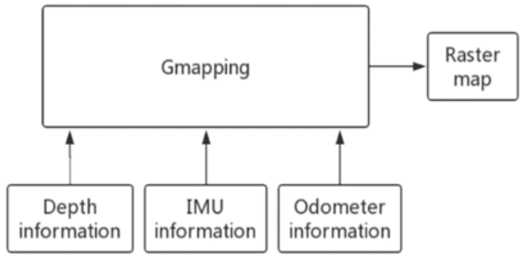


Fig.2. The overall framework of the Gmapping function package

2.2. The establishment of a semantic map

Traditional SLAM relies on feature extraction and matching algorithms based on low semantic levels such as points, lines, and surfaces to estimate motion. Not only lacks semantic information, but also the robustness of feature matching is low and it is easy to cause large estimation errors. Semantic maps help robots locate and navigate faster and more accurately in this environment by fusing the semantic information extracted by robots with traditional SLAM maps[4].

2.3. ROS framework design

At present, robot control research is costly and inefficient, and the robot operation platform ROS (Robot Operation System) is rich in open source programming resources, which is an effective tool for robot control research[5].
The text uses ROS as the simulation system. Realize the robot's SLAM and object recognition for unknown environment.
The calculation diagram of ROS in this paper is shown in Fig.3.

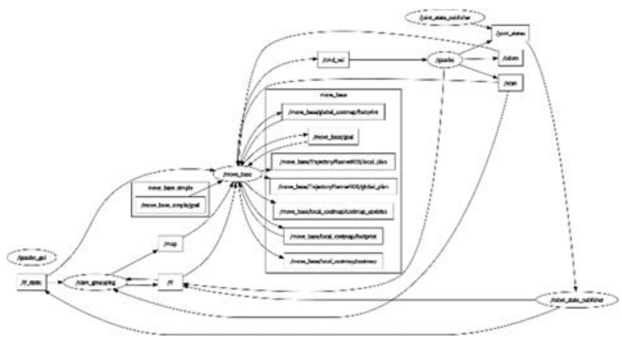


Fig.3. The calculation diagram of ROS

2.4. Construction of Gazebo-based simulation test environment

Gazebo simulation platform is suitable for the development of interactive robots, which can simulate the activities of recognition and positioning in space. Gazebo can not only simulate the motion of the robot, but also the data of the sensors, and the obtained data can be displayed directly in the rivz 3D visualization tool[6].
This article uses Gazebo to build a parking lot simulation environment and tests it to reduce the development cost and accelerate the development rate.

3. The results of experiment

This simulation is performed in the following environment. The ROS robot operating system selects the kinetic version, ubuntu16.04, and Gazebo7.0 simulation platform.
The simulation of parking lot in Gazebo is shown in Fig.4.

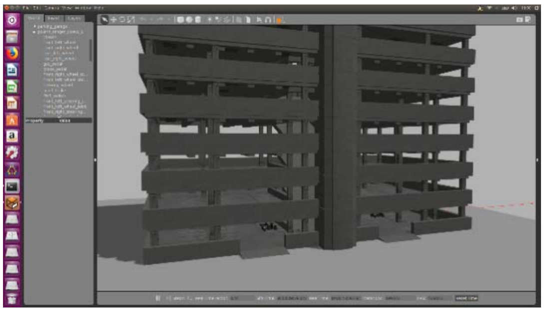


Fig.4. The simulation of parking lot in Gazebo

The schematic diagram in Fig.4. is a simulated multi-storey car park. This article only conducts experiments on the first floor. There are also five simulated cars added, which more accurately reflect the real scene. In order to avoid interference from other environments, the environment other than the parking lot is not simulated.
The simulation of SLAM robot in Gazebo is shown in Fig.5.

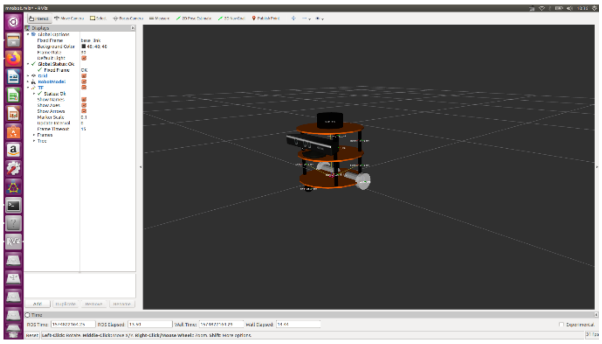


Fig.5. The simulation of SLAM robot

Fig.5. is a simulated SLAM robot, which is divided into three layers, the top layer is lidar, and the second layer is kinect. Point cloud information, depth information and image information are obtained through these two sensors, and SLAM and target recognition are performed.
The process of SLAM is shown in Fig.6.

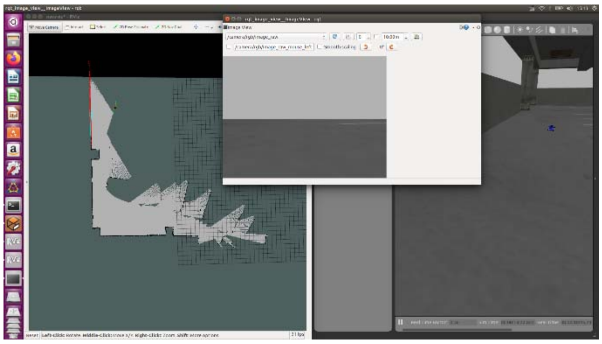


Fig.6. The process of SLAM

4. Conclusion

Autonomous localization and navigation is the key to autonomous mobile robots to solve various complex tasks, and it has important theoretical significance and application value for improving robot automation. SLAM technology can provide the structure information of the environment and its own location information, so as to achieve navigation and other functions. This paper studies and improves some disadvantages of traditional SLAM. First, by processing and comparing the multi-modal information collected by the robot, the accuracy of the mapping is improved. Second, the fusion of visual image information and lidar information through multi-sensors can better help robots to map and locate. Third, through deep learning methods, the robot extracts semantic information from the collected information and builds a semantic map to help the robot locate and navigate faster and more accurately in this environment.

References

1. Wang Dongshou, Wang Jia.Summarization of Research on Environmental Sensing Technology of Mobile Robots in Unknown Environment [J] .Machine Tool & Hydraulics, 2013,41 (15): 187-191.
2. Du Chen. Review of research on mobile robot-based synchronous positioning and map construction technology [A]. Network Application Branch of China Computer User Association. Network Application Branch of China Computer User Association 2018 The 22nd Annual Network New Technology and Application Conference Papers Episode [C]. Network Application Branch of China Computer Users Association: Beijing Union University Beijing Key Laboratory of Information Service Engineering, 2018: 5.
3. Xi Zhihong, Han Shuangquan, Wang Hongxu.Synchronous localization and semantic mapping of indoor dynamic scenes based on semantic segmentation [J] .Computer Applications, 2019,39 (10): 2847-2851.
4. Wu Hao, Chi Jinxin, Tian Guoli.Object Instance Recognition and Semantic Map Construction Based on Visual SLAM [J] .Journal of Huazhong University of Science and Technology (Natural Science Edition), 2019,47 (09): 48-54.
5. Li Tiejun, Wang Xiao, Yang Dong.Research on Robot Motion Control Based on ROS [J] .Modular Machine Tool & Automatic Manufacturing Technology, 2019 (10): 73-77 + 82.
6. Guo Mingyu, Qiu Hao.Research on ORB Feature Extraction and Comparison Based on Gazebo Simulation Environment [J] .Electronic Production, 2019 (15): 46-48.

Suvrey on Kinematics Calibration Technology of Manipulator

Zhou Yang[1,2], Yizhun Peng [1,2*], Nana Wang[1,2], Yuheng Zhang[1,2],Tianye Jian[1,2]

[1]College of Electronic Information and Automation, Tianjin University of Science and Technology, China;

[2]Advanced Structural Integrity International Joint Research Centre, Tianjin University of Science and Technology, China;

*E-mail: * pengyizhun@tust.edu.cn*

www.tust.edu.cn

Abstract

With more and more strict requirements on the quality and accuracy of industrial products, the requirements on the accuracy of robots are increased. Robot calibration is a common method to improve the accuracy. This paper mainly introduces the calibration based on model parameters, and summarizes the calibration based on kinematics model parameters from the aspects of kinematics modeling, attitude measurement and measurement instrument, parameter identification and error compensation. The modeling methods of various kinematics models, the applicable scope of measurement instruments and the algorithms of parameter identification and compensation are compared. Finally, the development trend of calibration is summarized and analyzed.

Keywords: Robot calibration, Kinematic model, Measurement, Parameter identification

1. Introduction

Mechanical arm has been widely used in the field of industry. For the use of mechanical arm, the most important factor is error. Error can be divided into two types: repeated positioning error and absolute positioning error. Repeated positioning error refers to the accuracy of the manipulator reaching the same point under the same operating instructions. Absolute positioning error is the error of manipulator between the desired target value and the actual value actually arrived. With the development and accumulation of technology, the repeated positioning error of the mechanical arm has been very small, usually up to 0.01mm or less[1]. But the absolute positioning error is still more than mm[2]. Therefore, it is necessary to calibrate the manipulator to improve the error of the manipulator.

Calibration is to measure the position and posture of the manipulator through the measuring instrument, then get the real geometric kinematics model of the manipulator through parameter identification, and finally improve the accuracy of the manipulator through error compensation.

The calibration process is as follows: First, the position and posture of the robot arm is measured by a measuring instrument. Then, the real geometric kinematics model of the manipulator is obtained by parameter identification. Finally, the error is compensated to improve the accuracy of the mechanical arm.

At present, there are three kinds of calibration: calibration based on model parameters, parameterless calibration and self-calibration. The common point of the first two types is that they need to establish a specific model for calibration. Calibration based on model parameters, its technology is mature and its calibration range is large. It is the current mainstream calibration method. Parameterless calibration does not require the establishment of an error model, but the quality of calibration mainly depends on the training data. Self-calibration focuses on automatic measurement and autonomic calculation, so it is also called closed-loop calibration. This paper mainly introduces the model-based calibration.

2. Calibration based on model parameters

The calibration based on model parameters needs to know the manipulator model and requires high data requirements. This kind of calibration can be divided into two types: the calibration based on kinematics model and the calibration based on dynamics model.

This paper mainly introduces the calibration based on kinematics model. It is a calibration method to establish kinematics model and solve the problem. The basic steps are the establishment of kinematics model, pose measurement, parameter identification and error compensation.

2.1. Modeling

The basis of the calibration of the manipulator is kinematic modeling, which represents the relationship of the manipulator from the base to the end. Generally, the more parameters in the model, the higher the accuracy of the established model will be, however, the complexity and computational complexity of the calibration also increase. At present, there are many methods to model the manipulator: DH model, MDH model, S model, CPC model, MCPC model, POE model, etc. The most classic model is DH model.

DH model is simple and easy to understand, but the defect is that when adjacent joints are parallel, the parameters are discontinuous and singular solutions will appear.

$$\begin{aligned} {}^{i-1}_{\ \ i}T = Rot(x_{i-1},\alpha_{i-1})Trans(x_{i-1},a_{i-1}) \\ Rot(z,\theta_i)Trans(z,d_i) \end{aligned} \tag{1}$$

Hayati[3] and Judd[4] et al. proposed a modified DH model(MDH model). In order to solve the problem of joint axis parallelization, the model added a rotation parameter about the y axis, so as to avoid the singular solutions caused by joint parallelization.

$$\begin{aligned} {}^{i-1}_{\ \ i}T = Rot(x_{i-1},\alpha_{i-1})Trans(x_{i-1},a_{i-1}) \\ Rot(z,\theta_i)Trans(z,d_i)Rot(y,\beta_i) \end{aligned} \tag{2}$$

Stone[5-6] et al. proposed another modified DH model (S model), in which one rotation parameter and one translation parameter were added. That is, each link was represented by six parameters.

$$\begin{aligned} {}^{i-1}_{\ \ i}T = Rot(x_{i-1},\alpha_{i-1})Trans(x_{i-1},a_{i-1}) \\ Rot(z_{i-1},\theta_i)Trans(z_{i-1},d_i)Rot(z_{i-1},\beta_i) \end{aligned} \tag{3}$$

To solve the integrity and continuity of model parameters, Zhuang [错误!未找到引用源。] and Schroer[8] et al. proposed CPC model, which can effectively describes differential transformation and eliminate singular solutions.

Chen[9] et al. raised product of exponentials model (POE model). It is also considered a zero reference to a method. The model establishes coordinates only at the base and end. The main advantage of this calibration model is that the local coordinate system can be placed at any position on the corresponding connecting rod, which greatly simplifies the model. It is not only suitable for traditional industrial robots, but also especially suitable for modular reconfigurable robots.

2.2. Measurement

Measurement is the most tedious but extremely important step in the calibration process. The measured data will directly affect the experimental results. In order to measure the position and posture of the robot, appropriate measuring instruments should be selected. In calibration, the following instruments are commonly used for measurement: coordinate measuring machine, theodolite, laser tracker, double ball-bar, draw-wire sensor, camera and so on. Different instruments vary in price, measurement accuracy, measurement method and range, and in the complexity of operating equipment and measurement. It is very important to select a proper measuring instrument for calibration experiment.

Coordinate measuring machine (CMM) is widely used in mechanical, electronic and instrument industries. It is one of the most effective methods to measure and obtain dimensional data. The current CMM are developing toward high precision and high speed. Dayong Yu[10] used a high-precision CMM to obtain the measurement residual, so as to estimate the actual value of geometric parameters. At the same time, he linearized the non-linear relationship between the position and posture of the parallel robot and geometric parameters, and used linear iterative least squares to estimate the actual value of geometric parameters.

Fig.1. Coordinate Measuring Machine

Theodolite is a kind of measuring instrument designed according to the principle of measuring Angle, which can measure horizontal Angle and vertical Angle.

Hanxu Sun[11] et al. proposed a new testing method of the position and posture of the end-effector of the manipulator based on the electronic theodolite. They used this method to measure the position and posture accuracy of the end actuator of a 6-degree-of-freedom manipulator, and the maximum difference is less than 1mm.

Fig.2. Theodolite

Laser tracker is a kind of high-precision and large-size measuring instrument in industrial measuring system. It is mainly used to track and measure the three-dimensional coordinate of the space moving object in real time. It has the advantages of high precision, high efficiency, real time tracking and measurement, quick installation, simple operation and so on. The tracker requires continuous data, but is susceptible to various effects and is cut off during the measurement process. This will affect work efficiency and the instrument is expensive. Hailong Bai[12] used the improved POE method to calibrate the robot arm using the laser tracker measurement method. The calibration results show that both the absolute positioning accuracy and the repeated positioning accuracy have been significantly improved. The accuracy of the calibrated robot can meet industrial requirements and process standards.

Fig.3. Theodolite

Double ball-bar (DBB)is a high-precision telescopic linear sensor with precision ball at each end. Two precisions magnetic bowl holders, one of which is connected to the measuring table and the other is connected to the extension rod at the end of the mechanical arm. The DBB measuring system is simple in operation, high in precision and low in cost. But due to its contact measurement characteristics, the calibration range using the DBB is small. Usually it has a measuring range of 1 mm and is not suitable for calibration with multi-degree-of-freedom manipulators.

Fig.4. Double Ball-bar

Draw-wire sensor is a sensor with small mounting size, compact structure, large measuring stroke and high accuracy. It measures from hundreds of millimeters to tens of meters. Ceccarelli[13] was the first to propose the application of draw-wire sensor in the measurement of space posture. He proposed a 3-2-1 cable posture measurement mechanism and applied it to the evaluation of robot workspace and other fields.

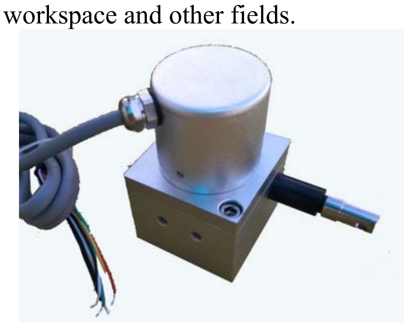


Fig.5. Double Ball-bar

The camera can be divided into monocular camera and multicamera camera. The camera itself needs internal calibration. There are two types of camera calibration: one is eye in hand, that is, the camera is fixed at the end of the mechanical arm, and the camera moves with the mechanical arm. The other is the eye to hand, that is, the camera is fixed in a suitable position outside the working space of the robot arm. The camera does not move with it. Haoyang Yu[14] used the robotic arm of the binocular vision servo as the research object and analyzed and studied the kinematics analysis theory of the manipulator and the binocular vision theory.

Fig.6. Camera

2.3. Parameter identification

Parameter identification is to calculate kinematics model parameters by measuring the position and posture of the robot and joint values. When the geometric parameter deviation is small, an error model can be established to describe the linear relationship between position and posture error and kinematic parameters.

The most commonly used method is the least square[15]. In addition, the extended kalman filter[16] (EKF)is also used in parameter identification. It is a classic method of solving nonlinear systems through Taylor series expansion in near estimates, EKF can better deal with the uncertainty of parameter identification problems. But EKF is easy to spread in the identification, and its identification accuracy is sometimes low. Levenberg-marquardt algorithm[17] combines Newton method and steepest descent method. It has strong local convergence performance, fast convergence speed and strong robustness. However, compared with other methods, it requires large memory under the same error convergence condition. Hu Zhang [18] respectively used the least square, parameter optimization, genetic algorithm and simulated annealing algorithm to calibrate and compare the robots.

2.4. Error compensation algorithm

Error compensation is to correct the parameter deviation in kinematics model. There are three widely used error compensation algorithms.

Joint space compensation: the calibrated results are directly corrected in joint space. For a reachable position and posture, the joint value is recalculated, so as to improve the accuracy of the position and posture of robot.

Differential error compensation:

Firstly, based on the idea of differential transformation, an error model is established. Secondly, by identifying the slight deviation between the nominal value and the true value of the geometric parameter, it is compensated the nominal parameter of the controller, thereby the accuracy of robot can be improved.

Real-time error compensation based on neural network: artificial neural network is established, parallel associative search solution space is carried out and adaptive reasoning is completed, the action rule of robot error source is obtained through training, and then kinematic inverse solution compensation is carried out.

3. The conclusion

The methods of robot kinematics calibration are more and more extensive. Many scholars have achieved ideal results in this research and improvement. This paper summarizes the common calibration methods. In the future, with the improvement of the accuracy of the measuring instrument, the improvement of the robot model and the enrichment of the identification and compensation algorithms, the accuracy of the calibration will increase. Robot calibration will tend to be intelligent. Algorithms such as neural networks and fuzzy inference will find more applications.

References

1. Kluz R, Trzepieciński, Tomasz, The repeatability positioning analysis of the industrial robot arm. *Assembly Automation*, 2014, 34(3): P285 - P295.
2. Guanhua Dong, Ying Yin and Xiaobing Hu, Research on Absolute Positioning Error of Robot Based on Mapping Theory. *Applied Mechanics and Materials*, 2014, 494-495: P1156 - P1160.
3. Samad A H, Hayati S, Robot arm geometric link parameter estimation. *IEEE Conference on Decision & Control IEEE*, 1983.
4. Judd R P, Knasinski A B, A technique to calibrate industrial robots with experimental verification. *Robotics and Automation. Proceedings. 1987 IEEE International Conference on*,1987.
5. Stone H W, Sanderson A C, A prototype arm signature identification system. *Robotics and Automation. Proceedings. 1987 IEEE International Conference on*, 1987.
6. Stone H W, Sanderson A C, Statistical performance evaluation of the S-model arm signature identification technique. *Robotics and Automation, 1988. Proceedings. 1988 IEEE International Conference on*, 1988.
7. Zhuang H, Roth Z S, Hamano F, A complete and parametrically continuous kinematic model for robot manipulators. *Robotics and Automation, 1990.*

Proceedings. 1990 IEEE International Conference on, 1990.

8. Klaus Schroer, Stephen L Albright, Michael Grethlein, Complete, minimal and model-continuous kinematic models for robot calibration. *Robotics and Computer Integrated Manufacturing*,1997, 13(1): P73 - P85.
9. I-Ming Chen, Guilin Yang, Chee Tat Tan, Song Huat Yeo, Local POE model for robot kinematic calibration. *Mechanism and Machine Theory*, 2001, 36(11): P1215 - P1239.
10. Dayong Yu, Junwei Han, Dacheng Cong, Kinematic calibration of parallel robot based on coordinate measuring machine. *National Advanced Manufacturing Technology Forum and Manufacturing Automation and Information Technology Seminar.* 2005:5.
11. Hanxu Sun, Yuesheng Tan, Qingxuan Jia, Zhiyu Shao, Study on novel kinematic accuracy test method for manipulator based on electronic theodolite. *Chinese Journal of Scientific Instrument,* 2007, 28(12): P2206 - P2209.
12. Hailong Bai, *POE based Kinematic Calibration for Industrial Robots* (Heilongjiang, Harbin Institute of Technology, 2015)
13. Ceccarelli M, Carrasco C A, Ottaviano E, Error analysis and experimental tests of CATRASYS (CassinoTracking System). *Industrial Electronics Society, 2000. IECON 2000. 26th Annual Confjerence of the IEEE,* 2000.
14. Haoyang Yu, *Study of Manipulator Motion Control Based on Binocular Vision* (Dalian, Dalian University of Technology, 2016)
15. Veitschegger W K, Wu C H, Robot calibration and compensation. *IEEE Journal of Robotics and Automation*, 1988, 4(6): P643 - P656.
16. Grotjahn M, Daemi M, Heimann B, Friction and rigid body identification of robot dynamics. *International Journal of Solids and Structures*, 2001, 38(10): P1889 - P1902.
17. Martinelli A, Tomatis N, Siegwart R, Simultaneous localization and odometry calibration for mobile robot. *Autonomous Robots,* 2007, 22(1): P75 - P85.
18. Hu Zhang, *The Modeling and parameter identification method for calibration of industrial robot* (Heilongjiang, Harbin Institute of Technology, 2015)

Visualization Analysis of Web Crawler Evolution Retrieval Research Based on KG

Zhenjiang Chen[1,2],Jiamian Wang[3], Yizhun Peng[1,2*],Di Li[1,2],Lianchen Zhao[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
[3]College of Artificial Intelligence, Tianjin University of Science and Technology, Tianjin, 300222, China)
** pengyizhun@tust.edu.cn*

Abstract

In order to understand the basic situation and future development trend of domestic research on web crawler technology. By using Citespace information visualization analysis software, 2892 web crawler technical literatures in CNKI information technology database from 2000 to 2018 were data mining. From the aspects of literature time distribution, inter-agency cooperation network analysis, co-citation of authors, co-occurrence of keywords and analysis of research frontiers, this paper draws a map of scientific knowledge and sorts out the research background. This paper intuitively reveals the research status, development path, core research groups and research fields of web crawler technology.
Keywords: web crawler; knowledge graph; CiteSpace; co-occurrence analysis; visualization

1. Introduction

Maybe for many people, "Web Crawler" sounds like a new concept like "big data" or "machine learning". But in fact, web crawlers have a long history, which can be traced back to the birth of the world wide web. At that time, the Internet had no search function. The Internet is just a collection of file transfer protocol (FTP) sites. Users can only browse certain websites to find specific shared files. So developers created an automated program called a web crawler or search engine robot. It can grab all the web pages on the Internet, and then copy the contents of all the pages to the database for indexing.

Web crawler technology is the first step to obtain massive network text resources, is an important acquisition technology tool, and is also an important topic in the field of data information. With the rapid development of the information age, network crawler technology has been continuously developed and improved, which has attracted widespread attention from emerging enterprises and national information units. In order to better show and explain the overall framework of research on web crawler technology, so that domestic scholars can quickly understand its research status and trends. In this paper, the research of web crawler technology is deeply analyzed by using the CNKI information technology database scientific literature and the visualization software CiteSpace of scientific knowledge graph.[1]

2. Data Sources and Analysis Tools

2.1. *Data sources*

CNKI is an international leading network publishing platform integrating journals, doctoral dissertations, master's dissertations, conference papers, newspapers, reference books, yearbooks, patents, standards, traditional Chinese studies and overseas literature resources. The data source of this paper is 2892 articles about describing web crawler technology from 2000 to 2018 in the information technology database of CNKI. On this basis, bibliometric statistics and information mining are carried out. This ensures the relative scientificity of the research results. As shown in Table 1.

Table 1. Data acquisition content

Category	Description
Data Source	CNKI China HowNet Database
Retrieval condition	topic='Web crawler'or'research'
Time limit	2000-2018
Document Category	Journal Papers & Doctoral Dissertations
Number of valid papers	2,892

2.2. *Analysis Tool*

CiteSpace is a visual analysis tool based on Java platform developed by Professor Chen Chaomei of Drexel University. Since it was put forward, it has been widely used in various fields because of its scientificity and universality.[2] In this paper, the author's co-occurrence analysis, mechanism co-occurrence analysis and keyword co-occurrence analysis of 2,892 papers are carried out by CiteSpace. Through clustering view and time zone view, we can draw a scientific knowledge graph of the current situation, hot spots and frontiers of network crawler research field.[3]

3. Results and Analysis

3.1. *Time Distribution of Selected Literatures*

In CNKI, 2892 articles were retrieved by topic retrieval, and the annual trend distribution map of research literature on web crawler technology was obtained by preliminary sorting out.

As shown in Figure 1, the number of CNKI papers published continues to rise, from zero in 2000 and 2001 to two in 2003, achieving zero breakthroughs in China. According to the analysis, in 2000, Salesforce and eBay launched their own API, which can be used by programmers to access and download some public data. Since then, many websites have provided Web APIs to enable people to access their public databases. Therefore, the technology of web crawler has gradually developed. In 2004, Python Beautiful Soup could grab some data without the API provided by the website. Since then, the number of CNKI papers has risen steadily, reaching its peak in 2017, and the number of CNKI papers has reached 356. With the continuous development of network technology, programmers are still developing a way to improve the network crawler to adapt to the expanding network data.

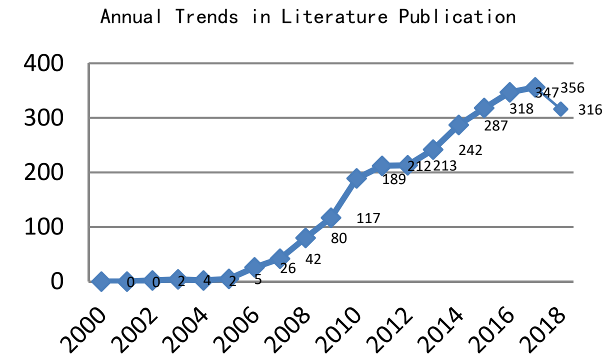


Fig. 1. Annual Trend Distribution of Web Crawler Technology Literature Publication

3.2. Analysis of Cooperation Network of Domestic Research Institutions

Introduce the data into CiteSpace software, select "Institution" as the node, select Top N = 50 as the threshold, set the value of "Years Per Slice" to 3, and select 20% of the data before 3 years for clustering. Run CiteSpace to get the knowledge map of the subordinate organizations of journal papers of web crawlers, and then get Figure 2 and Table 2 after sorting out. A node in the atlas represents an organization unit, and the size of the node represents the number of documents sent by the organization unit. Mediation centrality is an index to measure the importance of nodes in the whole network atlas, which can indicate the importance of documents in the database in CiteSpace visualization software. The greater the centrality of intermediary, the greater its influence in this field.[4]

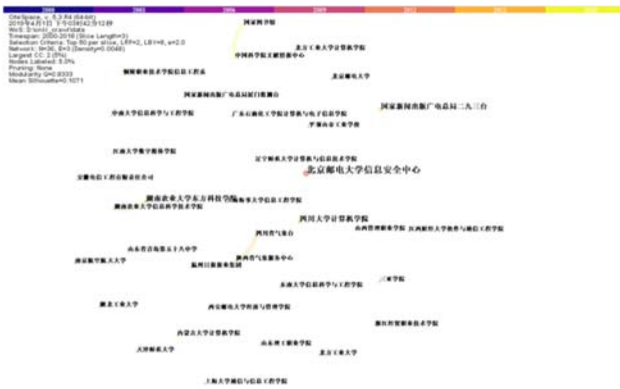


Fig. 2. Knowledge Graph of Publishing Agencies

By analyzing the statistics of the publishing agencies, we can get the universities and institutes which have more research on Web crawlers. As shown in the figure above, the top four in the research of web crawlers in China are the Information Security Center of Beijing University of Posts and Telecommunications, the Computer College of Sichuan University, the Oriental Institute of Science and Technology of Hunan Agricultural University and the 293 National Bureau of Press, Publication, Radio and Television. In terms of the number of articles published, although there are many institutions in our country, they are not prominent, and the overall academic strength is relatively weak. This also shows that more scholars are still needed to improve our country's development in this field.

Table 2. Top 6 in the Number of Institutional Publications

Institutional	Year	Quantity
Information Security Center of Beijing University of Posts and Telecommunications	2010	5
School of Oriental Science and Technology, Hunan Agricultural University	2016	3
School of Computer Science, Sichuan University	2016	3
State Administration of Press, Publication, Radio and Television 293	2016	3
Wenzhou Daily Newspaper Group	2016	2
Nanjing University of Aeronautics and Astronautics	2016	2

3.3. Co-citation Analysis of Authors

To a certain extent, the number of papers published by the author can roughly reflect the scientific research ability and level of the scholar in the relevant fields, and indirectly reflect the research maturity in the corresponding fields.

Based on the data in CNKI's information technology database, the authors of network crawler technology research are analyzed in time domain. Set the value of "Years Per Slice" to 1, and the selection criterion of node type to “Top N=50”. Run CiteSpace to get the distribution time zone map of the important authors in the domestic network crawler research. As shown in Figure 3. Table 3 lists the core researchers of 15 authors who have published more articles, such as Liu Qiang, Zhou Ping and Li Baoguo.

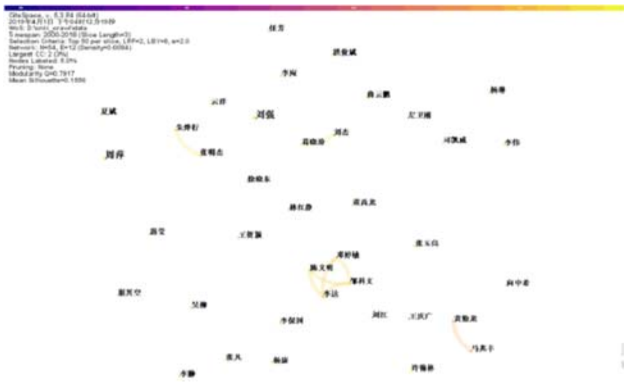


Fig. 3. Knowledge Graph of Publishers

Table 3. Top 10 Authors of Internet Crawler Research in China

Author's Name	Sending the papers amount	Author's Name	Sending the papers amount
Liu Qiang	3	Ren Fang	2
Zhou Ping	3	Chen Yiming	2
Li Baoguo	2	Yunyang	2
Li Ying	2	Xiang Zhongxi	2
Hong Junbin	2	Zhang Yugao	2

3.4. High Frequency Keyword Statistics and Co-occurrence Clustering Analysis

High frequency keywords can be used to analyze the research hotspots of scientific inquiry. Using CiteSpace software, the high-frequency keywords of 2,892 research papers from 2000 to 2018 are counted and visualized knowledge maps are drawn. Table 4 lists the top 10 and the middle 10 (parts) and the last 10 high-frequency keywords, years, frequencies and intermediary centrality (abbreviated as "neutrality").

Table 4. Keyword Distribution of Web Crawler Related Literature (Part)

Frequency	Year	Intermedia teness	Key word
713	2004	0.21	Internet worm
238	2006	0.12	Search Engines
135	2007	0.17	Topical crawler
122	2009	0.13	Reptile
101	2008	0.09	Vertical Search Engine
96	2008	0.13	Internet public opinion
85	2008	0.19	Text classification
79	2008	0.11	Chinese word segmentation
78	2010	0.11	Data mining

73	2008	0.07	Information extraction
37	2006	0.08	Information Retrieval
35	2012	0.05	Text clustering
35	2010	0.07	Distributed Reptilesl
33	2015	0.02	Big data
32	2007	0.07	Vector Space Mode
30	2010	0.01	Text Mining
29	2015	0.01	Scrapy
29	2011	0.05	Machine learning
27	2009	0.04	Ajax
26	2011	0.03	Topical crawler
6	2013	0	Web Services
5	2016	0	User Behavior
5	2017	0	Word vector
4	2013	0	Visualization
2	2013	0	Web Page Clustering
2	2015	0	Management information systems
2	2015	0	Task scheduling
2	2014	0	Tendency analysis
2	2015	0	Doctor's recommendation
2	2010	0	Web Page Purification

Fig. 4. Keyword Co-occurrence Map in the Field of Web Crawler Technology

The keywords with high frequency and high intermediary centrality represent the new research frontiers in the corresponding time interval in the field of network crawler research to a certain extent, and their diachronic evolution reflects the dynamic changes of hot spots in a discipline or research field.

Combining with the emergence of Fig. 4, Fig. 5 and Table 4, we can see that the research hotspots of web crawlers maybe: web crawler--vertical search engine--topic crawler--text categorization--distributed crawler--vector space model--machine earning--deep learning--knowledge graph (knowledge base). Of course, some parts are still in-depth research, such as topic crawler research will focus on how to improve the accuracy of web page correlation calculation, how to reduce the space-time complexity of calculation, and how to enhance the adaptability of crawler.[5] At the same time, it can be found that network data mining based on large data crawler technology provides basic data support for machine learning, deep learning and knowledge atlas.[6]

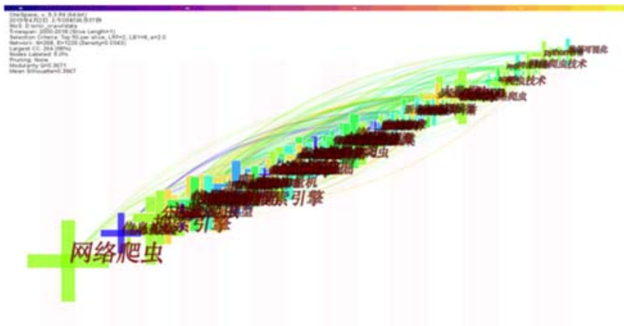


Fig. 5. Keyword Sequence Map Atlas

As shown in the figure above, the cold and warm colors of the graph indicate the relationship between time and time, that is, the colder the color is, the farther the research content is from now, and the warmer the color is, the trend of current research is indicated.

3.5. Research Frontier Analysis

In the era of big data, data sources are the first thing to do data analysis, and the network resources are the most abundant. So we often need to do network crawlers, which can let us get more data sources, and these data sources can be collected according to our purposes, removing a lot of irrelevant data.

Web crawler technology has now formed a relatively complete technical system, different types of Web sites for different forms of data capture. Then further processing of the captured data, such as data cleaning, is the reprocessing of the captured results, which can effectively improve the quality of the captured data, especially in the heterogeneous data source environment of the World Wide Web.[7] Secondly, different types of databases are used to store different types of data, such as relational databases and non-relational databases. Then the valid data is processed. Text mining, for example, is an interdisciplinary subject, involving data mining, machine learning, pattern recognition, artificial intelligence, statistics, computer linguistics, computer network technology, informatics and other fields.[8] It is a method and tool to discover implicit knowledge and patterns from many documents. It is developed from data mining, but it is different from traditional data mining.[9] Machine Learning, for example, refers to the process of using some algorithms to guide computers to use known data to obtain appropriate models, and to use this model to give judgments on new situations.[10] The idea of machine learning is not complicated, it is only a simulation of the learning process in human life, and in the whole process, the most critical is data.

4. Conclusions

This paper analyses the amount of literature, source journals, authors and organizations, keyword analysis and hot spot analysis of web crawlers, and draws the following conclusions:

4.1. Further research is needed

From the point of view of research basis, with the continuous development of the era of big data, there are 2892 core journals on Web crawler, and the research in the field of web crawler technology has begun to be saturated. However, the evolution of machine learning, in-depth learning and knowledge mapping needs further research and more attention.

4.2. Institutions and authors

Although there are many institutions in our country to study the technology of web crawler, they are less cited, the overall academic strength is relatively weak, and the communication between authors is not very close, and the total cited number of authors is also very small, indicating that more scholars are still needed to step into and improve the field.

4.3. High frequency keyword statistics and co-occurrence cluster analysis

Through keyword co-occurrence analysis, we can see that the research on the evolution of web crawler can be divided into three levels: basic crawler technology level, search crawler level and application level of crawler. Basic crawler technology includes subject crawler, information search, text categorization and distributed crawler. Search crawler level includes vertical search engine, information retrieval and subject retrieval. The application level of crawler includes text mining, machine learning, in-depth learning, personalized recommendation, natural language processing and knowledge graph.

4.4. Research hotspots and frontiers

Through the analysis of the research frontier, it is found that the research frontier of web crawler technology is mainly in the fields of data visualization, machine learning, in-depth learning and knowledge atlas.

References

1. Li Jie, Chen Chaomei. Cite Space: Technological Text Mining and Visualization [M]. Beijing: Capital University of Economics and Trade Press, 2016:32.
2. Li Jie. CiteSpace Chinese Guide [M/OL] http://blog.sciencenet.cn/blog-554179-1066981.html.
3. Chen Chaomei. CiteSpace Ⅱ. Detecting and Visualizing Emerging Trends and Transient Patterns in ScientificLiteratur[J]. Journalof the American Society for Information Science and Technology,2006(3).10.
4. Lin Deming, Chen Chaomei and Liu Zeyuan. A study of Zipf-Pareto distribution of network intermediary centrality [J]. Informatics, 2011, 30 (1): 76-82.
5. Zhong Weijin, Li Jia. Co-word analysis: the process and method of Co-word Analysis [J]. Information Journal, 2008, 27 (5): 70-72.
6. Pan Xiaoying, Chen Liu, Yu Huimin, Zhao Yizhao, Xiao Kangzhen. Summary of research on topic crawler technology [J/OL]. Computer application research: 1-6 [2019-04-02].https://doi.org/10.19734/j.issn.1001-3695.2018.11.0790.
7. Zhou Lizhu, Lin Ling. Review of focused crawler technology [J].Computer Applications, 2005 (09): 1965-1969.
8. Wang Minxiang. Research and implementation of topic-oriented crawling search strategy [D]. Shaanxi Normal University, 2012.
9. Feng Li.Overview of Reptilian Technology[J].Computer Knowledge and Technology, 2017,13(27): 213-214.
10. Shi Hongyi. Overview of Machine Learning [J]. Communication World, 2018 (10): 253-254.

Circuit Simulation of Synchronized Novel 4D Chaotic Systems

Yuhan Zhang, 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 four-dimensional (4D) autonomous chaotic systems, based on the center translation method, is presented. The analog circuit model of the synchronization system is constructed. The numerical and circuitry simulation results are given to illustrate the validity of the synchronization circuitry.

Keywords: novel 4D chaotic system, chaos synchronization, center translation method, synchronization circuitry

1. Introduction

Chao synchronization has great engineering significance in nonlinear dynamics, especially in secure communication.[1-3] Circuit implementation of chaotic synchronization system has been one of the major research subjects.

In this paper, synchronization of novel 4D chaotic systems, based on the center translation method, is proposed. Furthermore, corresponding analog synchronization circuitry is constructed to demonstrate the realizability of the synchronization control law.

2. Novel 4D Chaotic System

The novel 4D chaotic system has been discussed in Ref. 4. It is formulated as

$$\begin{aligned} \dot{x} &= a(y-x), \\ \dot{y} &= c(x+y)+z-xw, \\ \dot{z} &= mx-y-hz, \\ \dot{w} &= xy-bw, \end{aligned} \tag{1}$$

where $a = 25$, $b = 3$, $c = 18$, $m = 19$, and $h = 14$. The phase portraits of the system (1) are shown in Fig. 1.

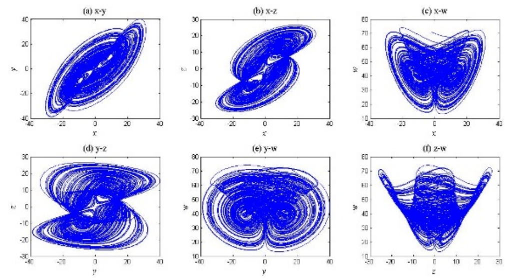


Fig. 1. Phase portraits of the novel 4D chaotic system: (a) *x*-*y*; (b) *x*-*z*; (c) *x*-*w*; (d) *y*-*z*; (e) *y*-*w*; (f) *z*-*w*

3. Chaos Synchronization Based on Center Translation Method

3.1. Control Law

Take the system (1) as the drive system. From Ref. 5, the response system can be formulated as

$$\begin{aligned} \dot{x}_2 &= a(y_2-x_2), \\ \dot{y}_2 &= c(x_2+y_2)+z_2-x_2w_2+u_{c1}, \\ \dot{z}_2 &= mx_2-y_2-hz_2, \\ \dot{w}_2 &= x_2y_2-bw_2+u_{c2}, \end{aligned} \tag{2}$$

where

$$u_{c1} = -k\left(y_2 - y\right) + x_2 w + x w_2 - 2xw, \quad (3)$$
$$u_{c2} = -x_2 y - x y_2 + 2xy, \quad (4)$$

and $k > 0$.

Theorem 1. *The state variables* x_2, y_2, z_2 *and* w_2 *of the response system* (2) *will be synchronized with those of the drive system* (1) *correspondingly as long as*

$$k > \frac{h\left(a+c\right)^2}{4ah - m^2} + c > \frac{\left(a+c\right)^2}{4a} + c.$$

Proof. Substitute Eq. (3) and Eq. (4) into the response system (2), and let

$$\boldsymbol{e} = \begin{bmatrix} e_1 & e_2 & e_3 & e_4 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} x_2 - x & y_2 - y & z_2 - z & w_2 - w \end{bmatrix}^{\mathrm{T}}$$

be the synchronization errors. Then the error system is presented as

$$\begin{aligned} \dot{e}_1 &= a\left(e_2 - e_1\right), \\ \dot{e}_2 &= c\left(e_1 + e_2\right) + e_3 - e_1 e_4 - k e_2, \\ \dot{e}_3 &= m e_1 - e_2 - h e_3, \\ \dot{e}_4 &= e_1 e_2 - b e_4. \end{aligned} \quad (5)$$

Take a positive definite function

$$\mathrm{V}\left(\boldsymbol{e}\right) = \frac{1}{2}\left(e_1^2 + e_2^2 + e_3^2 + e_4^2\right)$$

as a Lyapunov function candidate for the error system (5), then the derivative $\dot{\mathrm{V}}\left(\boldsymbol{e}\right)$ is derived as

$$\begin{aligned} \dot{\mathrm{V}}\left(\boldsymbol{e}\right) &= e_1\dot{e}_1 + e_2\dot{e}_2 + e_3\dot{e}_3 + e_4\dot{e}_4 \\ &= -a e_1^2 + \left(a+c\right)e_1 e_2 - \left(k-c\right)e_2^2 - h e_3^2 \\ &\quad + m e_1 e_3 - b e_4^2 \\ &= \boldsymbol{e}^{\mathrm{T}} \begin{bmatrix} -a & \frac{a+c}{2} & \frac{m}{2} & 0 \\ \frac{a+c}{2} & -\left(k-c\right) & 0 & 0 \\ \frac{m}{2} & 0 & -h & 0 \\ 0 & 0 & 0 & -b \end{bmatrix} \boldsymbol{e} \\ &= \boldsymbol{e}^{\mathrm{T}} Q \boldsymbol{e} \end{aligned}$$

where

$$Q = \begin{bmatrix} -a & \frac{a+c}{2} & \frac{m}{2} & 0 \\ \frac{a+c}{2} & -\left(k-c\right) & 0 & 0 \\ \frac{m}{2} & 0 & -h & 0 \\ 0 & 0 & 0 & -b \end{bmatrix}.$$

When

$$k > \frac{h\left(a+c\right)^2}{4ah - m^2} + c > \frac{\left(a+c\right)^2}{4a} + c,$$

the leading principal minors of the matrix Q satisfy all of the following conditions, i.e.,

$$\sigma_1 = -a < 0,$$

$$\sigma_2 = a\left(k-c\right) - \frac{\left(a+c\right)^2}{4} > 0,$$

$$\sigma_3 = \left(\frac{m^2}{4} - ah\right)\left(k-c\right) + \frac{h\left(a+c\right)^2}{4} < 0,$$

$$\sigma_4 = -b\sigma_3 > 0.$$

It means that the matrix Q is negative definite, so that the derivative $\dot{\mathrm{V}}\left(\boldsymbol{e}\right)$ is negative definite. Since the Lyapunov function $\mathrm{V}\left(\boldsymbol{e}\right)$ is positive definite and radially unbounded, it can be confirmed that the error system (5) is globally asymptotically stable at the origin, i.e., the synchronization errors $e1$, $e2$, $e3$ and $e4$ globally asymptotically converge to 0. Thus the state variables x_2, y_2, z_2 and w_2 of the response system (2) are synchronized with those of the drive system (1) correspondingly. The proof of Theorem 1 is completed.

3.2. Numerical Simulation

Remark 1. k = 50 *in the following numerical and circuitry simulation.*

Remark 2. *The initial values of the drive system* (1) *and the response system* (2) *are* (x_0, y_0, z_0, w_0) = (1, 1, 1, 1) *and* (x_{20}, y_{20}, z_{20}, w_{20}) = (5, 4, 7, 8) *respectively in this paper.*

The phase portraits of x_1-x_2, y_1-y_2, z_1-z_2 and w_1-w_2 before and after adding u_{c1} and u_{c2} 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 state variables x_2, y_2, z_2 and w_2 of the response system (2) are synchronized with those of the drive system (1) correspondingly under the control of u_{c1} and u_{c2}.

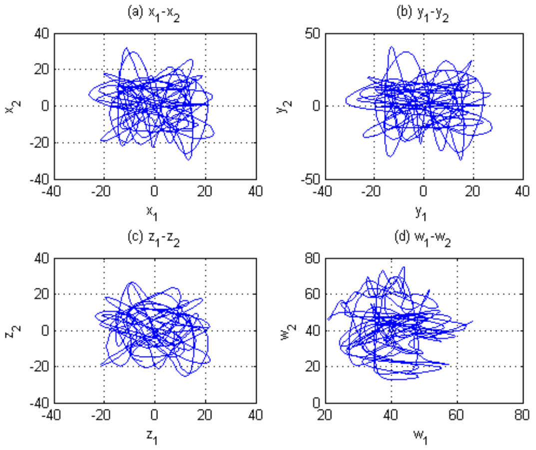


Fig. 2. Phase portraits before adding controllers: (a) x_1-x_2; (b) y_1-y_2; (c) z_1-z_2; (d) w_1-w_2

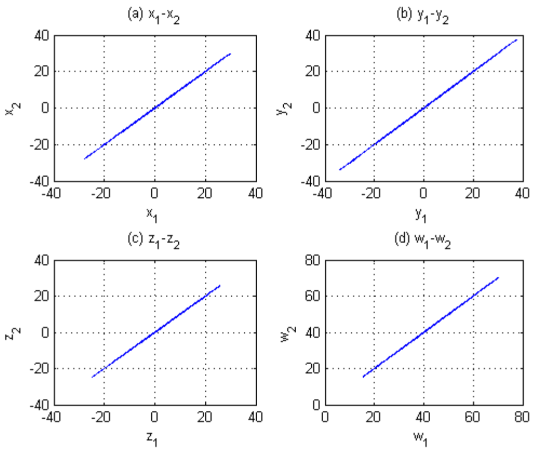


Fig. 3. Phase portraits after adding controllers: (a) x_1-x_2; (b) y_1-y_2; (c) z_1-z_2; (d) w_1-w_2

4. Design of Synchronization Circuitry

4.1. Circuit Design of Drive System

Circuit design procedure of the drive system (1) has been discussed in Ref. 6. The normalization model of the drive system (1) is represented as

$$\begin{aligned}
\frac{d\tilde{x}}{dt} &= -a\tau_0\tilde{x} - a\tau_0(-\tilde{y}),\\
\frac{d\tilde{y}}{dt} &= -c\tau_0(-\tilde{x}) - c\tau_0(-\tilde{y}) - \tau_0(-\tilde{z}) - 10\tau_0\tilde{x}\tilde{w},\\
\frac{d\tilde{z}}{dt} &= -m\tau_0(-\tilde{x}) - \tau_0\tilde{y} - h\tau_0\tilde{z},\\
\frac{d\tilde{w}}{dt} &= -10\tau_0(-\tilde{x})\tilde{y} - b\tau_0\tilde{w},
\end{aligned} \tag{6}$$

where $\tilde{x} = 0.1x$, $\tilde{y} = 0.1y$, $\tilde{z} = 0.1z$, $\tilde{w} = 0.1w$, and $\tau_0 = 100$.

From Ref. 6, the analog circuit model of system (6) is designed as

$$\begin{aligned}
\frac{d\tilde{x}}{dt} &= -\frac{1}{R_1C_1}\tilde{x} - \frac{1}{R_2C_1}(-\tilde{y}),\\
\frac{d\tilde{y}}{dt} &= -\frac{1}{R_3C_2}(-\tilde{x}) - \frac{1}{R_4C_2}(-\tilde{y}) - \frac{1}{R_5C_2}(-\tilde{z})\\
&\quad - \frac{1}{10R_6C_2}\tilde{x}\tilde{w},\\
\frac{d\tilde{z}}{dt} &= -\frac{1}{R_7C_3}(-\tilde{x}) - \frac{1}{R_8C_3}\tilde{y} - \frac{1}{R_9C_3}\tilde{z},\\
\frac{d\tilde{w}}{dt} &= -\frac{1}{10R_{10}C_4}(-\tilde{x})\tilde{y} - \frac{1}{R_{11}C_4}\tilde{w}.
\end{aligned} \tag{7}$$

where

$$\begin{aligned}
&C_1 = C_2 = C_3 = C_4 = 10\text{nF},\\
&R_x = R_{fx} = R_y = R_{fy} = R_z = R_{fz} = 10\text{k}\Omega,\\
&R_1 = R_2 = 40\text{k}\Omega,\\
&R_3 = R_4 = 55.5\text{k}\Omega,\\
&R_5 = R_8 = 1\text{M}\Omega,\\
&R_6 = R_{10} = 10\text{k}\Omega,\\
&R_7 = 52.6\text{k}\Omega,\\
&R_9 = 71.4\text{k}\Omega,\\
&R_{11} = 333.3\text{k}\Omega.
\end{aligned}$$

The circuit schematic diagram of the drive system (7) is shown in the left orange box of Fig. 4.

4.2. Circuit Design of Response System

Similarly, the normalization model of the response system (2) with Eq. (3) and Eq. (4) can be formulated as

$$\frac{\mathrm{d}\tilde{x}_2}{\mathrm{d}t} = -a\tau_0\tilde{x}_2 - a\tau_0\left(-\tilde{y}_2\right),$$

$$\frac{\mathrm{d}\tilde{y}_2}{\mathrm{d}t} = -c\tau_0\left(-\tilde{x}_2\right) - c\tau_0\left(-\tilde{y}_2\right) - \tau_0\left(-\tilde{z}_2\right) - 10\tau_0\tilde{x}_2\tilde{w}_2 - 10\tau_0\left[-\left(-\frac{k}{10}\right)\tilde{y}_2 - \left(-\frac{k}{10}\right)\left(-\tilde{y}\right) - \tilde{x}_2\tilde{w} - \tilde{x}\tilde{w}_2 - 2\left(-1\right)\left(\tilde{x}\tilde{w}\right)\right],$$

$$\frac{\mathrm{d}\tilde{z}_2}{\mathrm{d}t} = -m\tau_0\left(-\tilde{x}_2\right) - \tau_0\tilde{y}_2 - h\tau_0\tilde{z}_2,$$

$$\frac{\mathrm{d}\tilde{w}_2}{\mathrm{d}t} = -10\tau_0\left(-\tilde{x}_2\right)\tilde{y}_2 - b\tau_0\tilde{w}_2 - 10\tau_0\left[-\left(-\tilde{x}_2\right)\tilde{y} - \tilde{x}\left(-\tilde{y}_2\right) - 2\left(-1\right)\left(-\tilde{x}\tilde{y}\right)\right], \tag{8}$$

where $\tilde{x}_2 = 0.1x_2$, $\tilde{y}_2 = 0.1y_2$, $\tilde{z}_2 = 0.1z_2$, and $\tilde{w}_2 = 0.1w_2$.

Accordingly, the analog circuit model of the system (8) can be designed as

$$\frac{\mathrm{d}\tilde{x}_2}{\mathrm{d}t} = -\frac{1}{\mathrm{R}_{12}\mathrm{C}_5}\tilde{x}_2 - \frac{1}{\mathrm{R}_{13}\mathrm{C}_5}\left(-\tilde{y}_2\right),$$

$$\frac{\mathrm{d}\tilde{y}_2}{\mathrm{d}t} = -\frac{1}{\mathrm{R}_{14}\mathrm{C}_6}\left(-\tilde{x}_2\right) - \frac{1}{\mathrm{R}_{15}\mathrm{C}_6}\left(-\tilde{y}_2\right) - \frac{1}{\mathrm{R}_{16}\mathrm{C}_6}\left(-\tilde{z}_2\right) - \frac{1}{10\mathrm{R}_{17}\mathrm{C}_6}\tilde{x}_2\tilde{w}_2$$
$$-\frac{1}{\mathrm{R}_{31}\mathrm{C}_6}\left[-\frac{\mathrm{R}_{33}}{\mathrm{R}_{28}}\left(-\frac{\mathrm{R}_{23}}{\mathrm{R}_{24}}\right)\tilde{y}_2 - \frac{\mathrm{R}_{33}}{\mathrm{R}_{28}}\left(-\frac{\mathrm{R}_{23}}{\mathrm{R}_{25}}\right)\left(-\tilde{y}\right) - \frac{\mathrm{R}_{33}}{10\mathrm{R}_{29}}\tilde{x}_2\tilde{w} - \frac{\mathrm{R}_{33}}{10\mathrm{R}_{30}}\tilde{x}\tilde{w}_2 - \frac{\mathrm{R}_{33}}{10\mathrm{R}_{32}}\left(-\frac{\mathrm{R}_{27}}{\mathrm{R}_{26}}\right)\left(\tilde{x}\tilde{w}\right)\right],$$

$$\frac{\mathrm{d}\tilde{z}_2}{\mathrm{d}t} = -\frac{1}{\mathrm{R}_{18}\mathrm{C}_7}\left(-\tilde{x}_2\right) - \frac{1}{\mathrm{R}_{19}\mathrm{C}_7}\tilde{y}_2 - \frac{1}{\mathrm{R}_{20}\mathrm{C}_7}\tilde{z}_2,$$

$$\frac{\mathrm{d}\tilde{w}_2}{\mathrm{d}t} = -\frac{1}{10\mathrm{R}_{21}\mathrm{C}_8}\left(-\tilde{x}_2\right)\tilde{y}_2 - \frac{1}{\mathrm{R}_{22}\mathrm{C}_8}\tilde{w}_2 - \frac{1}{\mathrm{R}_{39}\mathrm{C}_8}\left[-\frac{\mathrm{R}_{40}}{10\mathrm{R}_{36}}\left(-\tilde{x}_2\right)\tilde{y} - \frac{\mathrm{R}_{40}}{10\mathrm{R}_{37}}\tilde{x}\left(-\tilde{y}_2\right) - \frac{\mathrm{R}_{40}}{10\mathrm{R}_{38}}\left(-\frac{\mathrm{R}_{34}}{\mathrm{R}_{35}}\right)\left(-\tilde{x}\tilde{y}\right)\right]. \tag{9}$$

where

$$\mathrm{C}_5 = \mathrm{C}_6 = \mathrm{C}_7 = \mathrm{C}_8 = 10\mathrm{nF},$$
$$\mathrm{R}_{x1} = \mathrm{R}_{fx1} = \mathrm{R}_{y1} = \mathrm{R}_{fy1} = \mathrm{R}_{z1} = \mathrm{R}_{fz1} = 10\mathrm{k\Omega},$$
$$\mathrm{R}_{12} = \mathrm{R}_{13} = 40\mathrm{k\Omega},$$
$$\mathrm{R}_{14} = \mathrm{R}_{15} = 55.5\mathrm{k\Omega},$$
$$\mathrm{R}_{16} = \mathrm{R}_{19} = 1\mathrm{M\Omega},$$
$$\mathrm{R}_{17} = \mathrm{R}_{21} = 10\mathrm{k\Omega},$$
$$\mathrm{R}_{18} = 52.6\mathrm{k\Omega},$$
$$\mathrm{R}_{20} = 71.4\mathrm{k\Omega},$$
$$\mathrm{R}_{22} = 333.3\mathrm{k\Omega},$$
$$\mathrm{R}_{23} = 50\mathrm{k\Omega},$$
$$\mathrm{R}_{24} = \mathrm{R}_{25} = \mathrm{R}_{32} = \mathrm{R}_{38} = 10\mathrm{k\Omega},$$
$$\mathrm{R}_{26} = \mathrm{R}_{27} = \mathrm{R}_{34} = \mathrm{R}_{35} = 10\mathrm{k\Omega},$$
$$\mathrm{R}_{28} = \mathrm{R}_{33} = \mathrm{R}_{40} = 200\mathrm{k\Omega},$$
$$\mathrm{R}_{29} = \mathrm{R}_{30} = \mathrm{R}_{36} = \mathrm{R}_{37} = 20\mathrm{k\Omega},$$
$$\mathrm{R}_{31} = \mathrm{R}_{39} = 100\mathrm{k\Omega}.$$

The circuit schematic diagram of the response system (9) is shown in Fig. 4, where the response circuit is in the right blue box, and the synchronization control circuits are in the middle green box.

4.3. Circuitry Simulation

The phase portraits of $\tilde{x}_1 - \tilde{x}_2$, $\tilde{y}_1 - \tilde{y}_2$, $\tilde{z}_1 - \tilde{z}_2$ and $\tilde{w}_1 - \tilde{w}_2$ before and after adding the synchronization control circuits to the response circuit are shown in Fig. 5 and Fig. 6 respectively. Comparing Fig. 6 with Fig. 5, it can be found that the state variables $\tilde{x}_2$, $\tilde{y}_2$, $\tilde{z}_2$ and $\tilde{w}_2$ of the response circuit are synchronized with those of the drive circuit correspondingly under the control of the synchronization control circuits.

Remark 3. *In Fig.* 5 *and Fig.* 6, *the scales of* $\tilde{x}_1$, $\tilde{y}_1$, $\tilde{z}_1$ *and* $\tilde{w}_1$, *as well as* $\tilde{x}_2$, $\tilde{y}_2$, $\tilde{z}_2$ *and* $\tilde{w}_2$, *are* 2 V/Div, 2 V/Div, 2 V/Div *and* 5 V/Div, *respectively.*

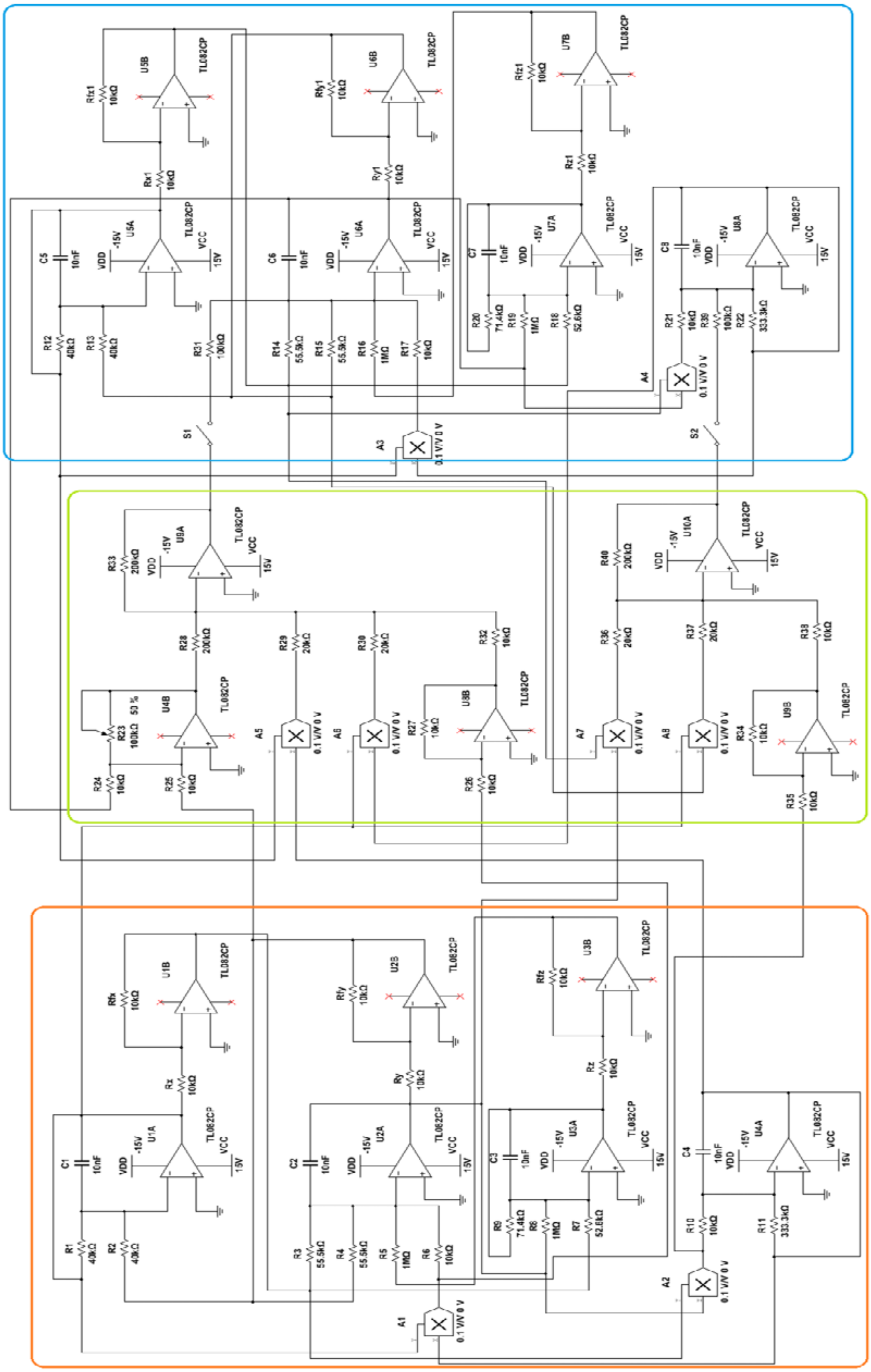


Fig. 4. Circuit schematic diagram of synchronization system

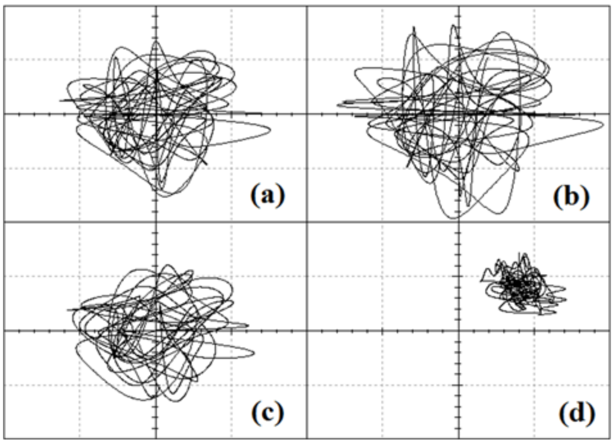


Fig. 5. Phase portraits before adding control circuits:
(a) $\tilde{x}_1 - \tilde{x}_2$; (b) $\tilde{y}_1 - \tilde{y}_2$; (c) $\tilde{z}_1 - \tilde{z}_2$; (d) $\tilde{w}_1 - \tilde{w}_2$

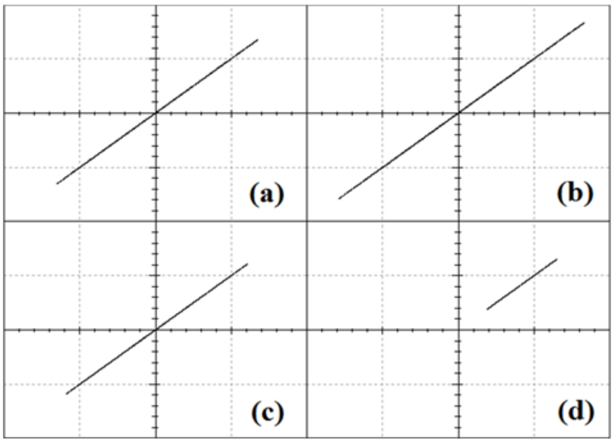


Fig. 6. Phase portraits after adding control circuits:
(a) $\tilde{x}_1 - \tilde{x}_2$; (b) $\tilde{y}_1 - \tilde{y}_2$; (c) $\tilde{z}_1 - \tilde{z}_2$; (d) $\tilde{w}_1 - \tilde{w}_2$

5. Conclusion

The phase portraits of numerical and circuitry simulation illustrate the good qualitative agreement between the mathematic model and the circuitry model. It shows that the designed synchronization circuitry is correct. However, circuit design of the control law is complex and difficult, because the synchronization controllers are nonlinear. In future research, the synchronization circuitry constructed in this paper will be implemented by hardware, and other control laws will be studied to conquer the complexity in circuit implementation.

Acknowledgements

The research is supported by the Innovation Fund for the Young Teachers of Tianjin University of Science & Technology.

References

1. Hamed Tirandaz, Saeid Saeidi Aminabadi and Hamidreza Tavakoli, Chaos synchronization and parameter identification of a finance chaotic system with unknown parameters, a linear feedback controller, *Alexandria Engineering Journal*, 2018, 57(3): 1519-1524.
2. Saleh Mobayen, Chaos synchronization of uncertain chaotic systems using composite nonlinear feedback based integral sliding mode control, *ISA Transactions*, 2018, 77: 100-111.
3. Hong Niu, Guoshan Zhang and Jiankui Wang, Chaos synchronization of Chua's circuit and Lorenz system based on strictly positive realness, *Proceedings of the 33rd Chinese Control Conference*, Nanjing, China, 2014.7.28-7.30.
4. Hong Niu, Analysis and control of a novel 4D chaotic system, *Proceedings of the 2017 International Conference on Artificial Life and Robotics*, Miyazaki, Japan, Jan. 2017, pp. 165-168.
5. Hong Niu, Synchronization of novel 4D chaotic systems with different control laws, *Proceedings of the 2019 International Conference on Artificial Life and Robotics*, Oita, Japan, Jan. 2019.
6. Dechu Tan and Hong Niu, Analog circuit design of a novel 4D chaotic system, *Proceedings of the 2018 International Conference on Artificial Life and Robotics*, Oita, Japan, Feb. 2018, pp.574-577.

Crowd Counting Network with Self-attention Distillation

Li Wang, Huailin Zhao, Zhen Nie, Yaoyao Li

School of Electrical and Electronic Engineering, Shanghai Institute of Technology, Shanghai, China
E-mail: 736036396@qq.com, zhao_huailin@yahoo.com
www.sit.edu.cn

Abstract

Context information is essential for crowd counting network to estimate crowd numbers, especially in the congested scene accurately. However, shallow layers of common crowd counting networks (i.e., Congested Scene Recognition Network (CSRNet)) don't own large receptive filed so that they can't efficiently utilize context information from the crowd scene. To solve this problem, in this paper, we propose a crowd counting network with self-attention distillation (SADNet). Each input image is firstly sent to the VGG-16 network for feature extracting. Then, the extracted features are processed by the dilated convolutional part for the final crowd density estimation. Specially, we apply self-attention distillation strategy at different locations of the dilated convolutional part to use the global context information from the deeper layers to guide the shallower layers to learn. We compare our method with the other state-of-the-art works on the UCF-QNRF dataset, and the experiment results demonstrate the superiority of our method.

Keywords: Self-attention Distillation, Dilated convolution, Crowd counting

1. Introduction

With the rapid growth of urban population, large-scale high-density assembly scenes are increasing, and the crowd gathering behavior is becoming more frequent and larger, which brings great difficulties and challenges to urban security systems. In order to deal with a large number of crowd monitoring data in a timely and effective manner, to prevent accidents and reduce hidden dangers in public places, crowd density estimation technology has become a research focus in the field of intelligent security.

The goal of the crowd density estimation algorithm is to estimate the number of individuals in the crowd in the entire image range through certain technical means. However, due to occlusion, perspective distortion, scale changes, and the diversity of group distribution, accurate crowd counting has always been a challenging problem in computer vision.

Traditional crowd counting algorithms are mostly based on detection and regression models. The crowd counting algorithm based on the detection model is more suitable for low-density crowd counting and has little effect on high-density scenes, similar to the pedestrian detection work. Pedestrians in the picture are detected by a pedestrian detector, and the number of detected persons is counted to calculate the total number of persons. In the crowd counting work based on the regression model, the main factors affecting the counting accuracy are the feature extraction method and the selection of the regression model. A regression algorithm is used on the extracted crowd features to establish a mapping relationship between the features and the number of people, so that the trained regressor has the ability to calculate the number of crowds. Although this method has made great progress in crowd counting, it cannot fully utilize the spatial information of the crowd, and it is still difficult to meet the accuracy requirements in dense scenes.

In recent years, due to the powerful feature expression capability and flexible architecture of neural networks, its application in the field of population density estimation has become more and more mature. By training a convolutional neural network, extracting crowd picture features, generating a corresponding crowd density map, and summing all pixels in the density map to calculate the total number of people. In order to solve the problem of population scale change, most previous counting networks used a multi-column network structure. For example, Zhang et al. [1] used multi-column convolutional neural network to capture different scales of crowd heads. Since then, in this way, [2] proposed a Hydra network structure, which used multi-column convolutional neural networks to extract population characteristics of different size tiles to obtain scale information. Reference [3] adds density pre-classification network on the basis of reference [2]. Although multi-column structures improved the counting performance, they usually had a large amount parameters, which needs much time to train. Therefore, [4] improved the convolution operation, and proposed a CSRNet network, which uses dilated convolution to expand the receptive field range, thereby extracting deeper feature information without increasing the amount of data. Reference [5] integrates attention mechanism on the basis of CSRNet to increase context information.

Contextual information is the key to accurately estimating the number of people in a crowd counting network, especially in crowded scenarios. However, the shallow layer of the common crowd counting network (such as a crowded scene recognition network (CSRNet[4])) does not have a large receiving field, so it cannot effectively use the context information in the crowd scene. To solve this problem, this paper proposes a crowd counting network with self-attention knowledge distillation (SADNet). Each input image is first sent to the VGG-16 network for feature extraction, and then the dilated convolution is used to process the extracted features. Three different attention generators have been added at different positions in the part of the dilated convolution. Use self-attention distillation strategy to obtain deeper global context information to guide shallow learning and obtain higher quality crowd density maps.

2. Proposed Method

The method of knowledge distillation was originally proposed by [6], with the purpose of transferring knowledge from large networks to small networks. By introducing soft targets related to large teacher networks as part of the total loss, the training of small student networks is induced to achieve knowledge transfer. Recent research has gradually expanded knowledge distillation to attention distillation. Based on the self-attention distillation method proposed in [7], this paper applies attention distillation to crowd density estimation and designs a new crowd counting network with self-attention distillation. The network model can learn from itself, without any additional supervision or label annotation, to obtain deeper global context information, make the population density map estimated by the network more similar to the true value, and improve counting accuracy. The overall network structure is shown in Fig.1.

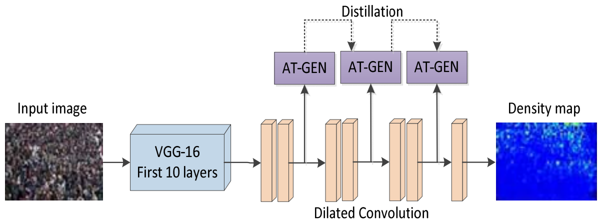


Fig.1.The overall flowchart of the algorithm.

As shown in Fig.1., we apply the first ten layers of VGG16 to extract features. The details of VGG-16 are shown in Fig.2.

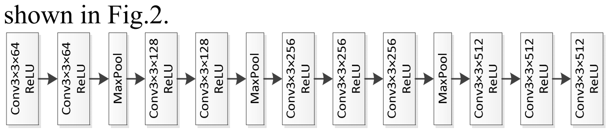


Fig.2. The first ten convolution layers of the VGG-16 network.

After shallow feature extraction, a six-layer hole convolution operation is used on the feature map, and then a 1×1 convolution is used to generate the final crowd density map, as shown in pink in Fig 1. The dilated convolution uses a 3×3 convolution kernel with an expansion ratio of 2, which is equivalent to six layers of 5×5 convolution layers. The number of channels in each layer is set to {512, 512, 512, 256, 128, 64}. The difference is that the dilated convolutional layer does not increase the amount of network calculations and avoids the loss of resolution caused by continuing the pooling operation.

The biggest feature of this paper is the application of the self-attention distillation strategy at different positions of the hollow convolutional layer, as shown in the purple part in Fig 1. After the second, fourth, and sixth dilated convolution operations, the feature map is transformed into three 1-channel attention maps by the attention generator. Using the second attention map as the true value and the first attention map as the output estimate, calculate the L2 loss between the two. Similarly, the third attention map is used as the true value, and the second attention map is used as the output estimation value. The L2 loss between the two is calculated. The calculated losses are fed back to the network to refine the low-level feature maps. At the same time, the feature information learned by the lower layers further improves the deeper performance of the network. Different context information is captured through attention mapping from different layers, which improves the similarity between the network-generated density map and the true-value density map. Using the network's own attention map as a distillation target does not require additional external supervision and does not increase the training time of the basic model.

Finally, the network outputs an estimated population density map. Sum all pixels in the density map to calculate the total number of people in the crowd picture.

3. Experiments

3.1. *Dataset*

This experiment uses the latest dataset UCF-QNRF[8] for crowd counting to train and test the network. The UCF-QNRF data set contains a total of 1,535 crowd pictures, which are divided into two parts, the training set and the test set. There are 1201 pictures in the training set and 334 pictures in the test set. These pictures not only contain a large number of people, but also complex background information such as buildings, vegetation, sky, and roads, making the data set closer to reality, and increasing the difficulty of counting people.

3.2. *Density Map Generation*

The crowd head annotations in the dataset are converted to true-value density maps, using the most popular adaptive Gaussian model currently available. Gaussian kernel blurs each head annotation, normalizes the sum to 1, so that all pixels in the final density map can be summed to obtain the total number of people. The formula is as follows:

$$F(x)=\sum_{i}^{N}\delta(x-x_i)*G_{\sigma_i}(x),\sigma_i=\beta\bar{d}^{\,l} \quad (1)$$

Where $G_{\sigma_i}(x)$ represents the 2D – Gaussian kernel, x_i is the position coordinates of the human head in the image, $\delta(x-x_i)$ is the Dirac function of the human head, N is the total number of people included in the image, and $\bar{d}^{\,l}=\frac{1}{m}\sum_{j=1}^{m}d^{i}{}_{j}$ represents the average distance of the m heads closest to the head. In denser cases it is approximately equal to the head size. β is a hyperparameter, here it takes 0.3.

The use of this density map makes counting network regression easier because it no longer needs to get accurate head-annotated points.

We chose the C3 framework[9,10] as the basic network to train and test the method proposed in this article.

In this paper, we uses learning rate adaptive optimization algorithm Adam to optimize the network training. At the same time, the Euclidean distance between the network estimated density map and the true value is used as the loss of the training network regression.

$$L(\theta)=\frac{1}{2N}\sum_{i=1}^{N}\left\|\hat{y}(x_i;\theta)-y_i\right\|_2^2 \quad (2)$$

Where θ represents the parameters to be optimized by the network, N is the number of image in the training set, x_i represents the input picture, $\hat{y}(x_i;\theta)$ represents the crowd density map estimated by the network, and y_i represents the the ground truth to the input image.

The loss function can capture counting errors between the estimated and true values. The network counting error is reduced by minimizing this loss function.

3.3. *Evaluation Metric*

We use the mean absolute error (MAE) and the mean square error (MSE) to evaluate our method. MAE is defined as:

$$MAE=\frac{1}{N}\sum_{i=1}^{N}\left|C_{X_i}-C_{X_i}^{GT}\right| \quad (3)$$

Where N is the total number of pictures in the test set, $C_{X_i}^{GT}$ represents the true number of people corresponding to the input picture Xi, and C_{X_i} represents the number of people estimated by the network.

MAE is representative of model accuracy. In addition, to calculate the estimated variance, MSE is calculated as follows:

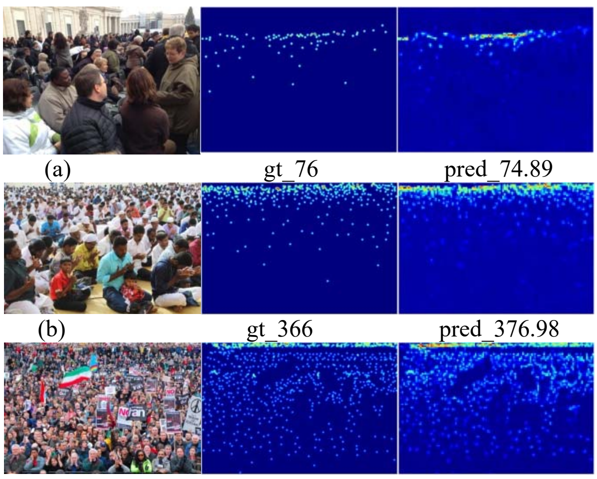
$$MSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(C_{X_i} - C_{X_i}^{GT})^2} \tag{4}$$

MSE is often used to indicate the robustness of counting predictions.

3.4. *Experimental Results*

Table 1.The experimental results on the UCF-QNRF dataset.

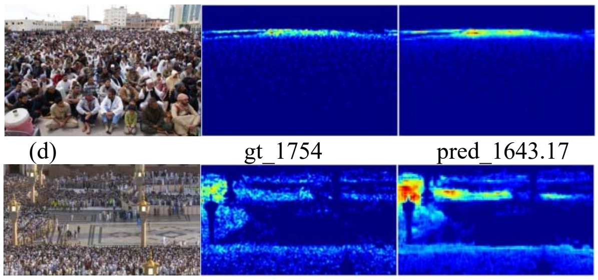
Method	MAE	MSE
Idrees 2013[11]	315	508
MCNN[2]	277	426
Encoder-Decoder[8, 12]	270	478
CMTL[13]	252	514
Switching CNN[14]	228	445
Resnet101[8, 15]	190	277
Densenet201[8, 16]	163	226
CL[8]	132	**191**
Our Proposed	**111.7**	198.2



(a) gt_76 pred_74.89

(b) gt_366 pred_376.98

(c) gt_797 pred_723.30



(e) gt_3302 pred_3436.02

Fig. 3. Visualization of density map.

The results of the self-attention distillation-based crowd counting network on the UCF-QNRF dataset are shown in Table 1.

As can be seen from the table, our proposed method has obtained relatively small errors on the UCF-QNRF dataset. The MAE is 111.7 and the MSE is 198.2. Compared with other more advanced crowd counting networks in recent years, MAE has improved significantly. Although MSE is not the lowest, it is similar to the MSE obtained by the method proposed in [8], and it is better than the other networks above. The experimental results demonstrate the effectiveness of self-attention distillation-based population counting networks.

Fig3. demonstrates the qualitative results of our method on different crowd congested scenes. From left to right, there are input images, ground truth crowd densities and the results of our method.

It can be seen from Fig.3 that the density map estimated by the network can well reflect the distribution of the crowd, but compared with the true density map, there are still some obvious differences. In this regard, further research is needed.

4. Conclusion

In this paper, we propose a crowd counting network with self-attention distillation, which is improved based on the CSRNet network. The network front end uses VGG-16 to perform basic feature extraction on the input image. In the dilated convolution part of the back end of the network, we selected three different locations to join the attention generator (AT-GEN). Use self-attention distillation strategy to obtain deeper global context information to guide shallow learning and obtain higher quality crowd density maps. Experiments on the UCF-QNRF dataset show that the method has superior performance and higher robustness than other advanced technologies.

5. References

1. Zhang, Y., Zhou, D., Chen, S., Gao, S., & Ma, Y.. Single-image crowd counting via multi-column convolutional neural network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2016), pp. 589-597.
2. Onoro-Rubio, D., & López-Sastre, R. J.. Towards perspective-free object counting with deep learning. In *European Conference on Computer Vision* (2016, October), pp. 615-629 . Springer, Cham.
3. Wang S, Zhao H, Wang W, et al. Improving Deep Crowd Density Estimation via Pre-classification of Density， In *Proceedings of the IEEE International Conference on Neural Information Processing*. (Springer, Cham, 2017), 260-269.
4. Li, Y., Zhang, X., & Chen, D.. Csrnet: Dilated convolutional neural networks for understanding the highly congested scenes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2018), pp. 1091-1100.
5. Wang L., Zhao H., Li Y.. Research on the multi-scale network crowd density estimation algorithm based on attention mechanism. *The 2019 International Conference on Intelligent Informatics and BioMedical Science (ICIIBMS2019)*, (Nov.21-24, 2019, Shanghai, China), pp.272-278.
6. Hinton G, Vinyals O, Dean J. Distilling the knowledge in a neural network[J]. *arXiv preprint arXiv*:1503.02531, (2015).
7. Hou, Y., Ma, Z., Liu, C., & Loy, C. C.. Learning lightweight lane detection cnns by self attention distillation. In *Proceedings of the IEEE International Conference on Computer Vision* (2019), pp. 1013-1021.
8. Idrees, H., Tayyab, M., Athrey, K., Zhang, D., Al-Maadeed, S., Rajpoot, N., & Shah, M.. Composition loss for counting, density map estimation and localization in dense crowds. In *Proceedings of the European Conference on Computer Vision (ECCV)*,(2018), pp. 532-546.
9. Gao, J., Lin, W., Zhao, B., Wang, D., Gao, C., & Wen, J. . C^ 3 Framework: An Open-source PyTorch Code for Crowd Counting. *arXiv preprint arXiv*:1907.02724.(2019)
10. Wang, Q., Gao, J., Lin, W., & Yuan, Y. (2019). Learning from synthetic data for crowd counting in the wild. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2019), pp. 8198-8207.
11. Idrees, H., Saleemi, I., Seibert, C., & Shah, M.. Multi-source multi-scale counting in extremely dense crowd images. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2013), pp. 2547-2554.
12. Badrinarayanan, V., Kendall, A., Cipolla, R.: Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *arXiv preprint arXiv*:1511.00561 (2015)
13. Sindagi, V. A., & Patel, V. M.. Cnn-based cascaded multi-task learning of high-level prior and density estimation for crowd counting. In*2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*(2017, August), pp. 1-6.
14. Sam, D. B., Surya, S., & Babu, R. V. . Switching convolutional neural network for crowd counting. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* , (2017, July), pp. 4031-4039.
15. He, K., Zhang, X., Ren, S., Sun, J.Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2016), pp. 770–778 .
16. Huang, G., Liu, Z., Weinberger, K.Q., van der Maaten, L.: Densely connected convolutional networks. *arXiv preprint arXiv*:1608.06993, (2016).

Path Planning Based on Improved Artificial Potential Field Method

Feifan Xu, Huailin Zhao, Zhen Nie, Xin Zhou, Zheheng Tao
School of Electrical and Electronic Engineering, Shanghai Institute of technology, No. 100, Haiquan Road, Fengxian District, Shanghai 201418, China
E-mail: zhao_huailin@yahoo.com,
www.sit.edu.cn

Abstract:

In this paper, the traditional artificial potential field method is improved. Aiming at the problem that the traditional algorithm cannot pass through the obstacles close to each other and is prone to oscillation near the obstacles, the angle function is added to match the original force field function base on the traditional algorithm, and the stability is enhanced by combining the idea of fuzzy control. Finally, a reasonable and smooth optimal path is obtained by MATLAB simulation. It is proved that the multi-function parallel and multi-algorithm hybrid algorithm is feasible in the field of mobile robot path planning.

Keywords: Path Planning；Artificial Potential Field；Fuzzy Control

1. Introduction

With the advent of the intelligent era, robot path planning [1-3] is gradually becoming a hot topic. The path planning, or more concretely speaking, is the global path planning, which is a behavior of a developer giving the optimal solution of a path according to a given starting point and end point and through an algorithm when he is fully informed of the environmental conditions. At present, domestic and foreign scholars have done a lot of research on the path planning of mobile robots, and have put forward a lot of algorithms. Commonly used path planning algorithms are: A* algorithm [4], fuzzy algorithm [5], genetic algorithm [6], artificial potential field method [7], ant colony algorithm [8], etc.

The artificial potential field (APF) is a virtual force method proposed by Oussama Khatib, a professor at Stanford University's Intelligent Robot Laboratory. The core idea is to imagine the operating environment of the robot as a virtual force field. This force field is separated into repulsive force field and gravitational field [9]. The idea of split is very important in the improvement of APF, because whether it is the im provement of multi-function or multi-algorithm, it is fundamentally the antonym of split. It is a kind of superposition. It constantly increases the original two-force field effect to multi-force field effect.

Compared with other algorithms, the APF has some advantages in terms of simple model, strong real-time performance, less calculation amount and low requirement for hardware platforms. However, the APF has the problems of target unreachability and local minimum, which leads to the failure of path planning. In this paper, aiming at the problem of path planning for robots, the problem of the traditional APF is improved in two ways. The one is to enrich the combination of multiple functions of repulsion function by adding angle function. The problem of the algorithm can be alleviated in some cases, but a lot of deficiencies in stability can still be found. The other one considers the angle function as well as fuzzy control [10]. By this algorithm, the robot can still have the external force to pull it out even if it is in the local minimum position, which further strengthens the stability of the algorithm. Finally, through the MATLAB simulation analysis, the improved APF will be fully competent for the path planning problem of robots.

2. Traditional APF

The APF introduces a virtual field into the space. The target point shows gravitation to the robot. The gravitation is increasing with the distance of the robot from the target point, and the direction is that the robot points to the target point. The obstacles show repulsive forces to robots. The repulsive force decreases with the distance between robots and obstacles, and the direction is that the obstacles points to robots. The robots maintain the motion state by the gravitation and repulsive force.

2.1. Gravitational Function

For the two force fields in the traditional algorithm, the gravitational field is originated from the gravitation of the target point to the robot. The gravitational field is shown in Figure 1. The starting point is high on the right side, while the target point is low on the left side.

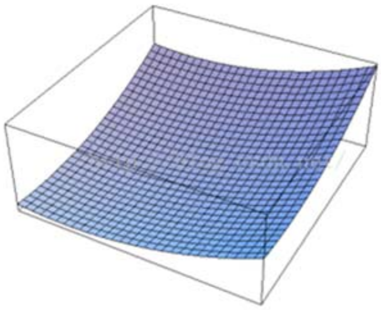


Fig.1. Gravitational Field Model

Such a model determines a fundamental principle, that is, the farther the robot is from the target point, the greater the gravitation, and as the target point slides down the hillside, the magnitude of gravitation decreases in the form of a power function. So the gravitation formula is expressed as follows.

$$U_{ttr}(q) = \frac{1}{2}\xi\rho^2\left(q, q_{goal}\right) \tag{1}$$

Where $\rho\left(q, q_{goal}\right)$ is the distance between robot and target point. It is calculated by $\rho\left(q, q_{goal}\right) = \left\|q - q_{goal}\right\|$. ξ is the positive proportional gain coefficient of the gravitational field.

According to the negative gradient of the gravitational field, the gravitation $F_{attr}(q)$ is as follows.

$$F_{attr}(q) = -\operatorname{grad}\left[U_{attr}(q)\right] \tag{2}$$

2.2. Repulsion Function

The obstacles have repulsive force on the robots. In order to prevent robots from approaching and colliding with obstacles, repulsion fields are formed. As shown in figure 2, the repulsive field is like a hill.

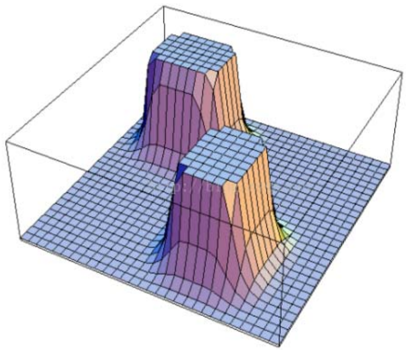


Fig.2. Repulsive Field Model

The farther the robot is from the obstacles, the smaller the repulsive force will be. Once the robot is so far away from the obstacles that it is beyond the influence of the repulsion field of the obstacle, the repulsive force will be zero. The formula for the repulsion function is as follows.

$$U_{rep}(q) = \begin{cases} \frac{1}{2}\eta\left(\frac{1}{\rho\left(q, q_{obs}\right)} - \frac{1}{\rho_0}\right)^2, & \rho\left(q, q_{obs}\right) \le \rho_0 \\ 0, & \rho\left(q, q_{obs}\right) \ge \rho_0 \end{cases} \tag{3}$$

Where η is the positive proportional gain coefficient of repulsion field. $\rho\left(q, q_{obs}\right)$ is the distance between the robot and the obstacles. It is calculated by $\rho\left(q, q_{0bs}\right) = \left\|q - q_{obs}\right\|$. ρ_0 is the influence range of the obstacles. The size of the value depends on the environment and the actual needs, and has no significant impact on the algorithm.

According to the negative gradient of the repulsive field, the repulsive force $F_{rep}(q)$ is as follows.

$$F_{rep}(q) = -\operatorname{grad}\left[U_{rep}(q)\right] \tag{4}$$

2.3. Resultant Function

The resultant force field, as its name implies, is a simulated force field formed by the gravitational field and repulsive force field. The resultant function is $U(q) = U_{att}(q) + U_{rep}(q)$. The resultant field model is shown in Figure 3.

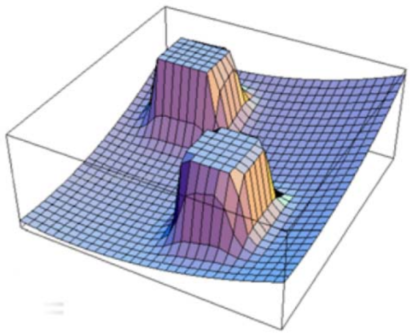


Fig.3. Resultant Field Model

Imagine that when a small ball is released at the top of the terrain, it will naturally reach its destination under resultant force. In essence, it is to control the resultant force to drive the robot to plan an optimal path.
When there are n obstacles around the robot, the resultant force F_{total} of the robot is as follows.

$$F_{\text{total}} = F_{attr} + \sum_{i=1}^{n} F_{repi} \tag{5}$$

The force model of the robot is shown in figure 4.

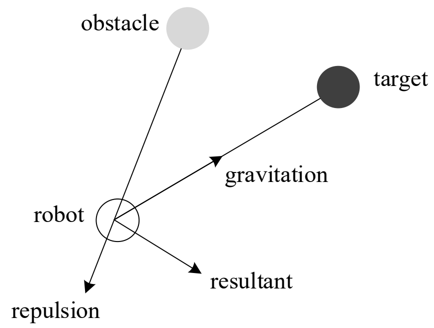


Fig.4. Force Model of Robot

2.4. Limitations of Traditional APF

When the path is planned by APF, if the position of the target point and the obstacles are random, special position may be generated, which may lead the path planning to be incorrect or impossible to complete. For these special cases, a total of four cases were summarized.
1) Place obstacles on both sides of path.
When obstacles are placed on both sides of the path that the robot will go through, repulsion sometimes counteracts gravitation and fail to path plan, as shown in figure 5.

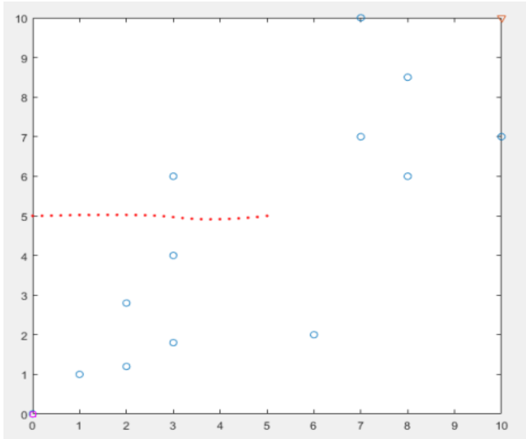


Fig.5. Obstacle on Both Sides of the Path

This problem indicates that the configuration of the coefficients may be uncoordinated, so long as the appropriate coefficients are adjusted, it can be solved.
2) Place obstacles in the middle of path
When obstacles are placed in the path that the robot is going through, the robot will be subjected to a strong repulsive force facing itself vertically. At a certain position, it will be offset by its own gravitation. Although it seldom fails to plan, it often appears local oscillation. As shown in Figure 6.

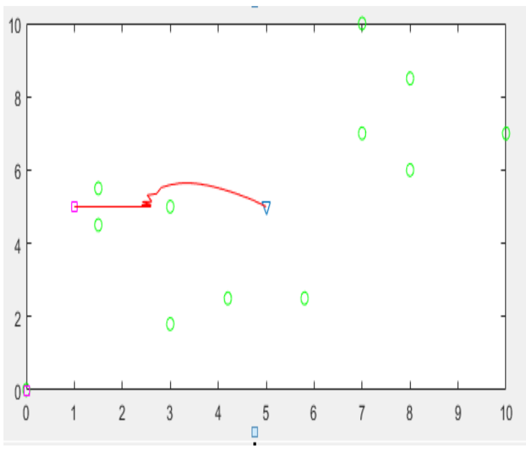


Fig.6. Obstacle in the Middle of the Path

3) Place obstacles near the target point
Because the repulsive force fields of the obstacles exist, if the target point is close to the obstacles, the target point is affected by the repulsion field. It will lead to the failure of the planning, which is as shown in Figure 7.

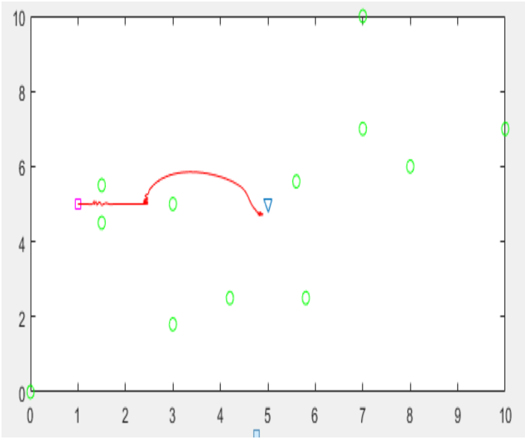
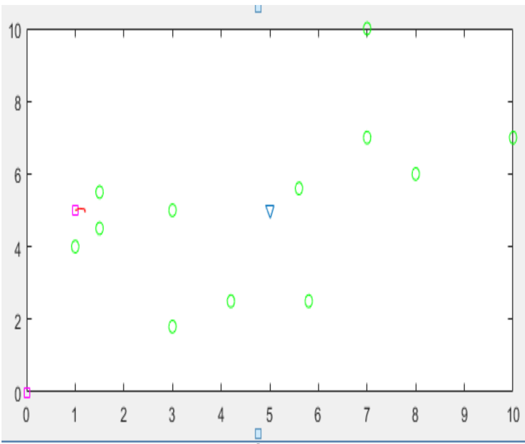


Fig.7. Obstacles Near the Target Point

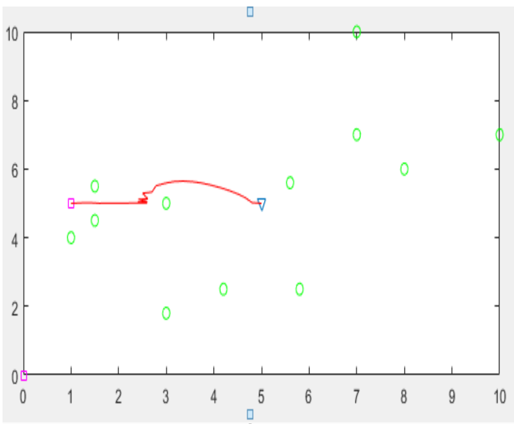
It can be seen from Figure 7 that the path plan fails and there is a violent oscillation in the path.

4) Place obstacles near the starting point

In the same way as in case 3, there are fewer cases where it is impossible to plan, because gravity is the most powerful here, and repulsion only interferes. This case is as shown in Figure 8.



(a)



(b)

Fig.8. Obstacle Near the Starting Point

The impact of the repulsion field appearing in Figure 8(a) leads to the failure of the path planning. Although the planning is completed in Figure 8(b), there is some local oscillations.

3. Improvement of APF

We have mentioned the word "superposition" constantly before. The basic principle of APF determines that superposition is the basic route of improvement. The traditional repulsive field and gravitational field can be combined to make path planning. This kind of superposition on the model is constantly stacking new force field. In actual application, it can be achieved by mathematical summation of the component force. That is to say, the original formula 6 is improved to formula 7.

$$\begin{cases} F_{xattr} + F_{xrep} = F_x \\ F_{yattr} + F_{yrep} = F_y \end{cases} \tag{6}$$

$$\begin{cases} F_{xattr} + F_{xrep} + F_{x1} + F_{x2} + \cdots = F_x \\ F_{yattr} + F_{yrep} + F_{y1} + F_{y2} + \cdots = F_y \end{cases} \tag{7}$$

Where F_x is the resultant force in the x-axis direction. F_{xattr} is the gravitation in the x-axis direction. F_{xrep} is the repulsive force in the x-axis direction. F_{xi}, $i = 1, 2, \cdots$ is the new force in the x-axis direction. F_y is the resultant force in the y-axis direction. F_{yattr} is the gravitation in the y-axis direction. F_{yrep} is the repulsive force in the y-axis direction. F_{yi}, $i = 1, 2, \cdots$ is the new force in the y-axis direction.

3.1. APF Considering Angle

For the traditional algorithm, the angle function is introduced as a consideration. The angle function is to express the angle between the line formed by the robot and the obstacles and the x-axis by the inverse trigonometric function through the difference between the position of the obstacles and the robot.

The gravitation function basically remains unchanged. The component of the x-axis and y-axis is still expressed by multiplying the coefficients by the distance between the axis and the target. The distance between the axis and the target is expressed by the sine or cosine of the distance for the sake of stability, which is essentially unchanged.

The repulsion function has changed a lot. The angle function is introduced mainly to use two sets of repulsion functions at the same time. The second function added on the basis of the original repulsion function is as follows.

$$U_{rep}(q)=\begin{cases}\frac{1}{2}\eta\left(\frac{1}{\rho(q,q_{obs})}-\frac{1}{\rho_0}\right)^2\times\rho\left(q,q_{goal}\right), & \rho(q,q_{obs})\le\rho_0\\ 0, & \rho(q,q_{obs})\ge\rho_0\end{cases} \quad (8)$$

Where η is the positive proportional gain coefficient of repulsion field. $\rho(q,q_{obs})$ is the distance between the robot and the obstacles. It is calculated by $\rho(q,q_{0bs})=\|q-q_{obs}\|$. ρ_0 is the influence range of the obstacles. $\rho\left(q,q_{goal}\right)$ is the distance between robot and target point.

The purpose of this improvement is to increase the repulsion force when the distance between the target point and the robot is increased, thus counteracting the strong gravitation here. Similarly, when the target point is very close to the robot, the gravitational force is very small, but as the distance decreases, the dragging effect on the repulsive force field can also be formed. In the model, it can be imagined that the hills will increase their steepness as the elevation of the hillside increases.

3.2. APF Combined with Fuzzy Control

The previous algorithm was built on the model by establishing a force field model and then using the free roll of the ball. The fuzzy logic method actually adds a driving idea, the fuzzy logic method adds a driving idea, which is like adding a driver to the ball to operate the steering wheel through the change of force field. Fuzzy control has strong robustness, and its idea fully combines physiological perception and action integration, and it has the advantage of lower computational complexity than the same type of algorithm.

The two inputs of the fuzzy control are the angle difference of gravitation and repulsion, and the resultant force difference between them. Both of them use the triangular membership functions. In practice, the shape of the membership function does not have a great influence on the control effect. On the contrary, the coverage of the universe by the fuzzy subset affects the performance of the whole control. However, theoretically, the shape of triangular membership function is only determined by the slope of the line, which is relatively simple in operation and less memory consumption. Therefore, it is selected as the membership function shape here. The fuzzy rules are shown in Table 1.

Table 1. Fuzzy Rule Table

Angle	Error	Output	Angle	Error	Output
-2	-2	-2	0	1	1
-2	-1	-1	0	2	1
-2	0	-1	1	-2	-1
-2	1	0	1	-1	0
-2	2	1	1	0	0
-1	-2	-2	1	1	1
-1	-1	-1	1	2	2
-1	0	0	2	-2	-1
-1	1	0	2	-1	0
-1	2	1	2	0	1
0	-2	-1	2	1	1
0	-1	-1	2	2	2
0	0	0			

The output obtained by the fuzzy rule is magnified 100 times and multiplied by the sine and cosine of the angle between the resultant force obtained previously and the x-axis. A new set of x-axis and y-axis components are obtained and added to the calculation of the total resultant force.

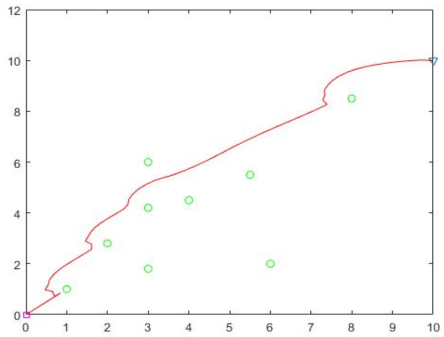
Similarly, the new resultant force only adds an additional set of components. The formula is as follows.

$$\begin{cases}F_{xattr}+F_{xrep1}+F_{xrep2}+F_{xf}=F_x\\ F_{yattr}+F_{yrep1}+F_{yrep2}+F_{yf}=F_y\end{cases} \quad (9)$$

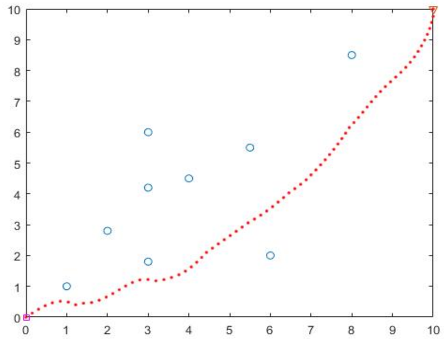
Where F_{xf} is the fuzzy control force in x-axis. F_{yf} is the fuzzy control force in Y-axis direction. When the repulsive force is zero, the repulsive force angle is forced to be equal to the gravitational angle, which can prevent some runaway situations.

4. Simulation Analysis

According to the path planning algorithm designed in the previous section, the simulation analysis is carried out on the MATLAB. Firstly, the environment model is built on MATLAB. The model is a 10m*10m simulation environment, including obstacles, starting point and end point. Firstly, the APF considering the angle is simulated and analyzed. The simulation results are shown in Figure 9.



(a)

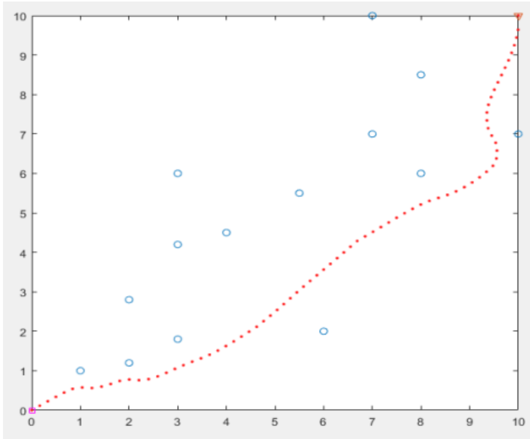


(b)

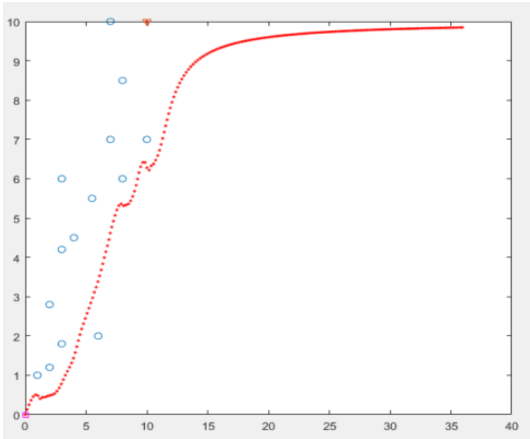
Fig.9. Path Planning Results：(a) The Path Obtained by Traditional APF; (b) The Path Obtained by APF Considering Angle

Fig. 9(a) shows the path obtained by using the traditional APF, but there is a problem of local oscillation. Fig. 9 (b) is the path obtained by the APF considering the angle. The influence distance of the obstacle is set to 5. The algorithm finds a better planning path than the original one, and the oscillation problem has been solved well. It can be seen from the simulation results that the traditional algorithm not only has local minimum point problems such as oscillation, but also cannot be optimal in path selection. The path length is 15.44024 m. The improved artificial potential field method by introducing the angle function and considering the distance between the robot and the target point not only solves the problem of local oscillation, but also plans a more optimal and smoother path. The path length is 15.04447m, which optimizes the path length by 5%.

In order to fully confirm the effect of the fuzzy control on the APF, the number of obstacles is increased that distributed near the target point and the starting point. The idea of this setting further strengthens the test of the stability performance of the algorithm. According to experimental experience, the repulsion field of obstacles near the target point and the starting point has a great influence on the path. The simulation results are shown in Figure 10.



(a)



(b)

Fig.10. Path Planning Results: (a) Path Obtained by APF Combined with Fuzzy Control; (b) The Path Obtained by APF Considering Angle.

Fig. 10 (a) is a path planned by using the APF combined with fuzzy control. The influence range of obstacles is set to 7. An optimal path is obtained by the algorithm between dense obstacles. Fig. 10(b) is a more complex map environment set for Fig. 9(b). The influence range of obstacles is set to 5, and the number of obstacles is increased. From the simulation results, although it has been greatly improved compared with the traditional APF, in the complex and harsh map environment, its lack of intelligence is maximized. If there is a more complex environment with a lot of obstacles, the algorithm would run-away phenomenon, and path planning fails.

5. Conclusion

In this paper, aiming at the path planning of robot in complex environment, the angle function and the fuzzy control are introduced to improve the traditional APF. The rapid and accurate path planning is realized. The simulation analysis on MATLAB fully demonstrates the superiority of the improved algorithm. Through this algorithm, even if the robot is at the local minimum position, there is still an external force pulling it out, which strengthens the stability of the algorithm. The improved APF can effectively complete the path planning of robots.

References

1. Huailin Zhao, Zheng Wu, Xiaoxing Wang. A Mul-ti-robot Rescuing System. The 2018 International Conference on Artificial Life and Robotics (ICAROB 2018), 2018, Japan, pp. 582-586;
2. Huailin Zhao, Zhen Nie. Design and Analysis of Multi-robot grouping aggregation algorithm. Journal of Robotics, Networking and Artificial Life, 2019, Vol. 6, pp. 60-65;
3. Bingbing Xu, Rongfei Hao. Current Situation and Development of Robot Path Planning Technology. Electronic Technology & Software Engineering, 2019, pp. 81;
4. Dianjun Wang. Indoor mobile-robot path planning based on an improved A* algorithm. Journal of Tsinghua University (Science and Technology), 2012, Vol.52, pp. 1085-1089;
5. Weiwei Shao, Yuxiu Wu, Dongdong Fan. Overview of Mobile Robot Path Planning Technology. Journal of Tongling Vocational & Technical College, 2018, Vol.17, pp. 57-60;
6. Mohd. Nayab Zafar, J. C. Mohanta. Methodology for Path Planning and Optimization of Mobile Robots: A Review. International Conference on Robotics and Smart Manufacturing (RoSMa2018), 2018, pp.141-152
7. Xianxia Liang, Zhaoying Liu, Xueling Song, Yingkun Zhang. Research on Improved Artificial Potential Field Approach in Local Path Planning for Mobile Robot. Computer Simulation, 2018, Vol.35, pp. 291-294;
8. Jiang Ming, Wang Fei, Ge Yuan, Sun Longlong. Research on path planning of mobile robot based on improved ant colony algorithm. Chinese Journal of Scientific Instrument. 2019, Vol. 40, pp. 113-121;
9. Qiang Luo, Haibao Wang, xiaojin Cui, Jingchang He. Autonomous Mobile Robot Path Planning Based on Improved ArtificialPotential Method. Control Engineering of China, 2019, Vol.26, pp. 1091-1098;
10. Weidong Chen, Baoxia Li, Qiguan Zhu. Application of fuzzy control in mobile robot path planning. Computer Engineering and Applications, 2009, Vol.15, pp. 221-223

Self-balancing Car based on Adaptive Fuzzy PID Control

Zhen Nie, Huailin Zhao, Lu Sun, Xiongfeng Zhong
School of electrical and Electronic Engineering, Shanghai Institute of technology, No. 100, Haiquan, Road, Fengxian District, Shanghai 201418, China
E-mail: zhao_huailin@yahoo.com, hbniezhen@foxmail.com
www.sit.edu.cn

Abstract:

The self-balancing car is widely studied for its advantages of convenient operation, flexible movement, energy saving and reliability. In this paper, the STM32F105 is used as the main control chip to design a self-balancing car which realizes remote control and video transmission through WIFI. Because the self-balancing car is an unstable nonlinear system, the Kalman filter optimization algorithm is used to fuse the data of gyroscope and accelerometer, and the adaptive fuzzy PID control algorithm is used to control the balance of whole system. The experimental result shows that the system is easy to control, has strong stability, low power consumption and high transmission efficiency.

Keywords: Self-balancing Car；Kalman Filter Optimization Algorithm; Fuzzy PID Control

1. Introduction

With the development of automatic control technology, more and more researchers are devoted to the research of two-wheeled mobile robots [1-2]. Compared with traditional wheeled mobile robots, the two-wheeled self-balancing car is located on the top of the axle. It maintains balance through movement. The two wheels are coaxial and driven independently. Moreover, it has the following advantages: The motion trajectory is flexible and changeable, which adapts to the complex working environment. The structure of car is simple, which realizes miniaturization and lightweight. It occupies a small area and is suitable for a variety of working environments. However, the two-wheeled self-balancing car has the complexities of multivariable, strong coupling, non-linearity, and highly coupled motion control. It is an ideal platform for the research of two-wheeled mobile robots and various control algorithms. The traditional two-wheeled self-balancing car uses first-order filtering and second-order filtering, and a higher level uses the Kalman filtering algorithm [3-4] and PID [4-5]. Although it can control the balance of the car, it has the disadvantages of slow response and poor stability.

In this paper, we design a self-balancing car which can use Android client to control the car and display the pictures collected by the car. It makes the self-balancing car more widely used. The system design not only upgrades the hardware, which reduces the load and power consumption, but also optimizes the algorithm. In this paper, the STM32F105 [5-6] is used as the main control chip of the system, the MPU6050 [7-8] is used to collect the angle and speed of the car, and Kalman filter optimization algorithm is used to integrate the data of gyroscope and accelerometer. Finally, self-balance control of the whole system is carried out by an adaptive fuzzy PID control algorithm [9-10]. After the self-balancing car is powered on, it can reach equilibrium and maintain in a short time, which improves the overall performance of the system.

2. Algorithm Design

The gyroscope is accurate but has zero-drift, and the measurement errors will accumulate continuously with the accumulation of time, which will affect the measurement accuracy. The accelerometer is affected by the vibration of the balance car, and the additional high-frequency vibration interference is aliased, but the drift is small. Therefore, a signal fusion method is needed to obtain an accurate inclination angle. Kalman filtering, which can set the assign trust weights, solve the zero drift of gyroscope and filter out the high-frequency vibration of the accelerometer. It also can recursive self-learning, smooth filtering and fast follow. Therefore, the Kalman filter optimization algorithm is used to fuse the gyroscope and accelerometer data, and the inclination angle and angular velocity are used as input variables to perform self-balancing control on the overall system through the adaptive fuzzy PID control algorithm. Figure 1 shows the flowchart of the whole program design.

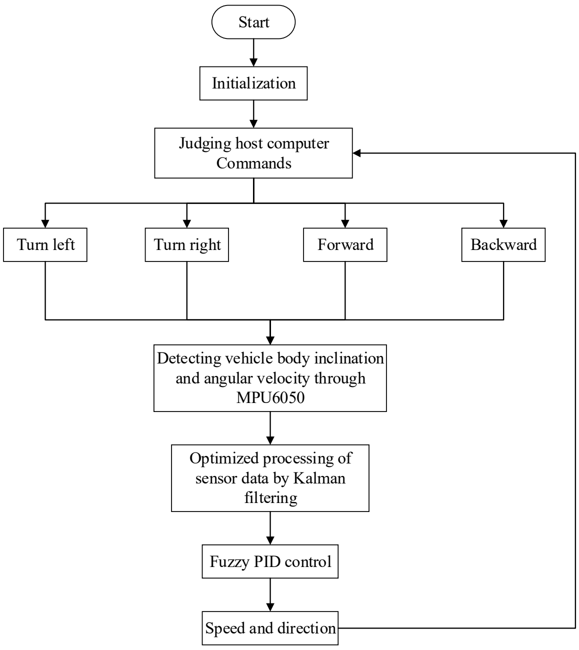


Fig.1 Control Program Flowchart

2.1 Kalman Filtering Algorithm

As mentioned earlier, in this paper, Kalman filter algorithm is used to fuse the data of the gyroscope and the accelerometer, and the optimal inclination angle and angular velocity are used as input for fuzzy control. In this way, the zero drift of the gyroscope and the dynamic error of the accelerometer will be overcome. The specific process of Kalman filtering is as follows:

Firstly, the deviation t between the tilt angle θ of the car and the angular velocity measured by the gyroscope observed by the accelerometer is taken as the state vector of the system, and the following state and measurement equation can be obtained:

$$\begin{bmatrix}\dot{\theta}\\ \dot{t}\end{bmatrix}=\begin{bmatrix}0 & -1\\ 0 & 0\end{bmatrix}\begin{bmatrix}\theta\\ t\end{bmatrix}+\begin{bmatrix}1\\ 0\end{bmatrix}u_k+\begin{bmatrix}\omega_{gyro}\\ 0\end{bmatrix} \quad (1)$$

$$\theta_{\text{acce}}=\begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}\theta\\ t\end{bmatrix}+\omega_{\text{acce}} \quad (2)$$

Where u_k is the output data of gyroscope. ω_{gyro} is the noise signal from gyroscope. ω_{acce} is the noise signal from accelerometer. θ_{acce} is the angle of the accelerometer after arcsine processing.

According to previous experience, we can know that ω_{gyro} and ω_{acce} obey Gaussian distribution. The above formula can be discretized to ontain:

$$X(k)=\begin{bmatrix}1 & -T\\ 0 & 1\end{bmatrix}\times X(k-1)+\begin{bmatrix}T\\ 0\end{bmatrix}\times U(k-1)+\begin{bmatrix}\omega_{gyro}(k)T\\ 0\end{bmatrix} \quad (3)$$

$$Z(k)=\begin{bmatrix}1 & 0\end{bmatrix}X(k)+\omega_{\text{acce}}(k) \quad (4)$$

Where T is the discrete period. $X(k)$ is the status of the balancing car at time k. $X(k-1)$ is the status of the balancing car at time k-1. $U(k-1)$ is the angular velocity measured by the gyroscope at time k-1. $\omega_{gyro}(k)$ is the noise from gyroscope at time k. $Z(k)$ is the angle obtained after the arcsine processing of the accelerometer at time k. $\omega_{acce}(k)$ is the noise from accelerometer at time k.

Let $A=\begin{bmatrix}1 & -T\\ 0 & 1\end{bmatrix}$ be the system prediction matrix, $B=\begin{bmatrix}T\\ 0\end{bmatrix}$ be the control matrix, and $H=\begin{bmatrix}1 & 0\end{bmatrix}$ be the measurement matrix.

In order to correct the Kalman filter, it is necessary to solve the covariance matrix of the system:

$$P(k\,|\,k-1)=A\times P(k-1\,|\,k-1)\times A^T+Q \quad (5)$$

Where $P(k\,|\,k-1)$ is the the covariance corresponding to $X(k\,|\,k-1)$. Q is the process noise covariance of the system, expressed as:

$$Q = \begin{bmatrix} q_{acce} & 0 \\ 0 & q_{gyro} \end{bmatrix} R \qquad (6)$$

Where q_{acce} and q_{gyro} are the covariances of accelerometer and gyroscope respectively. R is the covariance of measurement error.

The state equation modified by the sensor data is

$$X(k|k) = X(k|k-1) + K_h(k) \times [Z(k) - H \times X(k|k-1)] \qquad (7)$$

Where K_h is the Kalman gain, expressed as

$$K_h(k) = \left[P(k|k-1) \times H^T\right]\left[H \times P(k|k-1) \times H^T + R\right] \qquad (8)$$

Furthermore, the covariance at time k can be expressed as

$$P(k \mid k) = \left(I - K_h(k) \times H\right) \times P(k \mid k-1) \qquad (9)$$

Where I is the identity matrix.

Thus $X(k|k)$ and $P(k|k)$ are the latest optimal estimates.

2.2 Fuzzy Adaptive PID Algorithm

The fuzzy adaptive PID algorithm is an intelligent algorithm for online tuning of fuzzy rules based on empirical PID parameters. It takes the error e and the error change rate e_c as the input of the PID controller, and the parameters of the PID controller are adjusted online using a built-in expert rule base, so that the PID parameters in a better working state all the time.

As mentioned earlier, the inclination angle obtained by Kalman filter is used as the input of error e, and the angular velocity is used as the input of error change rate e_c. After the inputs are determined, the structure of fuzzy adaptive PID control is designed as shown in Figure 2.

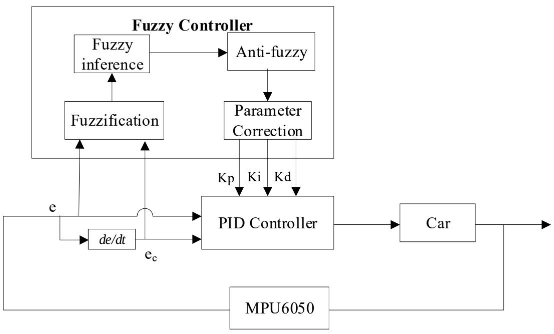
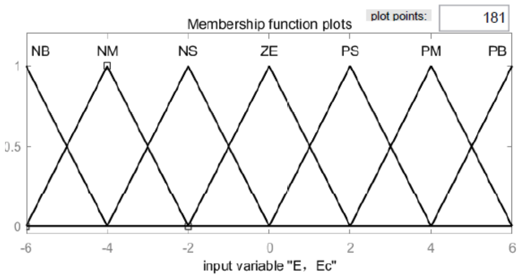


Fig. 2 Structure of Fuzzy Adaptive PID Control

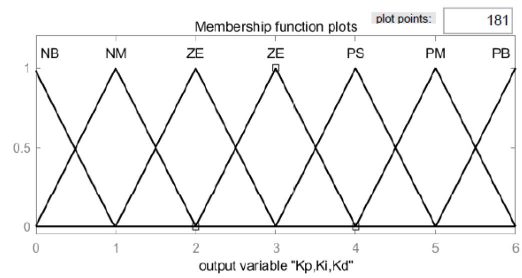
After the membership functions are established, according to the real-time step response of the system, the self-tuning rule of PID parameters is formulated as shown in Table 1.

In this paper, the barycenter method is used to anti-fuzzy, and the formula is as follows:

Set the basic fuzzy set to divide the input and output space. In this paper, the triangular membership function is used for input and output, and set the fuzzy subsets of e and e to {Negative Big (NB), Negative Medium (NM), Negative Small (NS), Zero (Z), Positive Small (SM), Positive Medium (PM), Positive Big (PB)} , and set the corresponding universe range as [- 6,6]. The fuzzy subsets of Kp, Ki and Kd are set as {NB, NM, NS, Z, PS, PM, PB}, and the universe range is set as [0,6]. as shown in Figure 3, these are the membership functions of the input and output.



(a) Membership Function of Input



(b) Membership Function of Output

Fig. 3 Membership Function

$$\mu = \frac{\sum_{i=0}^{n} \mu_c(z_i) \cdot z_i}{\sum_{i=0}^{n} \mu_c(z_i)} \qquad (10)$$

Where μ is the exact value of output after anti-fuzzy through the fuzzy controller. z_i is the value in the fuzzy control universe range. $\mu_c(z_i)$ is the membership value of z_i.

Table 1. Fuzzy Rule of Kp/Ki/Kd

ec \ e	NB	NM	NS	Z	PS	PM	PB
NB	PB/NB/PS	PB/NB/NS	PM/NM/NB	PM/NM/NB	PS/NS/NB	Z/Z/NM	Z/Z/PS
NM	PB/NB/Z	PB/NB/NS	PM/NM/NB	PS/NS/NM	PS/NS/NM	Z/Z/NS	NS/Z/Z
NS	PM/NB/Z	PM/NM/NS	PM/NS/NM	PS/NS/NM	Z/Z/NS	NS/PS/NS	NS/PS/Z
Z	PM/NM/Z	PM/NM/NS	PS/NS/NS	Z/Z/NS	NS/PS/NS	NM/PM/NS	NM/PM/Z
PS	PS/NM/PS	PS/NS/Z	Z/Z/Z	NS/PS/Z	NS/PS.Z	NM/PM/Z	NM/PB/Z
PM	PS/Z/PB	Z/Z/NS	PS/PS/PS	NM/PS/PS	NM/PM/PS	NM/PB/PS	NB/PB/PB
PB	Z/Z/PB	Z/Z/PM	NM/PS/PM	NM/PM/PM	NM/PM/PS	NB/PB/PS	NB/PB/PB

3. Hardware Design

In this design, an attitude sensor (MPU6050) is used to monitor the pitching state and state change rate of the car. The high-speed microcontroller STM32f105 uses Kalman filter algorithm and fuzzy PID control algorithm to complete the analysis and calculation. The output control signal drives the motor to generate forward or backward acceleration to control the car to maintain balance. The video information collected by the camera is directly transmitted to the host computer through WIFI module to realize the real-time monitoring function. The host computer sends commands to control the forward, backward and turning of the balance car. The system hardware structure block diagram is shown in Figure 4.

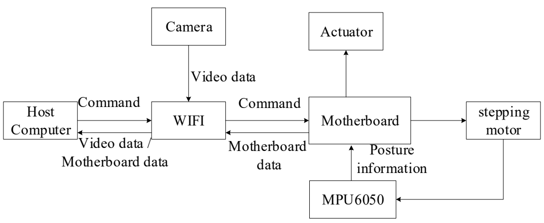


Fig. 4 System Hardware Block Diagram

The system has three core components: host computer, WIFI module and control motherboard. The host computer sends the control command to the WIFI module, and the WIFI module transmits the command signal to the motherboard through the serial port. The motherboard controls the balance car movement according to the control command. And the motherboard transmits the parameters to the host computer through the WIFI module for display. Then the command exchange and data exchange are completed.

3.1 WIFI / Camera Module

This module is a main component to realize video transmission and control command distribution. The camera is directly connected to the WIFI module, and a separate video stream is sent to the host computer for video display through the WIFI module. It has no direct relationship with the motherboard. Between the camera and the host computer, the image acquisition, transmission and display are directly completed by WIFI module. Here, we used an Android phone as the host computer to realize the remote control of the car and video display. The video transmission block diagram is shown in figure 5.

Fig. 5 Hardware Connection Diagram of Video Transmission

3.2 Camera Control

This part is designed with reference to the structure of the universal joint, using two steering engines to achieve the Omni-directional camera shooting. The platform structure is shown in Figure 6.

Fig. 6 Steering Engines Platform

The steering engine uses PWM to control the angle. The high level of PWM is 0.5ms-2.5ms, which corresponds to 0 ° -180 ° of the steering engine. In order to prevent the steering engine from working in the dead zone for a long time and causing damage to the steering engine, we control it to rotate between 20 ° -160 °.

3.3 Design of Host Computer

The host computer uses a common Android phone as the platform and is written in Java. The control and display interface is shown in Figure 7.

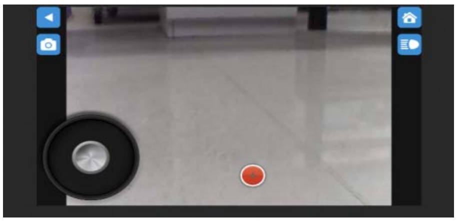


Fig. 7 Mobile Phone Client

The client and the car communicate with each other through the Wi-Fi module, which is used to receive the video stream transmitted by the WIFI module. And then the client performs decoding display to realize the video transmission. At the same time, the movement of the balancing car is controlled by the software buttons. The client and the car follow the established communication protocol, different buttons correspond to different control instructions. The steering wheel controls the movement direction of the balance car, sliding on the screen can control the camera rotation, and the photo button can take photos.

4. Experimental Verification

4.1 Simulation Test

In order to study the superiority of the balancing car system designed in this paper, we used SIMULINK environment in MATLAB to carry out a simulation comparison of traditional PID and fuzzy adaptive PID control algorithms, and obtained the simulation results shown in Figure 8.

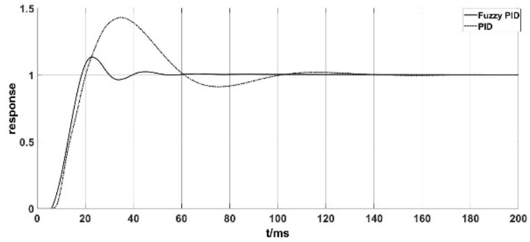


Fig. 8 Simulation Result

In Figure 8, we set a simulation time of 200ms, in which the solid line is the waveform obtained by fuzzy adaptive PID control, and the dashed line is the waveform obtained by traditional PID control. From the simulation results, it takes 70ms for the fuzzy adaptive PID control system to reach balance and 170ms for the traditional PID control system. From the perspective of overshoot, the fuzzy adaptive PID control has an overshoot of 0.1 and a traditional PID of 0.4. Obviously, the fuzzy adaptive PID control is better than the traditional PID control. It not only has a fast response speed, but also has small overshoot, small vibration, and good stability.

4.2 Experiment

After the simulation analysis, the designed algorithm and system have good stability and rapidity. According to the previous design, a self-balancing car that can realize video transmission is made, as shown in the figure 9.

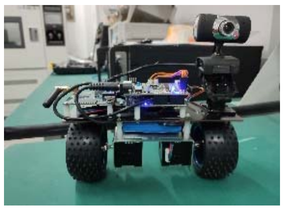


Fig. 9 Self-Balancing Car

In the experiment of the balance car, the balance, remote control and video display of the car are mainly tested. For the balance test of the car, we built a slope of about 20 ° in the laboratory, then let the car move on it to test its balance. The experiment is as shown in Figure 10. And the client control and video display is as shown in figure 7 in Section *3.3*.

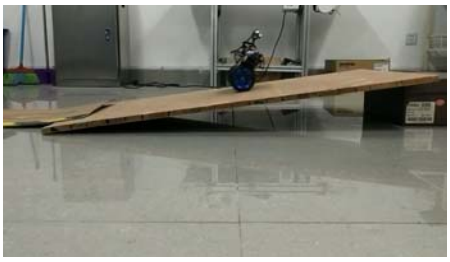


Fig. 10 Balance Test

Before the experiment, the balance car was placed on the flat ground. After the car was started, the client controlled the car to run on the flat ground. Then we use hands to interfere the car movement, the car can still maintain a good balance state, which proves that the car has a good stability. Thereafter, the car was controlled to move on the slope, and the car can keep running smoothly, which proved that the car had a good balance. For the test of the client, it can control the movement of the car very well, and the displayed screen is clear, which confirms the accuracy of the remote control of the client and the real-time performance of the video display.

5. Conclusion

This paper takes the self-balancing car as the research object, and takes the balance control and video transmission as the research purpose. The balance control and video transmission of the self-balancing car in a simple environment were realized. In this system, Kalman filter algorithm is used for data fusion, and fuzzy self-adaptive PID is used to control the balance of the car, so that the balance speed and stability of the car are improved. In this paper, not only simulation analysis in MATLAB was carried out, which confirms the feasibility of the algorithm, but also an experiment in the actual environment was carried out, and achieved the expected results. The experiment confirmed the stability and reliability of the system. It can be seen that the self-balancing car system has certain application value.

References

1. Huailin Zhao, Yangguang Guo, Masanori Sugisaka. Design of the Multi-Car Collaboration System[C], *the 2017 International Conference on Artificial Life and Robotics* (ICAROB 2017), 2017.01, Japan. pp. 186-190.
2. Huailin Zhao, Zhen Nie. Design and Analysis of Multi-robot grouping aggregation algorithm. *Journal of Robotics, Networking and Artificial Life,* 2019, Vol. 6: 60-65;
3. Li Zeyu. Design of Control System for Balanced Vehicle based on Kalman Filte and Double Closed Loops PID Control[J]. *Electronics World,* 2018(01):100-102.
4. Chen Pengzhan, Zhu Nianhua, Li Jie. Attitude Detection and Balance Control of Two-wheeled Self-balancing Vehicles[J]. *Control Engineering of China*, 2017,24(08):1574-1578.
5. Lin Feng, Cai Yanguang. Double Closed Loop PID Control System for Two-Wheel Self-Balancing Vehicle[J]. *Industrial Control Computer,* 2017,30(06):73-75.
6. Chen Peng-Zhan Zhu Nian-Hua Li Jie. Attitude Detection And balance Control Of Two-wheeled Self-balancing Vehicle[C]. *Chinese Process Control Conference* (CPCC2015), 2015.7.37, China, pp. 324
7. Huailin Zhao, Zheng Wu, Xiaoxing Wang. A Multi-robot Rescuing System[C]. *The 2018 International Conference on Artificial Life and Robotics* (ICAROB 2018), 2018.02, Japan. pp. 582-586
8. Sun Meng-fan Gui Wen-jie. Design of Control System for Two-Wheel Self-Balancing Robot[J]. *The Journal of Shandong Agriculture and Engineering University*, 2019,36(08):13-16.
9. Luo Can. Research on Control System of Two-wheeled Self-balanced Vehicle Based on Fuzzy Adaptive PID [D]. *Jilin University*, 2018
10. Liu Chengtao, Li Xianpeng. Self-balancing Car System Based on Attitude Detection and Fuzzy PID Algorithm[J]. *Microcontrollers & Embedded System*, 2019,19(07):78-80+85.

Crowd Counting Method Based on Improved CSRnet

Huailin Zhao, Shengyang Lu, Li Wang, Zhen Nie, Yaoyao Li
School of electrical and Electronic Engineering, Shanghai Institute of Technology, No. 100, Haiquan Road, Fengxian District, Shanghai 201418, China
E-mail: 405838753@qq.com
https://www.sit.edu.cn/

Abstract

Aiming at the problem of population counting, the research is getting deeper and deeper, CSRnet proposed a method of expanding convolution instead of convolution layer and pooling layer. This paper mainly proposes a multi-scale expansive convolutional neural network, which uses a method similar to the inception-ResNet module to calculate the population density of dense crowds with large scale changes and severe occlusion. And apply this method to ShanghaiTech dataset. The experimental results show that compared with CSRnet, the accuracy of this method has been improved, and the speed of feature extraction has also been improved.
Keywords: Crowd counting; CSRnet; inception-ResNet; multi-scale extended convolutional neural network

1. Introduction

With the continuous development of the national economy and modernization, the process of urbanization is accelerating, and the floating population in cities is increasing. Especially in metropolises such as Shanghai and Beijing, the flow of people has increased dramatically with the process of urbanization, the rapid increase of the floating population, the bustling streets have become more and more crowded, and the number of participants has increased during the opening of major exhibitions. In order to ensure social order and smooth urban road circulation, it is important to monitor the number of people in densely populated places [1]. Calculating the distribution of crowd density by detecting the number of people in various places can effectively prevent various security accidents. Crowd counting, as an effective way to calculate the number of people, has developed rapidly in recent years. Although the research on crowd counting methods by major institutions and scholars is getting deeper and deeper, there are still several major problems for crowd counting that have not yet been overcome: First, in the same picture, the scale of the crowd is different, which has a particularly large impact on the error of the crowd count; Second, the population distribution is different in different scenarios; Third, the problem of crowd counting under occlusion conditions arises; Fourth, the change of pedestrians at different times in the same scene[2]. In the same scene, the number, density and scale of the crowd are different, so for the counting method, it is necessary to have strong robustness for the crowd counting method in the same scene. In order to overcome the above problems, researchers at home and abroad have proposed various methods to improve the accuracy and adaptability of counting[3].

Traditional crowd counting methods mainly use detection-based methods and regression-based methods[4]. The detection-based method mainly uses a sliding window detector to detect the number of pedestrians in the image. This method mainly includes two detection methods. One is to use the overall detection method. This method detects the overall characteristics of pedestrians in the image and calculates the wavelet of the entire body of the pedestrian. It uses the edge features of HOG[5] to detect and then calculates the entire picture Number of pedestrians. The second way is to use the body part method[6]. This method mainly uses the characteristics of a part of the human body, and counts the entire population by counting this characteristic part. For dense crowds, the occlusion between the crowds is more serious, and the overall feature extraction is more difficult, so the part-based method is obviously more accurate than the whole-based method. The regression-based method is mainly divided into two steps. The first step is to extract features in the image, such as edge features, texture features, etc; The second step is to learn a regression model and use this model to map the characteristics of the crowd to the number of crowds. However, for a crowded crowd, the spatial scale of the crowd varies greatly, and the crowd's occlusion is serious. Therefore, the traditional method is not particularly effective and the accuracy of the crowd counting is low.

Li et. al [8]. Proposed a crowd counting method using VGG16 as the front-end network and hole convolution as the back-end network. This method has proved its superiority on the ShanghaiTech dataset, theUCF_CC_50 dataset, the World EXPO'10 dataset, and the UCSD dataset. Dilated convolution can effectively increase the perceptual field of the network, and more comprehensively extract the features of the entire network. The use of hole convolution can replace the convolution layer plus the maximum pooling layer, which can reduce the loss of information during the maximum pooling process. However, the network is not particularly good at the efficiency of feature extraction, this example verifies that multi-branch convolution can be more comprehensive, extract image features more efficiently, and improve the accuracy of crowd counting.

Recently, Chen Meiyun et al [9]. Proposed a crowd counting method based on pixel-level attention mechanism, which proposed a single-row crowd density estimation network based on pixel-level attention mechanism and improved. This network consists of two parts. First, the classification is performed at the pixel level to generate a high-quality local population density map. Classify the density map based on the density, and then use the classification results and a single-column population density estimation network to learn, learning more features with fewer parameters. However, the feature extraction of the entire picture is not very comprehensive, so although the effect is much improved compared to CSRnet, the effect of feature extraction is not particularly excellent.

In order to improve the accuracy and speed of the crowd counting method, this paper refers to the learning methods of the ResNet and inception-V1 modules, and proposes an improved CSRnet method. Each column of the convolutional neural network uses different size filters to obtain the same picture, for charactering people at different scales. It increases the robustness of the entire system, make the density map generated more accurate, and thus more accurately get the number of pedestrians in different pictures.

1. Related Work

Crowd counting using deep learning is generally divided into two types of methods. The first is an end-to-end counting method[6], which is to input images, go through a deep network, and then directly output the number of people. The second is
to get the density map of the crowd first, and then use the sum or integral method to count the crowding[7]. For the first method, the second method can more effectively get information about the human head. This article mainly uses the second method to obtain the density map of the crowd by using a multi-scale expanded convolutional neural network. The density map is integrated to calculate the number of crowdings[10].

2.1. Calculation Method of Crowd Density Map

Data sets are an essential part of a neural network architecture .For various datasets used for crowd counting, the position of the pedestrian in the picture is generally given, and the center position of the pedestrian is calculated by the counting method. Using the parameters of the data, the initial density map and the initial density can be labeled. The calculation formula of the graph is shown in the following formula (1).

$$M = \sum_{i=1}^{j} \sigma(P_i) \tag{1}$$

Where M is an $n \times m$ matrix with the same size as the original image, and the number of channels of the matrix is 1. j is the number of pedestrians. P_i is the coordinates of the pixel position of the ith pedestrian in the picture. $\sigma(P_i)$ means that there are pedestrians at P_i, so the value of this pixel is 1, and there are no pedestrians elsewhere, so the value of other pixels is 0. Therefore, a density matrix with the same size as M and the same number of channels can be formed. Filtering the matrix M by Gaussian kernel on M can form the true density map F. The specific density map formation formula is shown in the following Eq. (2).

$$F = M * G_{\delta i} \tag{2}$$

Among them, $G_{\delta i}$ is a Gaussian filter function with a total of 1. δi is the parameter of this Gaussian filter function. For the parameter setting of the specific Gaussian filter function, please refer to[5].Therefore, it can be found through the above two formulas that the sum of the elements of the density map M is related to the number of pedestrians. When the number of pedestrians is greater, the sum of the density maps is larger. The specific density map is shown in Figure 1. The density map is crucial for the framework of the entire convolutional neural network, we can figure out the number of crowing from the density map.

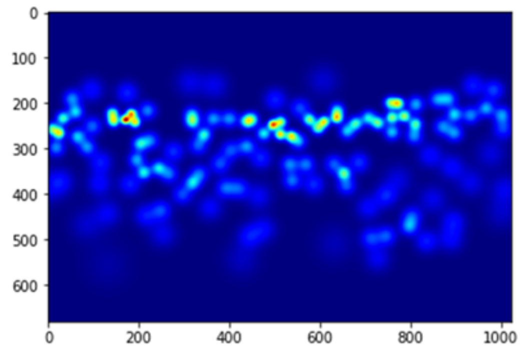


Fig. 1. Crowd Density Map

2.2 Construction of Loss Function

Referring to the most effective method at present, this paper adopts the construction method of Euclidean distance as a loss function, and uses the loss function to measure the difference between the predicted population density map and the real density map, and then uses the back propagation algorithm to adjust the parameters of the entire network. , Which ultimately optimizes the parameters. The specific loss function expression is shown in Eq. (3)

$$\mathrm{L}(\theta) = \frac{1}{2N}\sum_{i=1}^{N}\|V(x_i,\theta) - V_i\| \tag{3}$$

Where N is the number of training pictures. x_i represents the specific input picture, $V(x_i,\theta)$ represents the crowd density map obtained using this network, and $V_i$represents the real crowd density map. θ is the weight parameter of this network.

2. Multiscale Dilated Convolutional Neural Network

In this paper, ShanghaiTech dataset is used as the dataset. As ShanghaiTech dataset_A has a large crowd density and severe occlusion, also, the scale of the human head also changes greatly. The original input image is shown in Figure 2.

In close shots, the size of a pedestrian's head is larger, and in the distance, the size of the head is smaller. Therefore, considering only a single filter may not be very effective for the accuracy of crowd counting. But for the traditional convolutional neural network, the larger filter can reduce the error caused by the change of the human head scale. Therefore, using multi-layer parallel network can solve this problem well.

Fig.2. The original input image

The detailed structure of the multi-scale expansion convolutional neural network is shown in Figure 3. The overall network was inspired by CSRnet and inception-ResNet and ResNet. An improved convolutional neural network model based on CSRnet was designed. The network is based on CSRnet. By improving CSRnet and introducing multiple columns of parallel sub-neural networks, it can expand the perception of the entire network, improve the efficiency of feature extraction of the entire network, and use convolution kernels of different sizes to deal with head scale changes in pictures Big problem .The front-end network mainly uses the first ten layers of VGG16. The specific network structure diagram is shown in table 3. VGG16 has been widely used by many experts and scholars due to its strong learning and flexible architecture. However, VGG16 itself is not ideal for crowd counting networks, so we use the first 10 layers of VGG16. For the back-end network, a four-layer parallel neural network is designed.

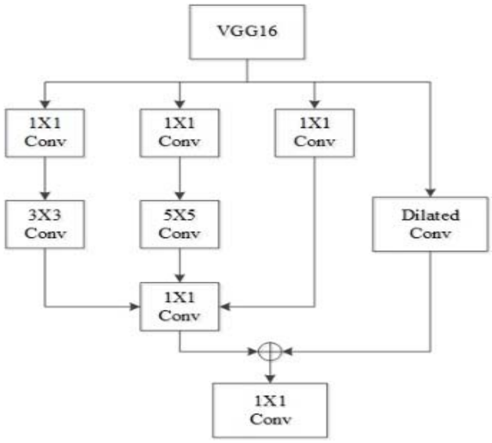


Fig. 3. The multi-scale expansion convolutional neural network

Table1. The First Ten Layers of VGG16

16 weight layers
Input image
Conv3-64
Conv3-64
maxpool
Conv3-128
Conv3-128
maxpool
Conv3-256
Conv3-256
Conv3-256
maxpool
Conv3-512
Conv3-512
Conv3-512

For the first branch, this paper uses an expanded convolution kernel, whose calculation formula is shown in Eq. 4.

$$y(\mathrm{m,n}) = \sum_{i=1}^{m}\sum_{j=1}^{n} x[m + r \cdot i, n + r \cdot i] \cdot \omega[i,j] \quad (4)$$

Among them, $y(\mathrm{m,n})$ is the output of the input feature map after convolution calculation, $x[m,n]$is the input feature map, $\omega[i,j]$ is the parameter of the convolution kernel, and r is the expansion Expansion factor of convolution,.When $r = 1$, the expansion convolution becomes a normal convolution. In the calculation of convolution, the larger the size of the convolution kernel, the larger the mapping area obtained by the calculation. For the calculation of a dilated convolution, it has greatly improved the area mapped by the traditional neural network, and the calculation rate has decreased a little. Not only that, through the calculation of the network, the calculation of an expansion convolution is equivalent to one convolution, one pooling, and one deconvolution operation.

The second branch is composed of three convolution kernels, one 1×1convolution kernel, one 3×3 convolution kernel, and one 1×1 convolution kernel. Because the convolution kernel of the second layer is small, the convolution operation of the second layer is mainly to extract the underlying information in the image, mainly to extract the information of the head of the person far away in the dense crowd, and the head of the person being blocked information.

The third branch is also composed of three convolution kernels, one 1×1 convolution kernel, one 5×5 convolution kernel and one 1×1 convolution kernel. The convolution kernel of the third layer is large, so it is mainly used to extract large-scale human head information. The size of the human head is larger near the camera, so the third layer can more comprehensively extract the feature information of the entire image.

The fourth branch only uses a 1×1 convolution kernel to extract the overall image. The size of the four-layer image is first integrated, and then a 1×1 convolution kernel is used to integrate the information of the four-layer image to extract the whole Image features to generate the final density map. Finally, the generated density map is used to integrate to calculate the number of pedestrians in the entire picture.

It is worth noting that the pooling layer of the back-end network of the traditional convolutional neural network designed this time is replaced by a full convolutional layer. This design idea is because the role of pooling is mainly to reduce the dimension and reduce the amount of calculation. However, using the pooling layer, whether it is the largest pooling layer or the average pooling layer, part of the information is missing, and the convolution calculation operation with a large scale can already realize the idea of dimensionality reduction. Therefore, in this network, the back-end network does not adopt a pooling structure. Not only that, the 1×1 convolution kernel is used in the designed network instead of the fully connected layer. This design is because the calculation parameters of the fully connected layer are complicated, the calculation amount is huge, and the network is prone to overfitting. If a 1×1 convolution kernel is used instead of the fully connected layer, not only the computing time is reduced and the computing efficiency is improved, but the input feature map can be an image of any scale.

3. Experimental Results and Analysis

4.1. Counting Performance Indicators

At present, a common way to evaluate the performance of crowd counting methods is to calculate their average absolute error (MSE) and mean square error (MAE), and evaluate the entire system by calculating and judging the average absolute error and mean square error. MSE is the expected value of the square of the difference between the estimated value of the parameter and the true value of the parameter, MAE is the average of absolute errors. For the entire system, the smaller the average absolute error is, the higher the accuracy of the crowd counting, and the smaller the mean square error, the better the robustness of the system. The calculation method of the average absolute error and the mean square error is shown in Eq. 5 and Eq. 6.

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|G(i) - P(i)| \quad (5)$$

$$\mathrm{MSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(G(i) - P(i)\right)^{2}} \quad (6)$$

Among them, N is the number of training pictures, and $G(i)$ and $P(i)$ are the predicted and true values of the number of pedestrians in the ith picture, respectively.

4.2. Processing of multi-scale extended convolutional neural network training

In this experiment, ShanghaiTech dataset is selected as the dataset. ShanghaiTech dataset is mainly divided into two parts, namely ShanghaiTech dataset_part_A and ShanghaiTech dataset_part_B. Among them, part_A is 482 pictures obtained from the Internet. The pictures are different in size, and the crowd is crowded. Therefore, 300 of them were selected as the training set and 182 were used as the test set. The part_B part is mainly a surveillance video from the streets of Shanghai. There are a total of 716 pictures. The pictures are of different sizes, but the density of the crowd is smaller, it is easier to count for the crowd counting network, and the counting accuracy is higher. . Therefore, 400 of them were selected as the training set and 316 were used as the test set.

4.3. Result Analysis

In order to obtain the performance indicators of the multi-scale expansive convolutional neural network more clearly, this experiment shows the training process of this network for part_A, and compares it with CSRnet training. The relationship between MAE and the number of iterations during training is shown in Figure 5.

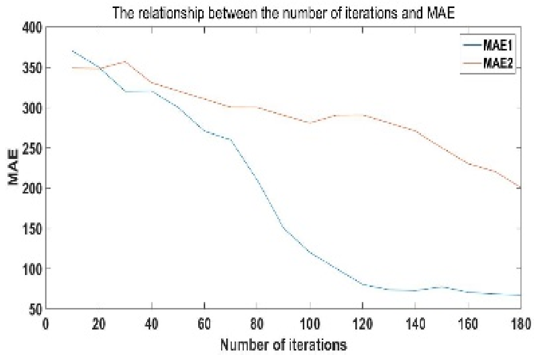


Fig. 5. The relationship between the number of iterations and MAE.

The experimental results show that the use of multi-scale dilated convolutional neural networks can greatly increase the rate of feature extraction of the entire network. The multi-scale dilated convolutional neural network decreases MAE to 73.4 when iterating 132 times, while CSRnet requires iteration about 320 Times. However, the convergence speed of multi-scale dilated convolutional neural network training is slower than that of CSRnet. This is mainly because the parameters have become more complicated, but the MAE of multi-scale dilated convolutional neural network is relatively low.

Tested on the ShanghaiTech dataset and compared with other algorithms, the performance of the multiscale dilated convolutional neural network has also improved to a certain extent. The test results of the multiscale dilated convolutional neural network and other different algorithms are shown in Table 2.

Table 2. Results Comparativation table

	PART_A		PART_B	
	MAE	MSE	MAE	MSE
MCNN[11]	110.2	173	26.4	41.3
SCNN[10]	90.4	90.4	21.6	33.4
Switch_CNN[12]	90.4	135	21.6	33.4
CSRnet[13]	68.2	115	10.6	16.0
ours	70.2	110	20.0	24.0

According to the comparison in the table above, compared with Switching-CNN, the multi-scale expanded convolutional neural network has MAE decreased by 20, MSE decreased by 24, compared with CSRnet, MAE increased by 2 and MSE decreased by 5. It can be seen that compared with other algorithms, the multi-scale dilated convolutional neural network has greatly improved the accuracy and robustness of crowd counting. Therefore, the multi-scale expanded convolutional neural network is suitable for places with dense crowds and large changes in crowd scales.

4. Conclusion

This paper designs a multi-scale expanded convolutional neural network. The network uses the first 10 layers of VGG16 as the front-end network, a 4-column parallel sub-neural network as the back-end network, and then uses a 1×1 convolution kernel Get crowd density. Compared with previous networks, this network can get more inf- ormation from pictures, reduce feature loss, and has a good counting effect under crowded crowds, large pedestrian scale changes, and crowd occlusion. Experiments on the ShanghaiTech dataset also show that compared with other networks, the multi-scale extended convolutional neural network has greatly improved both in accuracy and robustness, and is suitable for crowded crowds and large changes in crowd size .

References

1. Li Wang, Huailin Zhao, Yaoyao Li. Research on Research on the multi-scale network crowd density estimation algorithm based on attention mechanism. The 2019 International Conference on Intelligent Informatics and BioMedical Science (ICIIBMS2019), Nov.21-24, 2019, Shanghai, China, pp272-278.
2. Yawei Hou, Huailin Zhao. Handwritten Digit Recognition Based on Depth Neural Network. 2017 International Conference on Intelligent Informatics and BioMedical Science (ICIIBMS2017), Nov.24-26, 2017, Okinawa, Japan. pp35-38. (EI)
3. Shunzhou Wang, Huailin Zhao, Weiren Wang, Huijun Di, Xueming Shu. Improving Deep Crowd Density Estimation via Pre-Classification of Density. The 24th International Conference on Neural Information Processing (ICONIP2017), Nov.14-18, 2017, Guangzhou, China. pp260-269. (EI)
4. Ryan D, Demon S, Sridharan S, et al. An evaluation of crowd counting methods, features and regression mod-els[J]. Computer Vision and Image Understanding,2015(130): 1-17
5. Lu Jingang, Zhang Li. Dense crowd counting model based on multiscale and multicolumn convolutional neural network [J / OL] .Computer Applications: 1-7 [2019-12-09].
6. Guo Ruiqin, Chen Xiongjie, Luo Wei, Fu Changhong.A crowd counting method based on optimized Inception ResNet A module and Gradient Boosting [J] .Journal of Tongji University (Natural Science), 2019,47 (08): 1216-1224 .
7. Cao Jinmeng, Ni Rongrong, Yang Biao. Crowd counting based on multiscale and multitasking convolutional neural network [J] .journal of Computer Applications, 2019,39 (01): 199-204.
8. Chen Meiyun, Wang Bisheng, Cao Guo, Liang Yongbo.Crowd counting method based on pixel-level attention mechanism [J / OL] .Computer Applications: 1-9 [2019-12-09].
9. Yahan Wang, Huihui Bai, Lijun Zhao, Yao Zhao. Cascaded reconstruction network for compressive image sensing[J].EURASIP Journal on Image and Video Processing, 2018, Vol.2018 (1), pp.1-16
10. Zhang Y，Zhou D，Chen S，et al. Single- image crowdcounting via multi- column convolutional neural network[C]// IEEE Conference on Computer Vision and Pattern Recognition, 2016: 589597
11. Peng Shanzhen, Fang Zhijun, Gao Yongbin, Huang Bo, Wu Chenmou. Crowd counting based on multi-scale full convolutional network feature fusion [J / OL]. Journal of Wuhan University (Science Edition), 2018 (03): 249-254 [2019-12-09].https: //doi.org/10.14188/j.1671-8836.2018.03.008.
12. Sam D B, Surya S, Babu R V. Switching convolutionalneural network for crowd counting[C]// IEEE Conference on Computer Vision and Pattern Recognition, 2017, 1(3)6.
13. Li Y, ZHANG X.CHEN D.Csrnet:dilated convolutional neural for understanding the highly congested scenes[C]//IEEE conference on Computer Vision and Pattern Recognition.Salt Lake:IEEE,2018:1091-1100

Graph-based Global Reasoning Network for Crowd Counting

Li Wang, Huailin Zhao, Zhen Nie, Yaoyao Li
School of Electrical and Electronic Engineering,
Shanghai Institute of Technology, Shanghai, China
E-mail: 736036396@qq.com, zhao_huailin@yahoo.com
www.sit.edu.cn

Abstract

Convolutional neural network (CNN) has prompted the crowd counting task to massive progress in recent years. However, filters in the shallow convolutional layer of the CNN only model the local region rather than the global region, which can't capture context information from the crowd scene efficiently. To solve this problem, in this paper, we propose a Graph-based Global Reasoning Network (GGRNet). Each input image is processed by the VGG-16 network for feature extracting, and then the Graph-based Global Reasoning Unit (GGRU) reasons the context information from the extracted feature. Especially, the extracted feature firstly is transformed from the feature space to the interaction space for global context reasoning with the graph convolutional network (GCN). Then, the output of the GCN projects the context information from the interaction space to the feature space. The experiments on the UCF-QNRF dataset, and the results demonstrate the effectiveness of our method.

Keywords: Graph-based Global Reasoning Unit, Graph convolutional network, Crowd density estimation

1. Introduction

With the rapid development of society and economy, large-scale crowd gathering has become very common, and the significance of effective supervision of crowd gathering has become increasingly important. Counting the number of people in a high-density population is a very difficult task, because people appear in many different ways in various scenarios.

The current mainstream crowd counting algorithm is a crowd density estimation algorithm based on deep learning. It mainly uses the powerful feature expression ability of convolutional neural networks to learn the non-linear mapping relationship between input images and output density maps. Compared with the traditional crowd counting algorithms, convolutional neural networks have made significant progress in crowd counting. However, that convolution neural network only model the local region, and the effect of the global region is slightly poor. This leads to difficulty in extracting context information and poses a certain challenge for accurate crowd density estimation.

In order to overcome the inherent limitations of the convolution operation, a global reasoning network based on the graph convolution network (GCN) is proposed for crowd counting, and the global relation of the input crowd is modeled. The extracted crowd features are transformed from the feature space to the interaction space for global context inference, the context information is inferred and interpreted in the interaction space, and then the context-aware features are returned to the original coordinate space and combined with the features that were originally extracted. The use of global inference networks increases the relationship between different regions of the image and improves the accuracy of crowd density estimation.

1. Related Work

Traditional machine learning methods are difficult to design crowd features, and it is difficult to obtain one or a set of features to meet the challenges of crowd density, crowd occlusion, different scenarios and perspectives encountered in crowd counting problems. In recent years, due to the improvement of computer computing power, especially the significant enhancement of graphics card performance, deep learning technology has promoted many computer vision problems to achieve very great breakthroughs. Massive data training enables deep learning to obtain powerful feature representation capabilities, which can well cope with the challenges and difficulties commonly found in crowd counting, and obtain more accurate counting results.

In [1], for the first time, a non-linear mapping relationship between the density map and the input image was directly established using a convolutional neural network. A multi-column convolutional neural network MCNN was proposed. The MCNN network uses convolution kernels of different sizes to adapt to the head sizes of different scales, and then combines the three columns of convolutional neural networks to obtain the final crowd density map. Reference [2] added a density pre-classification network to the multi-column network. However, running a three-column convolutional neural network is slower. Since then, [3] proposed a hole convolutional neural network (CSRNet). Using a convolution kernel with holes instead of the pooling layer and the convolution layer reduces the amount of data and expands the receptive field range without losing the spatial resolution, which can extract more feature information. In order to regulate the multi-scale density map, [4] also proposed a new scale-perceived loss function to guide the network to specialize in specific head sizes.

In the last two years, attention models have been widely used in various types of deep learning tasks. Literature [5] proposed a new two-way multi-scale fusion network structure SFANet. This method generates the final high-quality and high-resolution density map through attention mechanism and multi-scale fusion. Reference [6] also uses attention mechanism.

The above research shows that the problem of crowd picture scale changes has always been the focus and difficulty of crowd counting tasks. In a picture, due to the change in the distance of the camera, two similar people show significant differences when they are at different distances from the camera. For this reason, we can consider increasing the connection between contexts to improve the accuracy of crowd density estimation. Based on the above research, this paper proposes a graph convolution-based global reasoning network (GGRNet). Global feature units are used to model long-range regional relationships of feature pictures, and context information is added for the final population density estimation.

2. Proposed Method

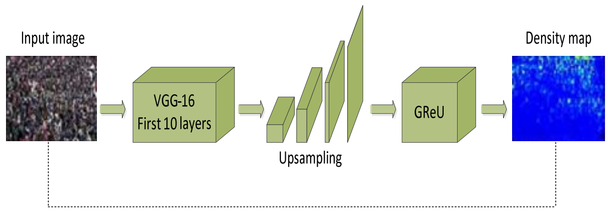


Fig.1.The overall flowchart of the algorithm.

Graph-based convolution-based reasoning methods have been widely used in deep neural networks in recent years, and have greatly improved network performance in many tasks. For example, the GCN network proposed in [7] is used for semi-supervised classification. [8] used graph neural networks to learn feature matching. This paper proposes a graph-based global reasoning network for crowd counting. The overall flowchart of the algorithm is shown in Fig.1.

In this paper, VGG-16 network is used for feature extraction of crowd images because of its strong learning ability and flexible architecture, which is convenient for connecting density map generation at the back end. Input the picture and extract the shallow feature information of the crowd image through the convolution operation of the first ten convolution layers of the VGG-16 network. Its model structure is shown in Fig.2. In the middle, it experienced three maximum pooling processes, and the output feature map size became 1/64 of the original input picture.

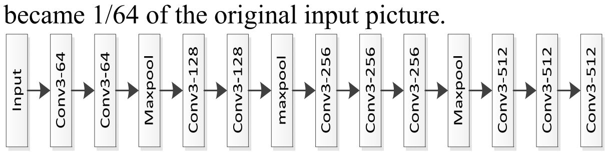


Fig.2. The first ten convolution layers of the VGG-16 network.

Since the feature image output by the VGG-16 network is only 1/64 of the original image size, and the global reasoning unit is used to model the long-range area of the crowd image, a design is designed between the VGG-16 network and the global reasoning unit Upsampling operation to enlarge the image. In this paper, a bilinear interpolation upsampling method is used, and a convolution process is added before and after it to help correct the feature deviation due to upsampling. The model structure is shown in Fig.3.

Fig.3. The model structure of upsampling.

Referring to the literature [9], a global reasoning unit GReU is added to the network. Since a single convolutional layer can only capture the local area relationships covered by the convolution kernel, and the relationship between long-distance areas needs to be achieved by superimposing multiple convolutional layers, this is not only cumbersome and inefficient, but also for network training Added difficulties.

In order to overcome this limitation and better obtain the global context feature information, a global reasoning unit is selected to be added to the network to project the crowd characteristics on the coordinate space to the interactive space of global context reasoning. In this interaction space, regions with similar semantic features are represented by the same feature node, and their features are stored to form a fully connected graph. Through the convolution operation, the interaction relationship between each node in the fully connected graph is modeled to obtain global context information. The network then maps the resulting global feature map back from the interaction space to the feature space, combines it with the crowd features that were originally extracted, and outputs the final crowd density estimate map. Sum all pixels in the density map to calculate the total number of people in the crowd picture.

3. Experiments

3.1. *Dataset*

In this paper, the UCF-QNRF dataset is used to train and test the network. This data set was first proposed and published in the ECVC2018 literature [10], and it is the new largest data set for population counting.

The UCF-QNRF dataset overcomes some of the shortcomings of previous datasets, including a variety of scenes with the most diverse perspectives, density, and lighting changes, making the dataset more realistic and representative. The UCF-QNRF dataset contains a total of 1535 pictures, and the number of labeled heads reached 1125642, with an average of 815 people per picture.

The crowd picture resolution in the dataset is relatively large, with an average resolution of 2013 × 2902. The head size of the crowd also varies greatly, and the number of people per picture ranges from 49 to 12,865, which is especially suitable for training deep convolutional neural networks.

3.2. *Density Map Generation*

Crowd density estimation During network training, crowd pictures are used as the input of the network, and crowd density maps are used as the output of the network. Therefore, the head position coordinates given in the data set must first be converted into corresponding crowd density pictures. In this paper, a method for generating a population density map based on an adaptive Gaussian kernel is used. Its formula is as follows:

$$F(x) = \sum_i^N \delta(x - x_i) * G_{\sigma_i}(x), \sigma_i = \beta \bar{d}^l \quad (1)$$

Where $G_{\sigma_i}(x)$ represents the Gaussian convolution kernel, x_i is the position coordinates of the human head in the image, $\delta(x - x_i)$ is the Dirac function of the human head, N is the total number of people included in the image, and $\bar{d}^l = \frac{1}{m}\sum_{j=1}^m d^i{}_j$ represents the average distance of the m heads closest to the head. In denser cases it is approximately equal to the head size. β is a hyperparameter, here it takes 0.3. The loss function formula to be optimized by the network is as follows:

$$L(\theta) = \frac{1}{2N}\sum_{i=1}^N \left\| \hat{y}(x_i;\theta) - y_i \right\|_2^2 \quad (2)$$

Where θ represents the parameters to be optimized by the network, N is the number of pictures in the training set, x_i represents the input picture, $\hat{y}(x_i;\theta)$ represents the crowd density map estimated by the network, and y_i represents the crowd density truth value picture corresponding to the input picture.

For network training and implementation, we chose the Pytorch framework. When training the network, the Adam algorithm was used to optimize the network, and the initial learning rate was set to 0.00001.

3.3. *Evaluation Metric*

At present, the evaluation of the pros and cons of the crowd counting algorithm mainly includes the following two indicators, mean absolute error (MAE) and mean squared error (MSE).

$$MAE = \frac{1}{N}\sum_{1}^{N} | z_i - \hat{z}_l | \qquad (3)$$

$$MSE = \sqrt{\frac{1}{N}\sum_{1}^{N} (z_i - \hat{z}_l)^2} \qquad (4)$$

Where N represents the total number of pictures in the test sequence, z_i represents the actual number of people in the picture, and $\hat{z}_l$ represents the estimated number of people in the picture.

3.4. *Experimental Results*

The graph-based global inference network is used for crowd counting. The experimental results on the UCF-QNRF dataset are shown in Table 1.

As can be seen from the table above, the method proposed in this paper is tested on the UCF-QNRF dataset, with MAE of 111.2 and MSE of 189.4. Compared with other more advanced crowd density estimation networks, this method has the smallest test error, and the network performance improves significantly.

Table 1. The experimental results on the UCF-QNRF dataset.

Method	MAE	MSE
Idrees 2013[11]	315	508
MCNN[1]	277	426
CMTL[12]	252	514
Switching CNN[13]	228	445
Resnet101[10, 14]	190	277
Densenet201[10, 15]	163	226
CL[10]	132	191
Our Proposed	**111.2**	**189.4**

In order to more clearly see the effect of this method on crowd density estimation, the density map output by the network was visualized and compared with the true value density map. Five images with different density levels were randomly selected for display, as shown in Fig.4.

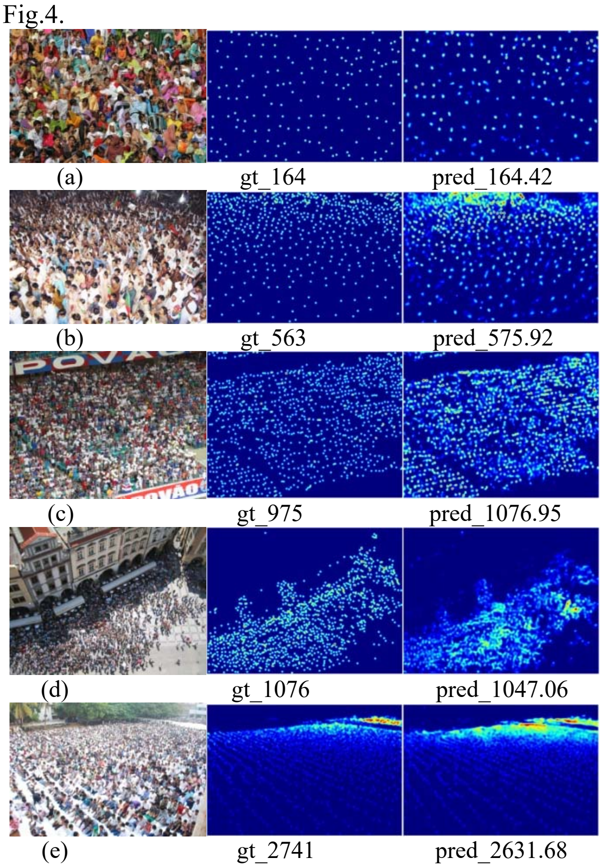


Fig. 4. Visualization of density map.

In Fig. 4, the first column is the original crowd picture, the second column is the true value density map, and the third column is the network-estimated crowd density map. It can be seen from the figure that the population density map estimated by the network can vividly reflect the population distribution.

4. Conclusion

In this paper, we propose a graph-based global reasoning network for crowd density estimation tasks. Model the crowd feature information extracted by the VGG-16 network within the global area. Use the Global Reasoning Unit (GReU) to establish the relationship between long-distance areas in the picture, add context information, and provide supplementary features for the generation of crowd density maps. The original features are combined to make the extracted crowd features more complete, so as to obtain a more accurate crowd density estimation map. Experimental results on the UCF-QNRF dataset show the effectiveness of the method.

5. References

1. Zhang, Y., Zhou, D., Chen, S., Gao, S., & Ma, Y.. Single-image crowd counting via multi-column convolutional neural network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2016), pp. 589-597.
2. Wang S, Zhao H, Wang W, et al. Improving Deep Crowd Density Estimation via Pre-classification of Density， In *Proceedings of the IEEE International Conference on Neural Information Processing*. (Springer, Cham, 2017), 260-269.
3. Li, Y., Zhang, X., & Chen, D.. Csrnet: Dilated convolutional neural networks for understanding the highly congested scenes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2018), pp. 1091-1100.
4. Varior R R , Shuai B , Tighe J , et al. Multi-Scale Attention Network for Crowd Counting. *arXiv preprint arXiv*:1901.06026, (2019).
5. Zhu L , Zhao Z , Lu C , et al. Dual Path Multi-Scale Fusion Networks with Attention for Crowd Counting. *arXiv preprint arXiv*:1902.01115 ,(2019).
6. Wang L., Zhao H., Li Y.. Research on the multi-scale network crowd density estimation algorithm based on attention mechanism. *The 2019 International Conference on Intelligent Informatics and BioMedical Science (ICIIBMS2019)*, (Nov.21-24, 2019, Shanghai, China), pp.272-278.
7. Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks[J]. *arXiv preprint arXiv*:1609.02907, (2016).
8. Sarlin, Paul-Edouard, et al. "SuperGlue: Learning Feature Matching with Graph Neural Networks." *arXiv preprint arXiv*:1911.11763, (2019).
9. Chen, Y., Rohrbach, M., Yan, Z., Shuicheng, Y., Feng, J., & Kalantidis, Y.. Graph-based global reasoning networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (2019), pp. 433-442.
10. Idrees, H., Tayyab, M., Athrey, K., Zhang, D., Al-Maadeed, S., Rajpoot, N., & Shah, M.. Composition loss for counting, density map estimation and localization in dense crowds. In *Proceedings of the European Conference on Computer Vision (ECCV)*,(2018), pp. 532-546.
11. Idrees, H., Saleemi, I., Seibert, C., & Shah, M.. Multi-source multi-scale counting in extremely dense crowd images. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2013), pp. 2547-2554.
12. Sindagi, V. A., & Patel, V. M.. Cnn-based cascaded multi-task learning of high-level prior and density estimation for crowd counting. In*2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*(2017, August), pp. 1-6.
13. Sam, D. B., Surya, S., & Babu, R. V. . Switching convolutional neural network for crowd counting. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* , (2017, July), pp. 4031-4039.
14. He, K., Zhang, X., Ren, S., Sun, J.Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2016), pp. 770–778 .
15. Huang, G., Liu, Z., Weinberger, K.Q., van der Maaten, L.: Densely connected convolutional networks. *arXiv preprint arXiv*:1608.06993, (2016).

A Design of Ocean Current Velocity Measuring Device

Jichao Zhao [1,3], Fengzhi Dai [1,2,3]*, Xin Ma [4], Fengkun Wang [1], Haokang Wen [1], Hongbo Hao [1], Qianqian Zhang [1]

[1] Tianjin University of Science and Technology, China,

[2] Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China,

[3] Advanced Structural Integrity International Joint Research Centre, Tianjin University of Science and Technology, Tianjin, 300222, China,

[4] National Oceanic Administration Beihai Marine Environmental Monitoring Center Station, China

*E-mail: * daifz@tust.edu.cn*
www.tust.edu.cn

Abstract

Ocean current refers to the relatively stable flow of the surface water of the ocean in a certain direction in a certain direction, and is the main regulator of the thermal environment on the surface of the earth. At present, the ocean current velocity detection at sea is basically a field measurement by a staff member in a boat, and the adjustment of the detector position during the measurement process is very inconvenient. This paper designs a new submarine current velocity measuring device, including measuring ship, cable, fixed anchor and ocean current velocity detector. The cable on the reel is retracted through the cable retracting room, and the end of the cable is fixedly provided with a fixed anchor. The fixed anchor is fixedly provided with a plurality of ocean current flow rate detectors, and the ocean current flow rate detector has a built-in control chip, a wireless communication module, and Pressure Sensor. This device can flexibly adjust the depth of the detector according to actual needs, so as to measure the flow velocity of different depths.

Keywords: Ocean current; Flow velocity; Speed measuring device; Ultrasonic velocimetry

1. Introduction

1.1. Ocean current

In order to develop and utilize the ocean, people have conducted long-term observation and research on the ocean. Among them, the current velocity is also a very important parameter in marine hydrological research. It is very important to grasp the law of seawater flow. It can directly serve national defense, marine transportation, fishery, and port construction.

Ocean currents refer to the relatively stable flow of large-scale ocean surface waters in certain directions throughout the year. Ocean currents are the main regulators of the thermal environment on the Earth's surface. Ocean currents can be divided into warm currents and cold currents. If the temperature of the ocean current is higher than the water temperature reaching the sea area, it is called warm current; if the temperature of the ocean current is lower than the water temperature reaching the sea area, it is called cold current [1]. Generally, the ocean current flowing from low latitude to high latitude is warm current, and the ocean current flowing from high latitude to low latitude is cold current. Sea vessels sailing along the ocean current can save fuel and speed up. When warm and cold currents meet, sea fog often forms, which is not good for sea navigation. In addition, ocean currents carry icebergs south from the Arctic, posing a greater threat to maritime shipping. The main global currents are shown in Fig.1.

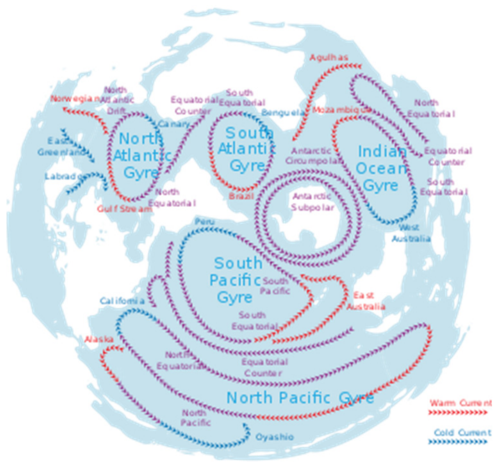


Fig.1. The main global currents

1.2. Common speed measurement methods

With the continuous development of science and technology and marine science itself, the methods of observing ocean currents have also been continuously improved and improved. From the original drift bottle to the modern ocean current meters, a large variety of current measuring instruments have been invented.

In recent years, with the increasing maturity and improvement of the following technologies, such as the digital signal processing, chip technology, optical fiber, laser technology, graphics and image processing, and computer technology, the flow velocity measurement has also undergone new developments, and acoustic Doppler has appeared A series of new flow velocity measurement methods such as flow velocity measurement, laser Doppler flow velocity measurement, radio wave velocity measurement, particle imaging velocity field measurement, etc [2].

In particular, acoustic flow measurement technology has become the most widely used solution in ocean velocity measurement with its unique advantages. At the same time, the study of acoustic current measurement technology is of great significance for exploring new approaches to ocean current measurement technology.

At present, the detection of ocean current velocity at sea is basically carried out by staff taking a boat to perform field measurements. However, it is inconvenient to adjust the position of the detector during the measurement [3]. It is prone to deviations due to the measurement of the device, so it is necessary to design a submarine ocean current speed measurement device to solve these problems.

2. Device design

In this paper, a new submarine ocean current speed measurement device is designed, which includes a measuring boat, cable, fixed anchor, and ocean current velocity detector. The cable on the reel is retracted through the cable storage room. The end of the cable is fixedly provided with a fixed anchor. The fixed anchor is fixedly provided with a plurality of ocean current velocity detectors. Pressure Sensor. This device can flexibly adjust the depth of the detector according to actual needs in order to measure the ocean current velocity at different depths.

2.1. Design notes

As shown in Fig.2: No.1 indicates the measurement ship; No.2 indicates the cable storage room; No.3 indicates the reel; No.4 indicates the cable; No.5 indicates the fixed anchor; No.6indicates the sliding sleeve; No.7 indicates the fixed No.8 indicates the sinking plate; No.9 indicates the ocean current velocity detector; No.10 indicates the rack; No.11 indicates the first motor; No.12 indicates the second motor; No.13 indicates the gear; and No.14 indicates the transmission belt.

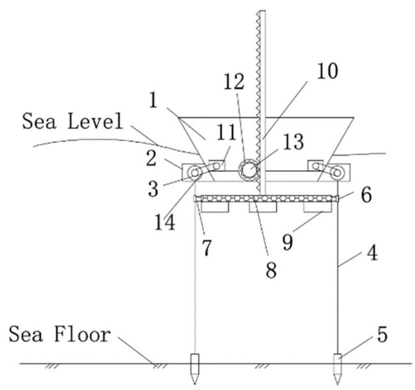


Fig.2. Design schematic 1

As shown in Fig.3: No.15 indicates a waterproof case; No.16 indicates a control chip; No.17 indicates a wireless communication module; No.18 indicates a pressure block; No.19 indicates a pressure sensor; No.20 indicates a connecting rod; and No.21 indicates a suspension.

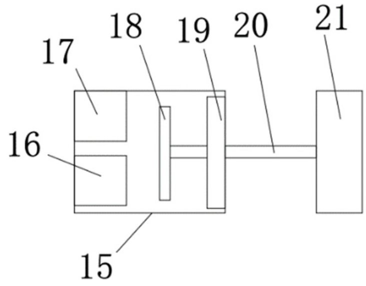


Fig.3. Design schematic 2

2.2. Instructions

A rewinding reel is provided by rotating inside the cable retracting room, and the reel is wound with a cable. The end of the cable is fixed with a fixed anchor. Two cables are provided with a sliding sleeve. The fixing rod is fixedly connected to the sinking plate. Several ocean current velocity detectors are fixedly arranged below the sinking plate. A rack is fixedly arranged above the sinking plate. A second motor and a second motor are fixedly arranged on one side of the rack. A gear is fixedly arranged on the output end of the gear [4]. The depth of the detector can be flexibly adjusted according to the actual needs, so that the ocean current velocity at different depths can be measured. At the same time, the ocean current velocity detector is provided with at least three groups. The average of the results measured by the detector is used as the measurement data, which further improves the accuracy of the measurement and is conducive to popularization and promotion.

3. Speed measurement principle

Acoustic waves are the propagation of energy in a propagation medium. Ultrasonic pulse signals generated by piezoelectric ceramic transducers propagate in the seawater medium in the form of phase velocity [5]. When the seawater medium itself has velocity, the sound waves travel downstream at the same distance. It is not the same as the time of countercurrent propagation [6]. The schematic diagram of the propagation plane is shown in the figure, where A and B are transceiver and replacement energizer, c is the ocean sound velocity at the measurement point, v is the velocity of the seawater, and d is the distance (channel axis distance) between the two transducers. During the measurement, the two transducer pairs are excited by the same electrical signal, that is, the two transducers transmit ultrasonic waves at the same time, and are ready to receive signals arriving at each other. Due to the existence of the velocity v, the ultrasonic pulse signals that cause the forward and backward propagation will not reach the transducer receiving surface at the same time [7]. According to the theory of sound wave retransmission in a fluid medium, it is known that the propagation time of the short ultrasonic pulses along the countercurrent and countercurrent directions on the biological axis are:

Downstream travel time:

$$t1 = d/(c + v) \tag{1}$$

Countercurrent propagation time:

$$t2 = d/(c - v) \tag{2}$$

Time difference between downstream and upstream:

$$\Delta t = t2 - t1 \approx 2vd/c^2 \tag{3}$$

Flow value:

$$v = c^2 \Delta t/(2d) \tag{4}$$

When c and d are known, as long as Δt is measured, the flow velocity v can be directly obtained, and the flow velocity v can be directly obtained. The error coefficient of this approximation is v^2/c^2. If the speed of sound and velocity are calculated according to typical values, the error magnitude brought by this approximation is $4 * 10^{-7}$. The ocean velocity in the submarine boundary layer is generally very low, ranging from about $0.1cm/s$ to about $1m/s$. Assuming that $d = 10\text{cm}$ and $c = 1500\text{m/s}$, the time difference range calculated from equation (4) is 1.33ns to 133ns. Such a small time difference can be ignored directly [8]. Speed measurement schematic is shown in Fig.4.

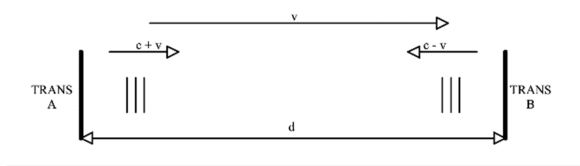


Fig.4. Speed measurement schematic

4. Conclusion

Compared with the prior art, the device can flexibly adjust the depth of the detector according to actual needs, so that the ocean current velocity at different depths can be measured. At the same time, there are at least three sets of ocean current velocity detectors, and the results measured by the three groups of ocean current velocity detectors are averaged as measurement data, which further improves the measurement accuracy and is conducive to popularization and promotion.

Acknowledgements

The research is partly supported by the Project of Tianjin Enterprise Science and Technology Commissioner to Tianjin Tianke Intelligent and Manufacture Technology Co., Ltd (19JCTPJC53700). It is also supported by the Industry-University Cooperation and Education Project (201802286009) from Ministry of Education, China.

References

1. Xiao Zhongjun, Zhou Qiang and Wang Cunrong, New Method to Measure Flow Velocity Based on Short Time Power Spectrum Analysis of Pulp Consistency Noise, *Proceedings of the 2008 Chinese Control and Decision Conference,* Yantai, Shandong, 2008, pp. 1608-1612.
2. K. H. Albusaidi, G. Lucas, Measurement of multiple velocities in multiphase flow. *IEE Colloquium on Advances in Sensors for Fluid Flow Measurement,* London, UK, 1996, pp. 12/1-12/4.
3. K. Jing, Z. Kan, F. Li, X. Wang, The study about flow velocity measurement with spatial filter principle based on the electrostatic sensor. *Proceedings of 2011 International Conference on Electronic & Mechanical Engineering and Information Technology,* Harbin, 2011, pp. 2865-2868.
4. X. Dong, C. Tan, Y. Yuan, F. Dong, Measuring Oil–Water Two-Phase Flow Velocity With Continuous-Wave Ultrasound Doppler Sensor and Drift-Flux Model. *IEEE Transactions on Instrumentation and Measurement,* 2016, 65(5): pp. 1098-1107.
5. C. Tan, X. Dong, F. Dong, Continuous Wave Ultrasonic Doppler Modeling for Oil–Gas–Water Three-Phase Flow Velocity Measurement. *IEEE Sensors Journal,* 2018, 18(9): pp. 3703-3713.
6. W. Xiaowei, Jiangyan, Liye, Algorithm Research on Velocity Measurement of Biphase Flow, *Proceedings of the 2009 Ninth International Conference on Hybrid Intelligent Systems,* Shenyang, 2009, pp. 471-474.
7. J. Hitomi, Y. Murai, H. J. Park, Y. Tasaka, Ultrasound Flow-Monitoring and Flow-Metering of Air–Oil–Water Three-Layer Pipe Flows. *IEEE Access,* 2017, 5: pp. 15021-15029.
8. X. Dong, C. Tan, F. Dong, Gas–Liquid Two-Phase Flow Velocity Measurement With Continuous Wave Ultrasonic Doppler and Conductance Sensor, *IEEE Transactions on Instrumentation and Measurement,* 2017, 66(11): pp. 3064-3076.

Device Design Based on TDS Water Quality Detection

Jichao Zhao [1,3], Fengzhi Dai [1,2,3]*, Di Yin [1], Yuhui Cheng [1], Fengkun Wang [1], Leixin Han [1], Qianqian Zhang [1]

[1] Tianjin University of Science and Technology, China,

[2] Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China,

[3] Advanced Structural Integrity International Joint Research Centre, Tianjin University of Science and Technology, Tianjin, 300222, China

*E-mail: * daifz@tust.edu.cn*
www.tust.edu.cn

Abstract

The traditional water quality detection method is sampling and detection by artificial nodes, which is easily interfered by weather and environment. The design is a new water quality detection device for marine waters and other waters. The device is spherical in shape and uses a built-in sensor and control system to detect the water quality in the form of drifting. This device has built-in micro control system, TDS detection module, power voltage regulator module, GPS module and wireless signal transmission module. The TDS module detects the concentration of total dissolved matter in the water, and the GPS module measures the current location of the device, which sends the data to the micro-control system. The micro control system sends the water quality information and the geographical location information to the network through the wireless transmitting module, and the user can view the water quality status by the device in real time through the Internet. The power supply voltage regulator module provides the different voltages required by the entire device to achieve the function of detecting the current water quality.

Keywords: TDS; Real-time detection; Micro control system; GPS module

1. Introduction

Looking at the increasingly serious environmental problems nowadays, the solution of water pollution is imminent. In order to solve these problems of pollution investigation and treatment, we have produced a device that integrates the detection function of pollutants, the traceability function of pollution sources, and the emergency treatment function. The device can collect water quality conditions in multiple directions for calculation and analysis through sensors at the same time, find the direction of the pollution source according to the principle of diffusion, trace it back, find the source of the pollution, put specific microbial algae after scientific calculation, and treat the pollution by biological methods[1]. The biggest feature of this device is the combination of pollutant detection, traceability, and treatment, and the biological treatment method adopted greatly avoids the problem of secondary pollution caused by water pollution in the expanded area. It can be applied to many aspects such as pollution detection and treatment in general waters, pollutant leakage detection in emergency and dangerous situations, etc. The application is very broad[2].

Water quality detection drift ball, a new water quality detection device for ocean waters and other waters. This device is spherical in shape, and has a built-in micro-control system, TDS detection module, power supply regulator module, GPS module and wireless signal transmission module. The TDS module detects the concentration of total dissolved substances in the water, the GPS module measures the current position of the device, and both send data to the micro control system.

The micro control system sends water quality information and geographic location information to the network through a wireless transmission module, and users can view the water quality status of the water area where the device is located in real time through the Internet. The power supply voltage stabilization module provides different voltages required by the entire device to implement the function of detecting the current water quality. In the form of drifting, the water quality of the water area is tested by built-in sensors and control systems. The device actual picture is shown in Fig.1.

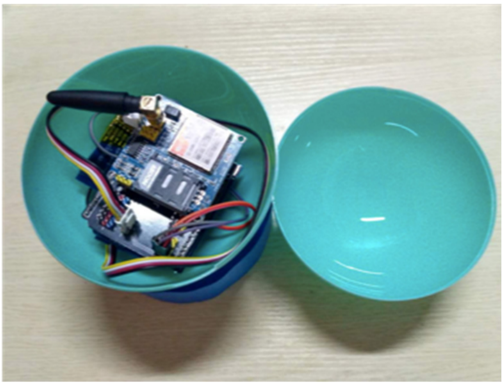


Fig.1. Device actual picture

2. The hardware structure design

2.1. Arduino Mega2560

Arduino is an open-source electronic prototyping platform enabling users to create interactive electronic objects.

ATmega2560 is a high-performance, low-power Microchip 8-bit AVR RISC-based microcontroller combines 256KB ISP flash memory, 8KB SRAM, 4KB EEPROM, 86 general purpose I/O lines, 32 general purpose working registers, real time counter, six flexible timer/counters with compare modes, PWM, 4 USARTs, byte oriented 2-wire serial interface, 16-channel 10-bit A/D converter, and a JTAG interface for on-chip debugging. The device achieves a throughput of 16 MIPS at 16 MHz and operates between 4.5-5.5 volts. The Arduino Mega2560 is shown in Fig.2.

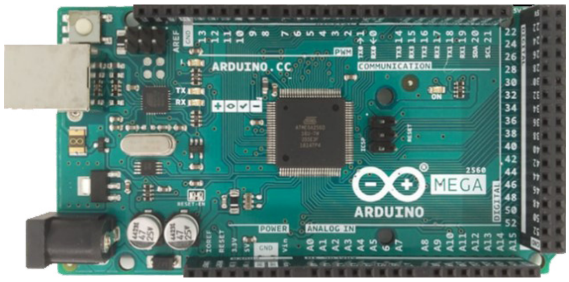


Fig.2. Arduino Mega2560

2.2. TDS Conductivity Sensor

The full name of TDS is Total Dissolved Solids, which refers to the concentration of total soluble substances in water, in milligrams per liter (mg/L). TDS mainly reflects the concentration of ions in water, and has a good corresponding relationship with the hardness and conductivity of water. TDS is commonly used to measure the concentration of pure water. When the TDS value is smaller, it means that the concentration of ions in the water is lower, the conductivity is smaller, and the purity of purified water is higher. In general, the TDS value of distilled water is zero. The size of the TDS value can also be used to reflect the concentration of harmful heavy metal ions in most water, the number of bacteria in the water, the level of organic matter, whether the nitrous acid concentration exceeds the standard, and whether there are crop residues.

Adopt the analog TDS water quality detection sensor module developed and produced by DFROBOT. It can detect the TDS value of various water quality in real time, as well as the chemical conductivity parameters of chemical water quality liquid. The module adopts the TTL serial port form, and has a certain protocol for query, and then reads the sensor module data. It can be directly connected to different development boards or serial ports, and it can also be directly connected to the serial port of the microcontroller. It has the following characteristics:

- Input voltage: DC 3.3-5.5V.
- Output signal: DC 0-2.3V.
- Working current: 3-6mA.
- TDS measurement range: 0-1000ppm.
- TDS measurement accuracy: ±10% F.S. (25 ° C).
- Size: 42 * 32mm.
- Module interface: PH2.0-3Pin.
- Electrode interface: XH2.54-2Pin.

The choice of the TDS Conductivity Sensor is shown in Fig.3.

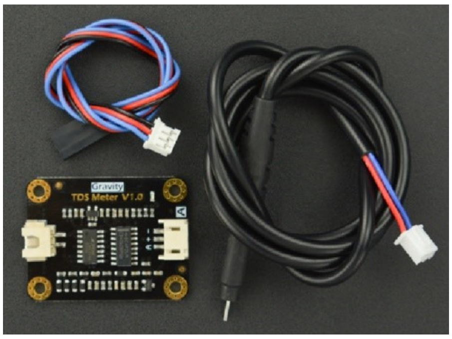


Fig.3. TDS Conductivity Sensor

2.3. pH sensor

The pH of water quality is an important indicator for measuring whether the water quality meets drinking water standards. The pH value of the water we want to drink in daily life should be between 6.5 and 8.5. The standard of pH value is based on the standard of hydrogen ions. If it is less than 7, it is significantly acidic, if it is more than 7, it is alkaline. The standard of drinking water is just neutral. However, the pH value does not change the color of the water quality. We cannot observe and distinguish directly with the naked eye, so we still need to use a sensor to measure the PH value.

The pH of a soil is also one of the most important characteristics for crop production. Agricultural soil should have the pH from 4.5 to 7. For high soil pH, the increase in waste water pH results in increase in combined pH of soil and waste water. Increased pH cause natural precipitation. As discussed earlier, for better plant growth the pH range required is from 5.5 to 7. So it is necessary to determine the soil pH first based on which the waste water need to be treated such that for a particular soil, the pH optimized waste water will ensure the suitable pH range of 5.5 to 7 for irrigation. By using fully automated method or semi-automated method it is possible to maintain the combined pH of waste water and soil in the range of 5.5 to 7[3]. The module we choose has the following characteristics:

- Heating voltage: DC 5±0.2V.
- Working current: 5-10mA.
- Detection concentration range: PH 0-14.
- Detection temperature range: 0-80℃.
- Response time: ≤5S.
- Stability time: ≤60S.
- Component power consumption: ≤0.5W.

The choice of the PH Sensor is shown in Fig.4.

Fig.4. pH sensor

2.4. Positioning module

ATK1218-BD uses SkyTra's GPS and Beidou dual-mode positioning module: S1216, external active antenna, positioning within 30 seconds. The module comes with a back battery, which can save ephemeris data[4]. It can be powered on again within half an hour after power failure, and can be repositioned in a few seconds. In terms of interface, it is connected to the outside through a 2.54mm pin with a pitch of 1 * 5Pin. It uses serial communication, and the configuration data can be saved, which is very convenient to use[5]. The choice of the Positioning module is shown in Fig.5.

Fig.5. Positioning module

2.5. GSM module

This module uses high-performance industrial-grade GSM/GPRS quad-band SIM800C module, and the working frequency band is GSM850/900/1800/1900MHz[6]. Not only can realize telephone voice, SMS (short message, multimedia message), GPRS data transmission functions, but also have DTMF decoding (can identify the other party's key), TTS (voice broadcast) and Bluetooth. This module supports 3.3V and 5V TTL serial ports, and supports 5V-18V working voltage[7]. The GSM/GPRS communication module is shown in Fig.6.

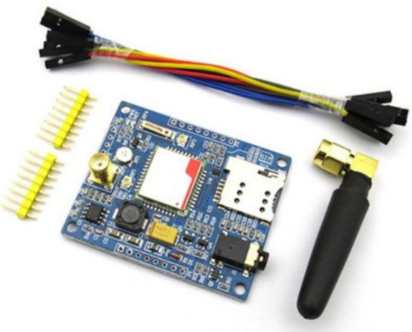


Fig.6. GSM/GPRS communication module

3. Conclusion

In the face of the current shortage of water resources and the serious water pollution environment and living conditions, most of the existing water pollution monitoring devices in the market perform separate detection and treatment[8]. This device integrates the detection and treatment of pollution in water, can effectively optimize the water quality for limited water resources, take different treatment measures for different pollution categories and pollution levels, and innovatively add a traceability link In order to form a chain structure for the governance of water pollution, the management loopholes and technical blind spots caused by divide and conquer are relatively reduced.

Acknowledgements

The research is partly supported by the Project of Tianjin Enterprise Science and Technology Commissioner to Tianjin Tianke Intelligent and Manufacture Technology Co., Ltd (19JCTPJC53700). It is also supported by the Industry-University Cooperation and Education Project (201802286009) from Ministry of Education, China.

References

1. G. Zhu, G. Zhao, Z. Zhang, X. Lu, Water quality of water source area in Taihu Lake and effect on water treatment process. *Proceedings of the 2011 International Conference on Consumer Electronics, Communications and Networks,* XianNing, 2011, pp. 3783-3786.
2. Wang Jingmeng, Guo Xiaoyu, Zhao Wenji, Meng Xiangang, Research on water environmental quality evaluation and characteristics analysis of TongHui River, *Proceedings of the 2011 International Symposium on Water Resource and Environmental Protection,* Xi'an, 2011, pp. 1066-1069.
3. R. M. Kingsta, A. S. Saumi, P. Saranya, Design and Construction of Arduino Based pH Control System for Household Waste Water Reuse. *Proceedings of the 2019 3rd International Conference on Trends in Electronics and Informatics*, Tirunelveli, India, 2019, pp. 1037-1041.
4. M. Ngui, W. Lee, Low Power Wearable Device with GPS and Indoor Positioning System. *Proceedings of the 2019 International Conference on Green and Human Information Technology*, Kuala Lumpur, Malaysia, 2019, pp. 125-127.
5. P. J. Zabinski, B. K. Gilbert, P. J. Zucarelli, D. V. Weninger, T. W. Keller, Example of a mixed-signal Global Positioning System (GPS) receiver using MCM-L packaging. *IEEE Transactions on Components, Packaging, and Manufacturing Technology: Part B*, 1995, 18(1): pp. 13-17.
6. Weiping Liu, Yanwen Liu, Ru Li, Pai Wang, Research and development of communication between PC and mobile base on embedded system and GPRS. *Proceedings of the 2011 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce*, Dengleng, 2011, pp. 4180-4183.
7. Z. Liangshui, W. Ai'hong, X. Jianyuan, Remote Monitor and Control System for Wreath Net Cabinet Based on GPRS Technology. *Proceedings of the 2006 International Conference on Power System Technology*, Chongqing, 2006, pp. 1-6.
8. Q. Wu, Y. Liang, Y. Sun, C. Zhang, P. Liu, Application of GPRS technology in water quality monitoring system. *2010 World Automation Congress,* Kobe, 2010, pp. 7-11.

Design of Film Forming Rate Measuring Instrument based on Polyurethane Material

Jichao Zhao [1], Fengzhi Dai [1,2,3]*, Fengkun Wang [1], Haokang Wen [1], Hongbo Hao [1], Qianqian Zhang [1]

[1] *Tianjin University of Science and Technology, China,*

[2] *Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China,*

[3] *Advanced Structural Integrity International Joint Research Centre, Tianjin University of Science and Technology, Tianjin, 300222, China*

*E-mail: * daifz@tust.edu.cn*
www.tust.edu.cn

Abstract

With the development and research of polyurethane materials, it has become widely used in the field of antibacterial, waterproof and moisture permeable, and medical. Therefore, this paper mainly studies on the determination of the film formation rate of polyurethane, and designs a film formation rate measuring instrument based on Atmega328P single chip microcomputer and BH1750 light intensity measurement chip. This design compares the advantages and disadvantages of the currently widely used film formation rate measuring instrument and improves it on the basis of it. By improving its multiple sensors into a high-precision main sensor, measurement requirements are reduced, and the size of the instrument is also reduced, making it easier and more portable. Moreover, it can be used not only to measure the film formation rate of polyurethane, but also to determine the film formation rate of other materials.

Keywords: Measurement instrument; Film formation rate; ATmega328P; BH1750

1. Introduction

Polyurethane microporous films are widely used in daily life as high-end or special fabrics due to their optical transparency, strong ductility, durability, ease of processing, and water permeability[1]. The film-forming rate of the casting film in different coagulation baths is different, and the film-forming rate has a significant effect on the membrane pore size and other indicators. At present, the instruments used to measure the film formation rate are manually measured by humans, which results in large errors and inaccuracy.

The new film-forming rate measurement instruments involved in this design include specific light source, light-intensity acquisition sensor, microcontroller and data processing computer. When making the film, on the glass plate will be cast film liquid scraped into a film, into the container's solidification bath. The specific light source directly above the polymer cast film liquid, light intensity sensor is located directly below the container, the intensity of the projected light measured by the light sensor, light intensity sensor connected to the acquisition controller, the acquisition controller connected to the data processing computer[2]. The device by measuring the change of light intensity through the membrane during phase separation, the membrane curve is generated. The test instrument is compact, small in size, easy to use and more suitable for a variety of measurement environments.

2. The hardware structure design

2.1. Controller selection

Arduino is a convenient, flexible and easy-to-use open source hardware product with rich interfaces, including digital I/O ports, analog I/O ports, and support for SPI, IIC, UART serial communication. It can sense the environment by connecting sensors and feed power and affect environment by controlling lights, motors and other equipment[3]. It has no complex micro-computer underlying code, no difficult assembly, just simple and practical functions. In addition, it has a simple integrated development environment, great freedom and great scalability. Standardized interface models provide a solid foundation for their sustainable development. Developers have developed simple features and many application libraries, so we don't need to operate registers directly, so people who don't have a good MCU knowledge can use Arduino to do what they want.

2.2. Data acquisition sensor selection

This device using BH1750FVI chip, power supply voltage 3-5V, light intensity range 0-65535lx, sensor built-in 16bit AD converter, direct digital output, omit complex calculations, omitted calibration. Without distinguishing between ambient light sources and close to the spectral characteristics of visual sensitivity, a wide range of brightness can be measured with high precision of 1 lux. Using the standard NXP IIC communication protocol, the module contains communication level conversion and can be directly connected to the 5V microcontroller IO. The GY-30 module is shown in Fig.1.

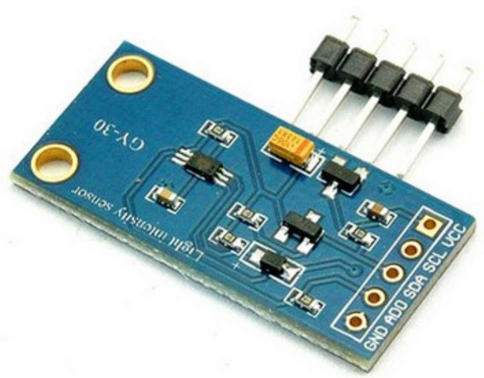


Fig.1. WIFI module

2.3. Development language

The C language is a process-oriented computer programming language that differs from Java, C++ and other object-oriented programming languages. The design goal of the C language is primarily to provide a programming language that compiles in a simple way, handles low-level memory, generates only a small amount of machine code and runs without the need for any support from the operating environment. C is faster than assembly language in describing problems, reducing workload, improving readability, easy debugging, modification and porting, and the quality of the code is comparable to the assembly language. C language is usually only 10%-20% less efficient than the target program generated by assembly language code. Therefore, the system is written in C language.

C# is a safe, stable, simple, and elegant object-oriented programming language, derived from C and C++. It inherits the power of C and C ++, while removing some complex functions (for example, no macros, multiple inheritance is not allowed). C# combines the simple visual operation of VB with the high efficiency of C++. With its powerful operation ability, elegant syntax style, innovative language functions and convenient support for component programming, it has become the preferred language for NET development. Therefore, the upper computer used in this system is developed by C# language.

3. Hardware circuit design

The hardware circuit design of the system is mainly composed of ATmega328P microcontroller control module, GY-30 light intensity acquisition module and key input module. GY-30 light intensity module is used to collect ambient light intensity information. The ATmega328P microcontroller control module is responsible for analyzing and processing the signals collected by the sensors, and then transmitting the data to the computer through a serial data transmission line.

3.1. ATmega328P microcontroller control module

This design uses the ATmega328P as the main control chip and the required drive chip operation of the crystal-vibration circuit, as well as the reset circuit[4]. The single-chip control module is mainly responsible for receiving the acquisition signal transmitted by the outside world, analyzing and processing the collected data, and sending control signals to the subsequent circuits to realize the control of each module. The ATmega328P chip is shown in Fig.2.

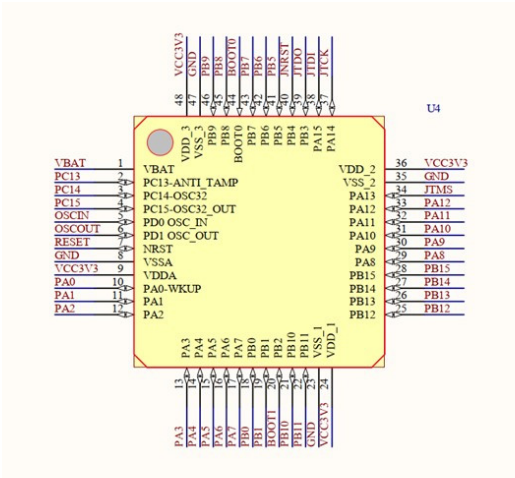


Fig.2. ATmega328P chip

3.2. Power circuit

In this design, we choose a RT9193 chip to power adapter, the main electric 12V through the transformer in the adapter to buck the pressure, and then through the bridge circuit to convert 12V into DC 5V output[5]. 5V voltage to the microcontroller, liquid crystal display module, buzzer module and other circuits to power.

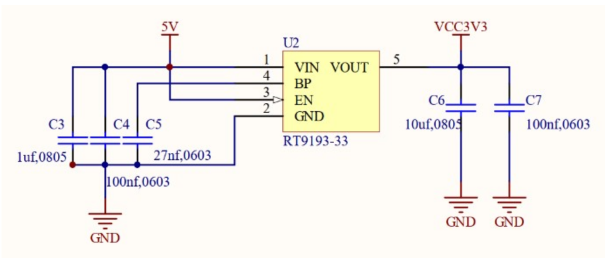


Fig.3. RT9193 regulation circuit

3.3. Data acquisition circuit

In this design, GY-30 light intensity collector with bh1750fvi chip is used, which can measure the light intensity in the range of 0-65535 lux and has a large measurement range. In the measurement process, it does not distinguish the environmental light source, and is close to the spectral characteristics of visual sensitivity. It can measure a wide range of brightness with high accuracy of 1 lux, and the response is more sensitive. And through the optimization design of the program algorithm, it can achieve the function of self-timing of the instrument, greatly reduce the error caused by manual measurement time, and make the data more accurate.

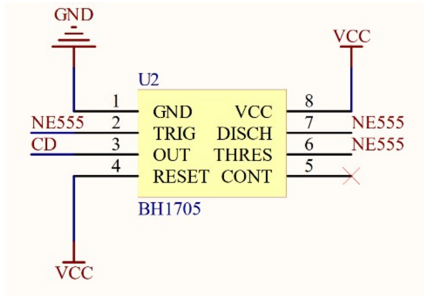


Fig.4. BH1705 chip

4. Experimental report

4.1. Experimental process

After the film formation rate tester and the computer are connected via a USB cable. Open the computer-side test software, first select the communication port connected to the instrument to connect the computer and the instrument, and then select the acquisition speed required for this test, the rate range is 0-99 times per second. Then place the sample to be tested in the corresponding position, click "Start Measurement", the tester starts to work, as the test time increases, the film formation rate curve gradually appears, and when the curve is stable, click "Stop Measurement". Fill in the environmental factors of this test on one side, choose whether to remove the influence of ambient light according to the test needs, and finally click "Save Image" to generate a test report.

4.2. Experimental result

When the liquid film of the casting film liquid contacts the coagulation bath, the solvent on the surface and inside of the liquid film is extracted by the coagulation bath. Due to the precipitation of the liquid film, the liquid film gradually changes from transparent to white, so the intensity of the light transmitted through the liquid film decreases accordingly. The start time of phase separation is the time when the transmission intensity starts to decrease. When the transmission intensity drops to a certain stable value, phase separation is considered complete, and the voltage value drops to a certain stable value. Over time, it gradually became stable.

This phase separation film formation rate tester is used to measure the change in light intensity transmitted through the film during the phase separation, that is, the film formation curve. This curve reflects the initial transmittance and end transmittance of the solidified liquid film and the time interval used, and the collection time can be adjusted appropriately according to the extraction rate of different coagulation baths. Import the data into an Excel spreadsheet and set it as the relationship curve between time and light intensity change. For example, when polyurethane is formed into a film, distilled water and methanol are used as coagulation baths, as shown in Figures 5 and 6.

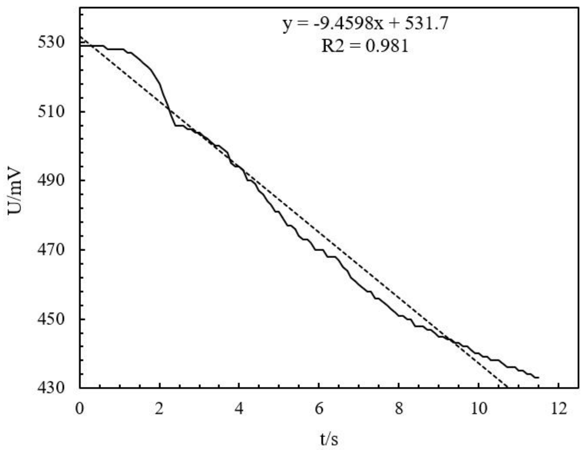


Fig.5. Distilled water bath curve

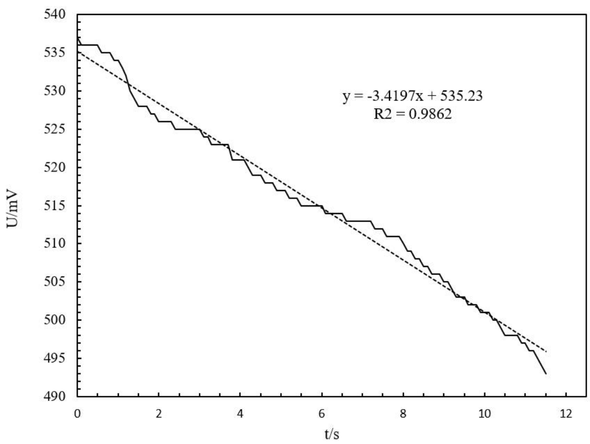


Fig.6. Methanol water bath curve

The absolute value of the slope can reflect the speed of the film formation rate of the process, that is, the greater the increase in the light intensity value during the same time period in the light projection curve, the faster the film formation rate of the system, the more solidified into The shorter the time taken by the film.

Acknowledgements

The research is partly supported by the Research Fund for the Doctoral Program of Higher Education of China (20131208110005). It is also supported by the Science & Technology Fund Planning Project of Tianjin Higher Schools (20120831, 20140710-1400020005).

References

1. Ma Ye, Xin Lian, Tan Huaping, et al., Chitosan membrane dressings toughened by glycerol to load antibacterial drugs for wound healing, *Materials science & engineering, C, Materials for biological applications*,2017, pp.81.
2. Eykens L, De Sitter K, Dotremont C, et al., Membrane synthesis for membrane distillation A review, *Separation and Purification Technology*, 2017, 182: pp.36-51.
3. M. Caruso, R. Miceli, P. Romano, G. Schettino, C. Spataro, F. Viola, A low-cost real-time monitoring system for PV plants based on ATmega 328P-PU microcontroller. *Proceedings of the 2015 IEEE International Telecommunications Energy Conference*, Osaka, 2015, pp. 1-5.
4. B. Lu, X. Wu, H. Figueroa, A. Monti, A Low-Cost Real-Time Hardware-in-the-Loop Testing Approach of Power Electronics Controls, *IEEE Transactions on Industrial Electronics*, 2007, 54(2): pp. 919-931.
5. Y. Jang, S. Cho, B. Kim, J. Sim, H. Park, A low-power LDO circuit with a fast load regulation, *Proceedings of the 2016 IEEE Asia Pacific Conference on Circuits and Systems*, Jeju, 2016, pp. 47-49.

Research on the Intelligent Aircraft Design based on STM32

Hongbo Hao [1], Fengzhi Dai [1, 2]*, Haokang Wen [1], Jichao Zhao [1], Qianqian Zhang [1]

[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 STM32f767 igt6 (M7 series) as the core, the intelligent aircraft designed this time will finally realize the intelligence of the aircraft by using the optimized attitude algorithm (cascade PID and single-stage PID share the optimized control), omni-directional ultrasonic radar detection barrier collision prevention technology, long-distance wireless transmission technology (to realize the timely transmission of images recognition), navigation technology (Beidou and GPS double positioning to realize more accurate positioning), voice recognition technology, man-machine interaction technology, and wireless local area network technology.

Key words: Aircraft, man-machine interaction, wireless transmission, PID

1. Introduction

1.1. Research background

In the inspection of the expressway, the UAV is used to conduct the inspection of the high-speed section. It is possible to carry out the first-time processing of various emergencies, and truly grasp the road conditions of the expressway in real time. In the field of forest fire prevention, various line electronic devices are not easy to install, so the application of the drone can immediately monitor the situation of the forest[1].

Another advantage of the four-rotor unmanned aerial vehicle is its small size. Unmanned aircraft can achieve autonomous flight in the air and perform certain tasks. Compared with ordinary aircraft, it has a simple structure, low cost, and it is easy to maintain and manufacture. It can be used for real-time battlefield investigation, target positioning, and unit tracking in military rain. It can be used for civilian disasters and technology to detect major disasters. It can find lost personnel rescue work and can also carry a variety of scientific equipment for scientific experiments.

1.2. This design innovation content

The designed aircraft, voice interaction, image recognition, and stereo radar design are the highlights. The main control core adopts STM32F767 as the first aircraft designed for M7 in China. M7 processing speed, excellent running effect and stable performance. It can carry a variety of smart sensors, and have multiple peripheral interfaces to support IIC, USART communication external device mounting and large-size screen display.

1.3. The main work of the paper

The control algorithms, software programs, and hardware circuits of large quadcopters are developed and designed.

In this paper, the communication protocols used by various quad-rotor aircrafts are analyzed in depth and optimized for program design.

According to the current mainstream algorithm of quadrotor, the attitude control is optimized.

This design analyzes the fusion algorithm that uses multiple types of sensors to achieve flight movement control.

2. Principle and Modeling of Four-rotor Aircraft

2.1. Four-rotor dynamics analysis

First, the four-rotor aircraft has a cross shape with four arms. In flight, there are generally two states, X-shaped or cross-shaped. The aircraft is equipped with an imaging device. The principle is to realize the action of the aircraft through the angle between the aircraft and the horizontal plane. First of all, the quadrotor has a cross shape and has four arms.

Generally, there are two states when flying, X-type or cross-type X. The angle between the arm and the forward direction is 45 degrees, and the adjacent arms are perpendicular for each other. The arms coincide with the forward direction, and two adjacent vertical arms are crossing[2].

2.2. Four-HelicalMathematical Modeling

The four-helical aircraft is a nonlinear, multiva-riable, highly coupled, underactuated system. The motion state of the aircraft has 6 degrees of freedom, and there are only 4 inputs. The idealized model is assumed as follows[3].

2.2.1. DC motor modeling

The phase voltage of each winding of the three-phase brushless DC motor is composed of the winding induced potential and the copper wire's own resistance[4]. Therefore, the following voltage balance equation can be used for each voltage.

$$U = IR + L\frac{\mathrm{di}}{\mathrm{dt}} + E \tag{1}$$

Then the three-phase DC state equation is

$$\begin{bmatrix} U_a \\ U_b \\ U_c \end{bmatrix} = \begin{bmatrix} R_a & 0 & 0 \\ 0 & R_b & 0 \\ 0 & 0 & R_c \end{bmatrix} \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix} + \begin{bmatrix} L_{aa} & L_{ab} & L_{ac} \\ L_{ba} & L_{bb} & L_{bc} \\ L_{ca} & L_{cb} & L_{cc} \end{bmatrix} \begin{bmatrix} \dot{i}_a \\ \dot{i}_b \\ \dot{i}_c \end{bmatrix} + \begin{bmatrix} e_a \\ e_b \\ e_c \end{bmatrix} \tag{2}$$

The equivalent circuit diagram of the three phase brushless motor can be obtained from the voltage equation, as shown in the Fig.1 below.

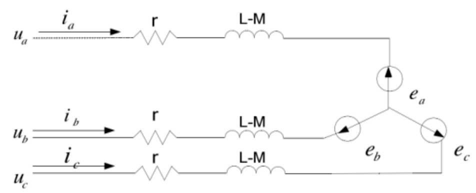


Fig.1 Equivalent circuit diagram of three-phase brushless motor

2.2.2. Establishment of the torque equation

The current magnitude and electromagnetic tor -que of a phase brushless motor are proportional to the flux.

$$P = T_e \omega = \sum_x i_x e_x \tag{3}$$

The equation of motion of the motor is as follows:

$$T_e = T_L + J\frac{d\omega}{dt} \tag{4}$$

2.2.3. Establishment of the aircraft model

The four-rotor has six flight states: rising, descending, forward, backward, leftward, and rightward. In order to stabilize the hover, it is necessary to detect the angle between the four motors and the horizontal plane. The aircraft uses the quaternion and Euler angles in 3D stereology to represent the angle of rotation. The calculation formula uses a 3D Cartesian coordinate system, as shown in the Fig.2 below.

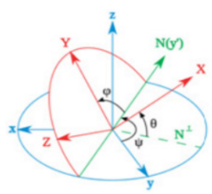


Fig.2 3D Cartesian coordinate system Definition of quaternion

$$\mathrm{q} = [\mathrm{w} \;\; \mathrm{x} \;\; \mathrm{y} \;\; \mathrm{z}]^T \tag{5}$$

$$|q|^2 = w^2 + x^2 + y^2 + z^2 = 1 \tag{6}$$

Conversion of Euler angles to quaternions:

$$q = \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \cos(\phi/2)\cos(\theta/2)\cos(\psi/2) + \sin(\phi/2)\sin(\theta/2)\sin(\psi/2) \\ \sin(\phi/2)\cos(\theta/2)\cos(\psi/2) - \cos(\phi/2)\sin(\theta/2)\sin(\psi/2) \\ \cos(\phi/2)\sin(\theta/2)\cos(\psi/2) + \sin(\phi/2)\cos(\theta/2)\sin(\psi/2) \\ \cos(\phi/2)\cos(\theta/2)\sin(\psi/2) + \sin(\phi/2)\sin(\theta/2)\cos(\psi/2) \end{bmatrix} \tag{7}$$

Quaternion is converted to Euler angle:

$$\begin{bmatrix} \varphi \\ \theta \\ \psi \end{bmatrix} = \begin{bmatrix} \mathrm{arc\ tan}(\dfrac{2(wx + yz)}{1 - 2(x^2 + y^2)}) \\ \arcsin(2(wy - zy)) \\ \arctan(\dfrac{2(wz + xy)}{1 - 2(y^2 + z^2)}) \end{bmatrix} \tag{8}$$

3. Aircraft scheme design

3.1. Overall scheme design

The aircraft modules used in this design include M7 series control cores, attitude sensors, intelligent interactive equipment, wireless remote control equipment, obstacle avoidance equipment, positioning equipment, and upper computer display equipment[4].

3.2. Equipment selection

This article frame (F450), motor (B2212 brushless motor), ESC (SimonK ESC), NRF24L01 wireless module, ESP8266 WIFI module, voice recognition module LD3320A, ultrasonic module HC-SR04, OV5640 camera module, GPS Beidou dual positioning Module S1216, attitude sensor MPU6050 AK8975.

4. Hardware scheme design

4.1. Overall design

After the selection is completed, the overall circuit design scheme is born accordingly. Each module we use has its own communication method. Each module occupies different internal resources of the core main control board according to different design schemes.

4.2. STM32F767 minimum system design

This system includes a reset circuit. The hardware reset button for testing the system; the Jlink download interface is used to download programs and online testing. The SDRAM external expansion circuit uses the W9825G6KH chip to expand the 256K memory capacity to 32M, which enhances the demand for memory of each program. The FLASH chip used in the FLASH external circuit is MT29F4G08, which adds 512M of storage space to the main control.

It can be used to store pictures and key information.

4.3. Power supply design

This power supply contains 5V to 3.3V regulated power supply, which mainly supplies power to 3.3V modules and core STM32F767 chips. Expand the power supply 5V-3V, in order to prevent the voltage from being too high, the voltage regulator is used to stabilize the voltage.

4.4. Module Interface Circuit Design

As each module transmits information through digital signals, the common interfaces are IIC, UART, and motor interface, voice interface, camera interface, etc.

5. Software scheme design

5.1. Single-level PID control algorithm

The PID algorithm of input and output response is controlled, and the parameters are adjusted to make the input and output response reach an ideal state[5]. The PID is divided into positional and incremental, and the positional PID is related to the error of the whole past, and the integral link is used. The incremental PID is only related to the error of the first two beats, and the output is the control increment.

5.2 . Cascade PID control algorithm

Cascade PID is divided into inner ring and outer ring. The input and feedback are angle data, but the inner ring input feedback is the angular velocity data, which is quickly corrected according to the expected speed. The outer ring is input to the inner ring according to the angular deviation, and the inner ring determines how much the speed should be operated according to the angular deviation.

Generally, we only use PI control in the outer ring. Proportion corrects the aircraft from the deviation angle to the desired angle; Integration eliminates the angular static difference. In the inner ring, we adopt PID control. Proportion corrects the yaw angular velocity of the aircraft to the desired angular velocity. Integration eliminates the static acceleration difference[6]. Different coefficient suppresses the system motion.

6. Testing and conclusion

The design satisfactorily completed the design requirements, realized the stable flight of the aircraft. It detected the environmental parameters and realized the function of human-computer interaction. It provided more convenient control through the remote control of the mobile phone, and supported real-time one-megapixel image transmission. It can record or take pictures and support GPS/Beidou global positioning and synchronize positioning in real time in Google Earth.

Google Earth is a simulated earth that contains information about various roads and buildings on the entire earth. A 3D version of the earth model, we will use the NMEA protocol format positioning information to configure the corresponding serial port with Google Earth . Then it can show Google Earth the location of its aircraft. The aircraft positioning information is shown in Fig.3 and The Google Earth interface is shown in Fig.4.

Fig.3 Aircraft position

Fig.4 Google earth

Acknowledgements

The research is partly supported by the Project of Tianjin Enterprise Science and Technology Commissioner to Tianjin Tianke Intelligent and Manufacture Technology Co., Ltd (19JCTPJC53700). It is also supported by the Industry-University Cooperation and Education Project (201802286009) from Ministry of Education, China.

References

1. OV5640 Auto Focus Camera Module Application Notes R2.15, Beijing: Electronic Industry Press, 2015.
2. Guiyun Tian, F*oundation and Application of Microcontrolle*, Beijing Higher Education Press, 2004.
3. Lin Li, Zhang Junliang, *Principle and Application of Single Chip Microcomputer*, Beijing: Electronic Industry Press, 2015, (6): pp.54-60.
4. V. Yu. Teplov, A. V. Anisimov, Thermostatting System Using a Single-Chip Microcomputer and Thermoelectric Modules Based on the Peltier Effect, IEEE AQTR, 2002(4): pp. 45-50.
5. Mian A Wang Daobo, Nonlinear flight control strategy for an underactuated quadrotor aerial robot. *IEEE AQTR*, 2008, 1: pp.8-16.
6. Li Wang, Huailin Zhao, Yaoyao Li, Research on the multi-scale network crowd density estimation algorithm based on attention mechanism, *Proceedings of the 2019 International Conference on Intelligent Informatics and BioMedical Science*, Shanghai, China, Nov.21-24, 2019, pp.272-278.

Research and Design of Gain Controllable System in RF Receiver

Haokang Wen [1], Fengzhi Dai [1,2,3]*, Jichao Zhao [1], Hongbo Hao [1], Hong Niu [1], Qianqian Zhang [1]

[1] Tianjin University of Science and Technology, China,

[2] Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China,

[3] Advanced Structural Integrity International Joint Research Centre, Tianjin University of Science and Technology, Tianjin, 300222, China

*E-mail: * daifz@tust.edu.cn*

www.tust.edu.cn

Abstract

This paper studies and designs a gain controllable radio frequency amplifier system. As the hardware core of the whole system, the variable gain amplifier realizes the data function of the amplifier combined with the single chip microcomputer. After adjusting the gain of RF amplifier, the input signal can be amplified or attenuated, and the ideal output signal can be obtained. Using ADS simulation, the gain controllable RF amplifier system designed in this paper has better control effect, meets higher index, and has stable performance. The experiment shows that the gain controllable RF amplifier system has certain practicability.

Keywords: radio frequency amplifier, gain controllable, single chip microcomputer

1. Introduction

1.1. Research background and significance

As a transmission module in a communication system, RF amplifiers are now widely used in various communication systems, such as various types of communication equipment and digital image transmission. Gain control technology is an essential content inside these systems.

In recent years, more low-noise amplifier designs and applications have appeared, such as an optical amplifier that has low noise characteristics for a wide range of input signal electric power エラー! 参照元が見つかりません。, and low-power low-noise amplifiers for portable Electrocardiogram recording system applications [2], but different for other types of design schemes, this topic is designed a stable and mid-low frequency RF signal gain controllable amplification system. The single-chip microcomputer implements the numerical control function on this system. Theoretically, it can meet higher indicators and can better control gain. It can be applied to a variety of occasions and has a wide range of practical significance.

1.2. Main content of this paper

In the hardware part, full reference is made to various types of RF amplifier modules to build a gain-controllable amplifier system that meets the requirements. The ADS software is used to simulate the feasibility of the scheme. In the software part, the AT89C52 is used as the main control microcontroller to implement the key control function for the entire variable gain system. At the same time, a display module is added to display the input signal, output signal and gain of the amplifier in the receiver in real time.

2. RF Receiver Overview

The common basic architecture of the RF receiver is shown in Fig.1. It is mainly composed of five parts: filter, low-noise amplifier, mixer, band-pass filter, and analog-to-digital conversion.

The LNA low-noise amplifier module in Fig. 1 has multiple design schemes. According to the design of the controllable gain system in this article, this LNA can be replaced with VGA.

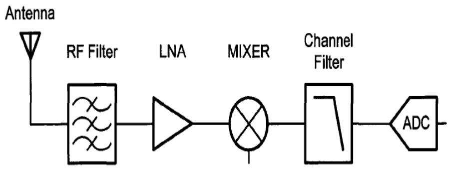


Fig.1 RF receiver basic structure

3. Controllable gain system design

Fig. 2 shows the block diagram of the system designed in this paper, which mainly includes detectors, variable gain amplifiers, AD converters, DA converters, and microcontrollers.

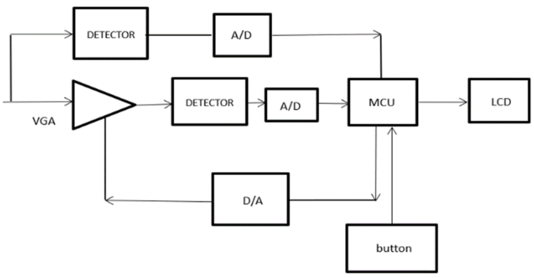


Fig.2 System structure block diagram

3.1. Detector

The AD8361 selected in this paper is an average response power detector, suitable for high-frequency receivers up to 2.5GHz. This device can convert a complex modulated RF signal up to 2.5 GHz into a dc voltage representing the rms level of the RF signal [3]. The circuit design of the module is shown in the Fig.3.

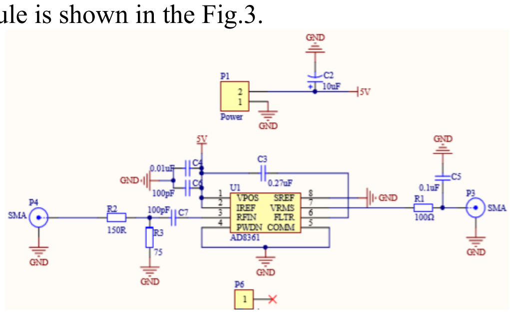


Fig.3 AD8361 module circuit diagram

3.2. Variable gain amplifier

In this subject, because the goal is to achieve the amplification of low-frequency RF signals, the amplifier module with AD8367 chip as the core is selected in this subject. This module is a variable gain intermediate frequency amplifier. The module is mainly composed of two parts, one is a 9th order resistance attenuation network, and the other is a fixed gain amplifier. The AD8367 is an example of a VGA that uses variable attenuation followed by a post-gain amplifier [4]. Providing the module with DC voltage enables it to work in VGA mode, which is the working mode used in this project. The module's peripheral circuit is designed with a shielded shell and an anti-reverse protection diode, which has a good protection function for the module. As shown in Fig.4.

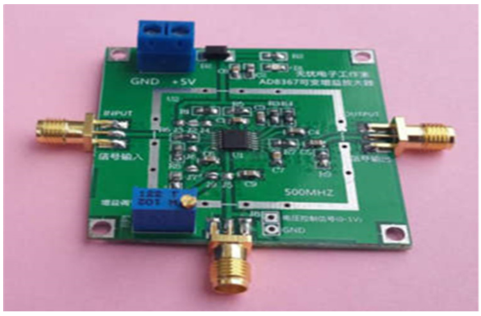


Fig.4 Variable gain amplifier module

3.3. ADS simulation design

ADS, Advanced Design System, is a commonly used RF circuit simulation software. Using this software, the variable gain amplification system of this subject is simulated. The simulation is shown in Fig.5.

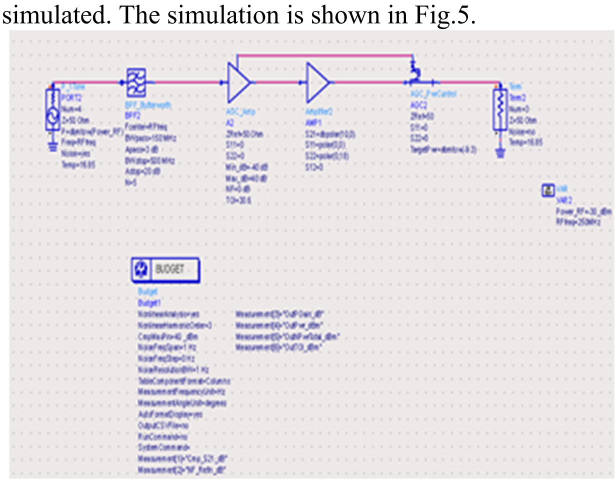


Fig.5 Simulation circuit diagram

In this model, by setting different output powers, the simulation of the variable gain system of this subject is completed. Because the working frequency of the variable gain amplifier is 0-500MHz, in this simulation, the input RF signal frequency is set to 250MHz and the input power is -30dBm. At the same time, when the input signal is 250MHz obtained in the actual test, the output RF signal power of the system is -9.3dBm.

3.4. AD module

The analog-to-digital conversion module in this subject is the PCF8591 module. PCF8591 is an 8-bit data acquisition device [5]. All data in the module is serially input or output through the I2C. Through the setting of the single-chip microcomputer, data collection and conversion functions can be realized.

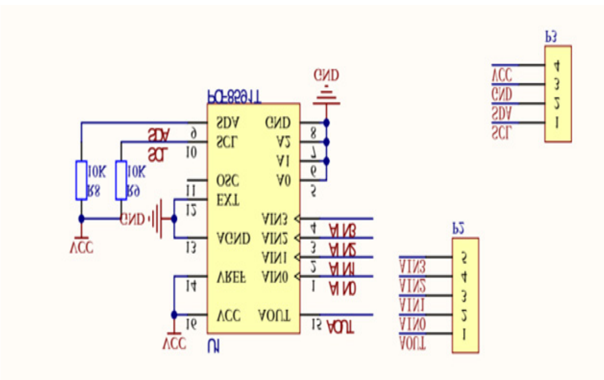


Fig.6 PCF8591 schematic

3.5. DA module

The system hardware designed in this paper adopts a modular design, taking the 89C52 single-chip microcomputer as the core of the system, controlling the DA conversion circuit, AD conversion circuit, and key control circuit to form the entire control system. The function of DA is to output the analog voltage signal through the single chip microcomputer, and send it back to the variable gain amplifier VGA module. Through the control of the single-chip microcomputer, the outputs are analog voltage value, so as to control the variable gain amplifier and realize different gains.

DAC0832 is an electric current exports type 8 bits D/A converter, adopting to reversing T type resistor network [6]. Through this module, it is possible to output different point flows with three keys and output corresponding analog voltage values through the load resistor. This voltage signal is supplied to the variable gain amplifier, which can make the variable gain amplifier module AD8367 have different gain, and finally realize the control function of the microcontroller to the amplifier. Combining key control with DA analog-to-digital conversion [7], through actual tests, the actual system requirements can be achieved, the test results are good, and the control performance is stable.

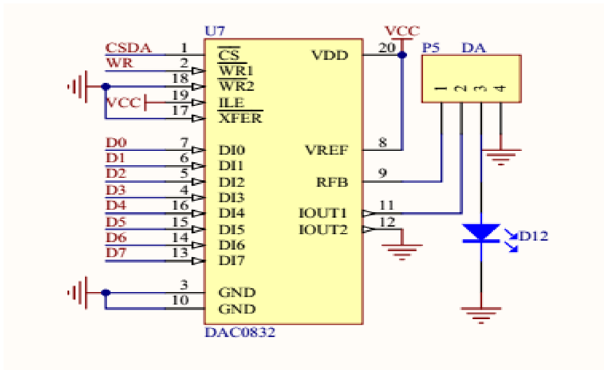


Fig.7 DAC0832 schematic diagram

4. System test

After building the entire system and connecting the modules, the actual tests were performed. The experimental effect is obvious, and obvious waveform changes can be obtained on the spectrum analyzer. Next, list some data, objects and graphics in actual tests.

In this subject, first set the RF input signal: 250MHz, -30dBm.

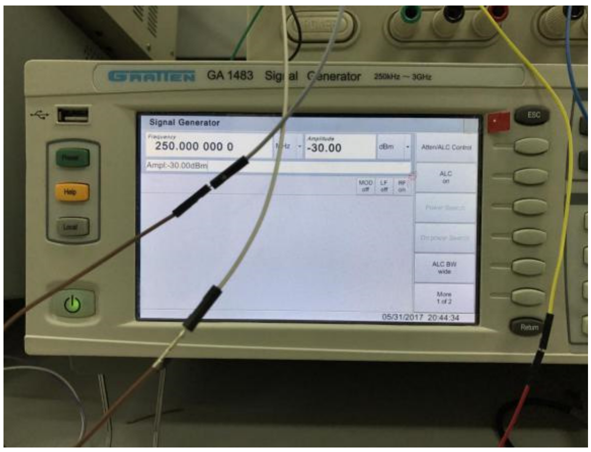


Fig.8 Input signal

As shown in Fig.9, the output signal passing through the amplifier is: 250MHz with an amplitude of -9.28dBm.

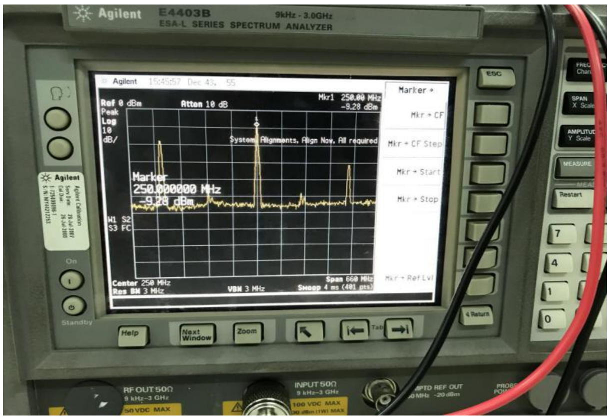


Fig.9 Output signal

As shown in Fig.10, The liquid crystal display of the single-chip microcomputer: IN: 026; OUT: 060; DEV: 00043.

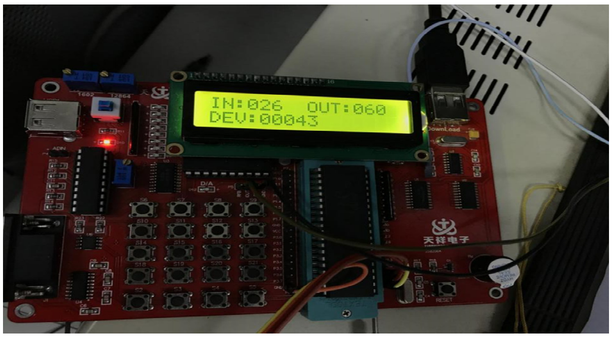


Fig.10 LCD display

In the liquid crystal display section, IN is the input RF signal source (V) of the system, and OUT is the amplified output value (V) of the system. DEV is the gain, which is the ratio of the OUT value to the IN value. The displayed value is 100 times the actual ratio.

Under the test conditions, the input signal is 250MHz with an amplitude of -30dBm and the output signal is 250MHz with an amplitude of -9.28dBm. It can be deduced that the actual gain is 20.72dB and the actual magnification is about 11 times. The single chip microcomputer collects and converts the value on the LCD after being collected by AD and converted into a digital quantity. The ratio of the two numbers is 0.43, and the gain of the amplifier is 2.3 times.

5. Conclusion

The full text starts from the design of the variable gain amplifier and aims to achieve the gain controllability of the amplifier, and realizes the design of a variable gain amplifier system at the front end of the RF receiver. The simulation of the amplifier was completed, and the single-chip microcomputer controlled the gain of the amplifier accurately and stably. However, there were some discrepancies in the liquid crystal display and gain calculation of the system's analog quantity transmitted to the single chip microcomputer.

Acknowledgements

The research is partly supported by the Project of Tianjin Enterprise Science and Technology Commissioner to Tianjin Tianke Intelligent and Manufacture Technology Co., Ltd (19JCTPJC53700). It is also supported by the Industry-University Cooperation and Education Project (201802286009) from Ministry of Education, China.

References

1. Yamamoto, Shu, et al., Gain-controllable optical amplifier and optical gain-control method, U.S. Patent No. 5,710,660. 20 Jan. 1998.
2. Xiao Yang, Qi Cheng, Li-fei Lin, Wei-wei Huang, Chao-dong Ling, Design of low power low noise amplifier for portable electrocardiogram recording system applications, *Proceedings of the 2011 IEEE International Conference on Anti-Counterfeiting, Security and Identification*, 2011, pp. 89-92.
3. Pilotte, M., Operation of RF Detector Products at Low Frequency. *Analog Devices App Note* AN-691, 2005.
4. Newman E J., X-Amp,™ A New 45-dB, 500-MHz Variable-Gain Amplifier (VGA) Simplifies Adaptive Receiver Designs. *Analog Dialog*, 2002, 36(1): pp. 1-3.
5. Gay, W., *PCF8591 ADC*, In Custom Raspberry Pi Interfaces (pp. 67-90). Apress, Berkeley, CA, 2017.
6. Xu, Y. H., Zhang, D. T., Research on One Kind of Intellectualization High Accuracy Digital Adjustment Direct Source. *Applied Mechanics and Materials*, 2014, 599: pp. 608-611.
7. Huailin Zhao, Zhen Nie, Research of the ARM-based Multi-robot aggregation Control, *Proceedings of the 2019 International Conference on Artificial Life and Robotics*, Beppu, Japan, Jan. 2019, pp.352-357.

Design of Packaging Bottle Recycling System based on Internet of Things

Yujie Yan [1,3], Fengzhi Dai [1,2,3]*, Shiwei Wu [1], Yuanyuan Xi [1], Huanhuan Li [1]

[1] *College of Electronic Information and Automation,*
Tianjin University of Science and Technology, China;

[2] *Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China*

[3] *Advanced Structural Integrity International Joint Research Centre, Tianjin University of Science and Technology*

*E-mail: * daifz@tust.edu.cn*
www.tust.edu.cn

Abstract

At present, the problems of environmental pollution and resource waste in China need to be solved. The existing packaging bottle recycling system in the market has the problems of high cost and maintenance cost, imperfect function and inflexible rebate mode. Based on the technology of intelligent detection, wireless communication and software development, this paper designs a comprehensive system of packaging bottle classification and recycling, which includes intelligent classification and recycling machine, Mini Program and Web management platform. The system can automatically identify and classify metal bottles and plastic bottles, and can give users integrals, which has certain environmental protection significance.

Keywords: classification, Mini Program, Web, integrals

1. Introduction

A large number of domestic recyclable resources are put into the garbage system, resulting in the waste of resources. According to statistics in 2017, discarded beverage bottles accounted for up to 30 to 40 percent of household waste. Waste beverage bottles have a high recycling value. It can be seen that the waste beverage bottles that seriously pollute the environment are actually valuable resources in the wrong place, and effective measures must be taken to recycle them.

At present, due to the imperfect recycling system, the recycling rate of waste beverage bottles in China is far lower than that in other developed countries. The beverage bottle recycling system in the market can realize the round-the-clock recycling function, which improves the resource recovery to a certain extent, and also brings economic benefits to investors. But its cost and maintenance cost is high, the function is not perfect, the rebate mode is not flexible enough, has not been widely promoted in the market. Therefore, it is necessary for the society and the market to research and develop a set of intelligent packaging bottle classification and recycling system which is easy to popularize.

2. Overall design scheme of the system

Combined with practical application scenarios and market needs, this paper uses intelligent detection, wireless communication, software development and other technologies to develop a system of automatic beverage bottle classification and recycling, which is based on the Internet of things, including intelligent classification and recycling machine, small WeChat program and Web management platform.

The intelligent sorting and recycling machine can identify and classify the plastic and metal bottles automatically, avoiding the secondary pollution caused by the mixing mode and reducing the cost of later manual sorting. The recovery machine is also equipped with automatic counting, voice prompt, overflow alarm and the detection of ambient temperature and humidity and other functions, and through wireless communication to upload all data information to the server in real time.

User can use the WeChat applet to scan the code of the equipment on the recycling machine and then place the beverage bottles. The system will return the corresponding points according to the material and number of the beverage bottles. The method of using small program to get points is more flexible than the traditional method of swiping IC card to get points, which effectively improves the user experience. The web management platform can monitor the running state of the recovery machine remotely, and the administrator can receive the alarm information details in time and take timely measures. The overall architecture of the packaging bottle recycling system based on the Internet of things is shown in Fig 1.

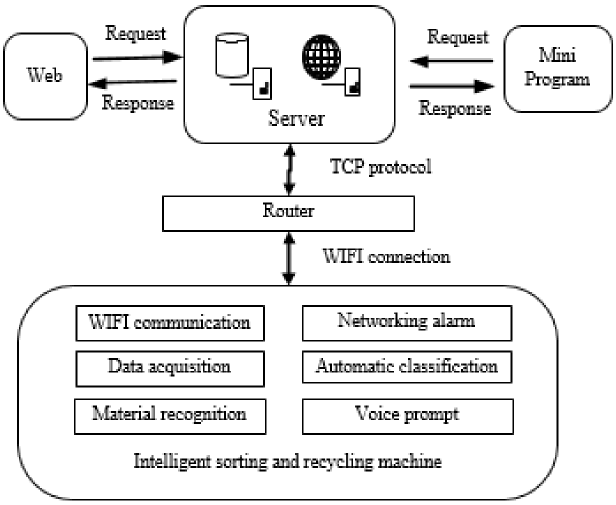


Fig.1. The design of the mechanical structure model

3. Design of the control system of the reclaimer

In this system, STM32F103 single-chip microcomputer is selected as the core controller, and the control system functions are realized by combining the peripheral modules. The hardware control system mainly includes: core controller, power module, WiFi communication module, each sensor module and other peripheral modules. The following describes the parts of the hardware control system.

3.1. Main control chip

In this system, STM32f103vet6 is selected as the core controller of the hardware control system, which is mainly used to realize the functions of data acquisition and analysis, motion control of executive components, etc. The processor of STM32f103vet6 is Arm Cortex-M3 high-performance kernel, with the maximum working frequency of 72mhz, flash memory of 512KB, SRAM of 64KB, and packaging form of lqfp64. Its peripheral configuration is powerful, supporting SPI interface, I2C interface, USB interface, USART interface and other communication interfaces, as well as ADC, timer and other peripheral devices [1].

3.2. Voltage stabilizing module

The power supply module is responsible for power supply for each module of the system. The power supply voltage required by each module of the system is different: the power supply voltage required by the actuator actuator is 12V, the power supply voltage required by MCU and Esp8266 WiFi module is 3.3V, and the power supply voltage required by other peripheral modules is 5V. Therefore, the main power supply of the system is 12V DC power supply, and the voltage stabilizing module adopts Ams1117 series voltage stabilizing chip to stabilize the voltage, so as to meet all power supply requirements.

Ams1117 voltage regulator chip belongs to linear voltage regulator chip (LDO), and the pressure difference between input and output shall not be less than 1.3V. This series of chips have two versions: adjustable voltage stabilizing output and fixed voltage stabilizing output. This system uses Ams1117-5.0 and Ams1117-3.3 chip to stabilize voltage, and its package form is SOT-223 chip package. As ripple may occur during power supply and the peak voltage may cause the main control chip to burn down, it is necessary to filter the output voltage, and the control voltage amplitude is within 5%. When using Ams1117 voltage stabilizing chip, a capacitor shall be connected in parallel at the input and output terminals respectively for filtering. The voltage stabilizing circuit of Ams1117-5.0 and Ams1117-3.3 is shown in Fig 2.

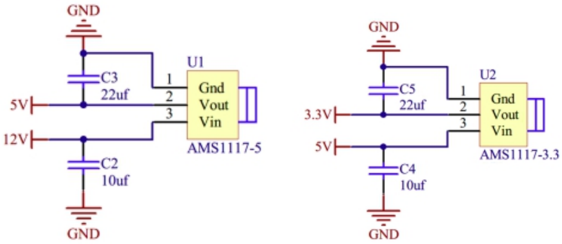


Fig.2. 5V and 3.3V voltage stabilizing circuit

3.3. Executive element

Zsm14k12 steering gear is selected as the actuator of intelligent classification Reclaimer. The actuator is composed of DC motor, control circuit board, variable speed gear set and adjustable potentiometer. The working voltage of the steering gear is 9.5V-12V, the electrical standard of the control signal is 5.0V TTL signal, and the maximum working angle is 270° Zsm14k12 actuator works fast and has stable performance. Its control mode can be serial control or PWM control. Using this actuator as the executive component of the intelligent sorting reclaimer can efficiently complete the action of the turning mechanism.. The appearance of the steering gear is shown in the Fig 3.

Fig.3. The appearance of the steering gear

3.4. Material identification module

Material identification is an important function of the hardware control system of the intelligent sorting and recycling machine. The material identification function depends on the metal sensor module. The metal sensor is also called inductive proximity sensor. LJ18A3-8-Z/B type metal sensor is used in this system. The sensor has the advantages of fast response, strong anti-interference, waterproof and corrosion resistance, and very stable performance. It is often used in automatic assembly line, intelligent hardware and security system.

The working voltage of the sensor is 6-36vdc, which is composed of high frequency oscillation circuit, detection circuit, amplification circuit and output circuit. After the normal power supply of the sensor, the oscillator generates an alternating electromagnetic field on the detection surface of the sensor. When the metal beverage bottle is close to the magnetic field, it will absorb the energy of the alternating magnetic field, which will cause the vibration attenuation and stop. The changes of oscillator oscillation and stop oscillation are processed by the subsequent circuit and converted into a switch signal. The physical drawing of metal sensor is shown in Fig 4.

Fig 4 Physical figure of metal sensor

3.5. Wireless communication module

Esp8266 WiFi wireless communication module is selected as the bridge between intelligent hardware and server. The Esp8266 module provides a complete WiFi network solution with a built-in 32-bit low-power processor. Because of its low cost, high performance and small size, the module is widely used in smart home, wearable electronic equipment, wireless positioning and other fields [2].

TCP is a link oriented and reliable protocol. It needs to wait for the confirmation of the receiver to establish a connection [3]. It is safe and reliable and can ensure data integrity. In this system, the esp8266 module uses TCP protocol to set the WiFi application mode of Esp8266 as sta mode and select TCP connection through at command, so as to realize the function of establishing data connection with ECS. The Esp8266wifi module is shown in Fig 5.

Fig.5. ESP8266 Profile Diagram

4. System server design

This paper selects Alibaba cloud server and builds a Wamp framework for server development, providing services such as communication management, data storage and processing, logical analysis, API interface, etc. for the whole system. The applet is configured with the relevant domain name, and the self built server can be accessed through HTTPS protocol to realize communication. The developed web management platform is deployed to the ECS, and the administrator can log in to the web management platform with a browser to remotely monitor the real-time running status of the recycler.

4.1. Socket network communication

This system selects alicloud server as the TCP server, and designs the TCP console software according to the socket communication principle. Socket socket is in the form of (host IP address: port number), which is located between the application layer and the transport layer of TCP/IP protocol [4]. Using socket can quickly and stably establish the communication connection between TCP client and server.

The ECS waits for and accepts the TCP connection request from the hardware terminal by running the socket.py file. Execute the corresponding SQL statement to save the client data into the database, and complete the real-time collection of the hardware terminal data of the intelligent classification reclaimer by the server.

4.2. WAMP framework

The common Wamp application framework is selected for the server development of this system, which is composed of windows, Apache, MySQL and PHP. Windows is the operating system. Apache is the website server. MySQL is the relational database. PHP is the background development script language. They are independent in themselves, but they are often combined to build the server due to their good compatibility [5].

5. Client development

This project uses wechat developer tools to develop mini programs based on Mina framework. After the user logs in to the applet, click the scan code button on the navigation bar at the bottom of the main page to scan the QR code of the device on the recycling machine. After the scanning, user can enter the throwing page. The function of small program scanning QR code mainly depends on calling wx. Scancode ( ) interface.

The web management platform of intelligent classification and recycling system is developed based on B/S mode, and Ajax technology is used to realize asynchronous request operation of web pages. Ajax implementation does not load the entire page, the part of the content of the page is updated quickly [6]. After logging into the website, the administrator can enter the dynamic monitoring interface, which can remotely monitor the operation status of the intelligent sorting reclaimer in real time.

Acknowledgements

The research is partly supported by the Project of Tianjin Enterprise Science and Technology Commissioner to Tianjin Tianke Intelligent and Manufacture Technology Co., Ltd (19JCTPJC53700). It is also supported by the Industry-University Cooperation and Education Project (201802286009) from Ministry of Education, China.

References

1. Honghua Liao, Hao Fu, Weichuang Yu, et al. Design and Implement of Pyroelectric Energy Harvester Experimental Measurement System Based on STM32F103VET6. *Sensors & Transducers.* 2016, 196(1): pp.69-74.
2. Craig A, Lindley. Meet the ESP8266. *Nuts & volts,* 2015, 36(10):pp.42-47.
3. Joonhyouk Jang, Jinman Jung, Yookun Cho, el at. Design of a Lightweight TCP/IP Protocol Stack with an Event-Driven Scheduler. *Journal of information science and engineering: JISE,* 2012, 28(6): pp.1059-1071.
4. Mehmet S. Unluturk, Mehmet H. Ozcanhan, Gokhan Dalkilic. Improving communication among nurses and patients. *Computer Methods and Programs in Biomedicine*, 2015, 120(2): pp.102-112.
5. Yasir Shoaib, Olivia Das. Web Application Performance Modeling Using Layered Queueing Networks. *Electronic Notes in Theoretical Computer Science*, 2011, 275: pp.123-142.
6. Bhaskar Kumar, Kumar Abhishek, Akshay Deepak, M. P. Singh. Implementation of Interactive Real Time Online Co-shopping Using Push AJAX. *Procedia Computer Science*, 2016, 89: pp.473-482.

Research on Image Super-Resolution Reconstruction Based on Deep Learning

Lingran An [1,3], Fengzhi Dai [1,2,3]*, Yasheng Yuan [1]

[1] College of Electronic Information and Automation,
Tianjin University of Science and Technology, China;

[2] Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China;

[3] Advanced Structural Integrity International Joint Research Centre,
Tianjin University of Science and Technology, Tianjin, 300222, China
*E-mail: * daifz@tust.edu.cn*
www.tust.edu.cn

Abstract

This paper mainly applies the relevant theories of deep learning to image super-resolution reconstruction technology. By comparing four classical network models used for image super-resolution (SR), finally a generative adversarial network (GAN) is selected to implement image super-resolution, which is called SRGAN. SRGAN consists of a generator and a discriminator that uses both perceived loss and counter loss to enhance the realism of the output image in detail. The data sets used by the training network are partly from the network and partly from the artificial. Compared with other network models, the final trained SRGAN network is above average in PSNR and SSIM values. Although it is not optimal, the output high-resolution images are the best in the subjective feelings of human eyes, and the reconstruction effect in the image details is far higher than that of other networks.

Keywords: Super-resolution, deep learning, neural network, Generative Adversarial Networks

1. Introduction

Image resolution describes the number of pixels in an image, which is also a measure of the amount of information in the image. It is an important indicator of image detail information presentation ability. In a large number of electronic image applications, high resolution images are often expected. However, in practice, constrained by many factors, we usually cannot directly obtain the ideal high-resolution image with edge sharpening and non-block blur.

The most direct measure to improve the image resolution is to improve the optical hardware equipment in the imaging system, but this method is expensive and technically difficult Therefore, it is very important to improve the image resolution from the aspect of software and algorithm. In this context, super-resolution reconstruction (SR) technology emerged.

The idea of image super-resolution was first proposed by Harris[1] and Goodman[2] in the 1960s, who adopted interpolation to improve the spatial resolution of a single image. With the rise of deep learning, more and more visual fields are trying to apply it. In 2014, Dong et al. used convolutional neural network to achieve image super-resolution reconstruction (SRCNN)[3]. Subsequently, Dong et al. proposed FSRCNN on the basis of SRCNN[4]. This model is the acceleration of SRCNN, achieving a breakthrough in both the speed and quality of reconstruction. Kim et al. proposed the DRCN algorithm to apply the recursive neural network structure to the super-resolution processing[5]. Wenzhe et al. proposed a real-time super-resolution reconstruction method based on convolutional neural network and named the model

ESPCN[6]. In 2016, Ledig et al. used the generative adversarial network (GAN) for the super-resolution reconstruction problem (SRGAN), and used the perceptual loss and adversarial loss to improve the authenticity of recovered images[7].

2. Image super-resolution reconstruction

2.1. Principle

Image super-resolution technology is to transform low-resolution data into high-resolution data through a certain method on the basis of unchanged image detection system, so as to obtain image observation of high definition images. The formation of low resolution image is often caused by the bad environment, which is often called image degradation process. It can be expressed as:

$$L = DBMH + n \tag{1}$$

Where H and L represent high and low resolution images, M is the matrix after displacement, B is the fuzzy matrix after degradation, D is the matrix for down-sampling, and n is the additional noise pollution.

The process of low-resolution image imaging is a forward process, while the reverse process is the process of image reconstruction. The lost information can be recovered according to the imaging principle to obtain high-quality images.

2.2. Classification

The traditional image super-resolution methods can be divided into three categories: super-resolution technology based on interpolation， reconstruction and learning.

- Interpolation

The algorithm based on interpolation uses the gray value of adjacent pixels to generate the gray value of the pixels to be interpolated so as to realize the super-resolution reconstruction of the image. The classical interpolation methods include neighborhood interpolation, bilinear interpolation and bicubic interpolation.

- Reconstruction

The method based on reconstruction can realize the estimation and reconstruction of high-resolution images by establishing the imaging model of low-resolution images and constructing the prior constraint of high-resolution images. The main research methods include iterative back-projection (IBP), projection onto convex set (POCS) and maximum a posteriori (MAP).

- Learning

The learning-based algorithm can train the images with high and low resolution and master the relationship between them, so as to establish the mapping model. It mainly includes neighborhood embedding (NE), sparse representation (SR), anchored neighborhood regression (ANR), etc. All these methods belong to the field of machine learning.

3. Super-resolution based on deep learning

3.1. Deep learning theory

A simple neural network model is shown in Fig.1. The input layer is on the left, the hidden layer is in the middle, and the output layer is on the right.

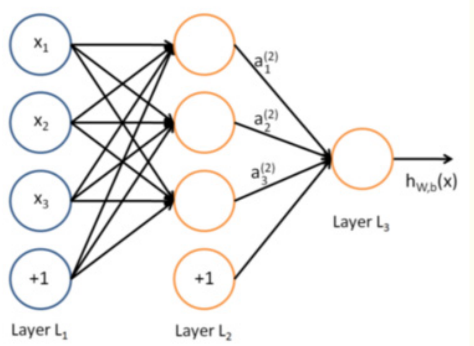


Fig.1. Simple neural network model schematic diagram

The convolutional neural network consists of multiple layers stacked one on top of the other. It takes the raw data that's coming in, and it extracts the high-level information from it, and abstracts it, and that's the feedforward operation of convolution. The error obtained by comparison is fed back to the front layer continuously until the error is reduced to the minimum, so that the model converges to achieve the purpose of training. Its network structure generally includes input layer, convolution layer, pooling layer, full connection layer.

The structure of generative adversarial network is composed of a generator and a discriminator. The generator is used to synthesize the network data, and the discriminator is used to judge whether the network data generated by the experiment is feasible and effective to approach the real value.

3.2. Four super-resolution models

In this section, four deep neural network models are introduced for image super-resolution reconstruction, namely SRCNN, ESPCN, DRCN and SRGAN.

- SRCNN

SRCNN uses only three layers of network to achieve image super resolution It uses mean square error (MSE) as the loss function, which is beneficial to obtain a higher PSNR. The network structure of SRCNN is shown in Fig.2.

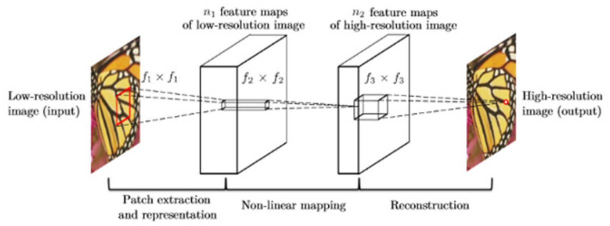


Fig.2. The network structure of SRCNN

- ESPCN

ESPCN is an efficient method to extract features directly from low-resolution image sizes and calculate high-resolution images. The core concept of the network is the sub-pixel convolutional layer, which greatly reduces the computation, saves time and improves the speed of the experimental process. The activation function is tanh function, and the loss function is MSE The network structure of ESPCN is shown in Fig.3.

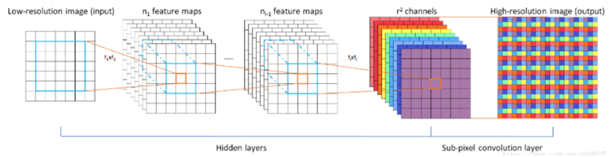


Fig.3. The network structure of ESPCN

- DRCN

DRCN applies recursive neural network in image super-resolution processing, and at the same time uses the idea of residual learning (Skip-Connection) to deepen the network structure (16 recursion), increase the network receptive field, and improve the performance. The network structure of DRCN is shown in Fig.4.

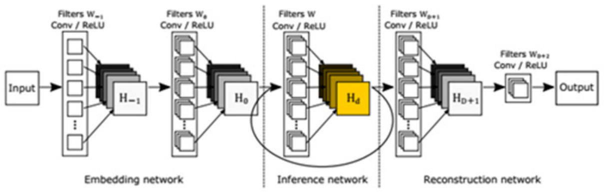


Fig.4. The network structure of DRCN

- SRGAN

SRGAN applies the generative adversarial network to solve the super-resolution problem. The network is composed of a generator network and a discriminator network, which can improve the sense of reality of the generated image by both the loss of perception and the loss of resistance. The network structure of SRGAN is shown in Fig.5.

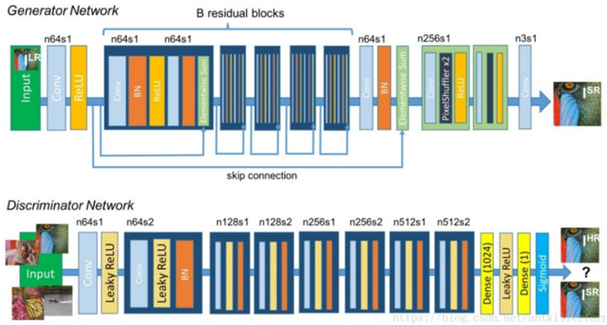


Fig.5. The network structure of SRGAN

3.3. Model selection

Among the above four models, the network model structure of SRCNN is simple and the information obtained is very limited. DRCN network has complex structure and large computation. ESPCN can only deal with images with a small magnification degree. SRGAN can still generate the details in the image in the case of 4 times or more magnification. Therefore, we chose SRGAN as the base network.

4. Model training and testing

4.1. Experimental environment

In this paper, the image super-resolution reconstruction network model based on the generative adversarial network is implemented on the TensorFlow deep learning framework, which provides a Python programming interface to conveniently implement the method proposed in this paper. The environment used in this paper is configured as Ubuntu 16.04 operating system, 32GB DDR4 memory, graphics card GPU model TITAN XP, CPU model e5-2640v4.

4.2. The data set

The data set used by the training network consists of two parts: one part is 8156 high-resolution images from the RAISE data set, and the other part is artificially collected, 844 high-resolution images are randomly shot by camera. A total of 9000 images were collected from the two data sets to form the high-resolution training set, and then the corresponding 9000 low-resolution training sets were obtained by sampling down 4 times. The data sets used by the test trained network are Set5, Set14, and BSD100 test sets.

4.3. Model training process

The training process of SRGAN model is as follows:

1. Train the SRResnet with 1000000 iterations.
2. Train the SRGAN with the weights from the generator of SRResnet for 500000 iterations using the MSE loss.
3. Train the SRGAN with the weights from the generator and discriminator of SRGAN (MSE loss) for 200000 iterations using the VGG54 perceptual loss.

4.4. Evaluation index of experimental results

In evaluating the effect of image super-resolution, the evaluation criteria mainly lie in the gap in data between the reconstructed image and the expected real image. Common indicators include Structural Similarity (SSIM) and Peak Signal to Noise Ratio (PSNR). The value range of SSIM is [0,1], and the closer it is to 1, the better. PSNR is in dB, and the bigger the value, the better.

4.5. Experimental results and analysis

We tested the trained SRGAN network model on the Set5, Set14 and BSD100 test sets, and compared the performance of SRGAN with SRCNN, ESPCN and DRCN, and summarized the results in Table 1. Where the bold part is the maximum value.

Table 1 The performance comparison

		SRCNN	*ESPCN*	*DRCN*	*SRGAN*
Set5	*PSNR*	30.07	30.76	**31.52**	30.95
	SSIM	0.8627	0.8784	**0.8938**	0.8836
Set14	*PSNR*	27.18	27.66	**28.02**	27.78
	SSIM	0.8627	0.8784	**0.8938**	0.8792
BSD100	*PSNR*	26.68	27.02	**27.21**	27.12
	SSIM	0.7219	0.7442	**0.7493**	0.7449

As can be seen from the above table, SRGAN's PSNR and SSIM values are only in the upper middle level, but the high-resolution images generated by SRGAN are the most realistic compared with other methods. A detailed comparison of a set of output images is shown in Fig.6.

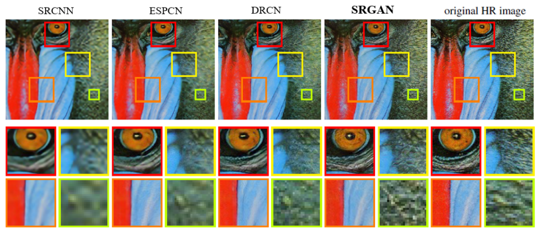


Fig.6. Output picture detail comparison diagram

5. Conclusion

In this paper, an image super-resolution reconstruction network based on GAN model is trained, and the trained network is tested on the open data set and compared with three classical SR networks. It can be seen from the experimental results that although the PSNR and SSIM values of this network are not the highest, the reconstructed image is closer to the real image. It's a more subjective visual network.

Acknowledgements

The research is partly supported by the Project of Tianjin Enterprise Science and Technology Commissioner to Tianjin Tianke Intelligent and Manufacture Technology Co., Ltd (19JCTPJC53700), and the Industry-University Cooperation and Education Project (201802286009).

References

1. Harris J L, Diffraction and resolving power, *JOSA*, 1964, 54(7): pp. 931-933.
2. Goodman J W, Cox M E, Introduction to Fourier Optics, *Physics Today*, 1969, 22: pp 97.
3. Dong C, Loy C C, He K, et al, Learning a deep convolutional network for image super-resolution, *Proc of European Conference on Computer Vision*, 2014: pp 184-199.
4. Dong C, Loy C C, He K, et al, Image super-resolution using deep convolutional networks, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 38(2): pp 295-307.
5. Wenzhe Shi, Jose Caballero, et al, Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network, arXiv preprint arXiv:1609.05158, 2016.
6. Kim J, Lee J K, Lee K M, Deeply-Recursive Convolutional Network for Image Super-Resolution, arXiv preprint arXiv:1511.04491, 2016.
7. Christian Ledig, Lucas Theis, et al, Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network, arXiv preprint arXiv:1609.04802, 2016.

Research on Fatigue Detection Method Based on Deep Learning

Yasheng Yuan[1,3], Fengzhi Dai[1,2,3]*, Lingran An[1], Di Yin[1], Yuxuan Zhu[1], Yujie Yan[1]
[1] Tianjin University of Science and Technology, China;
[2] TianjinTiankeIntelligent and Manufacture Technology CO., LTD, China;
[3] Advanced Structural Integrity International Joint Research Centre, Tianjin University of Science and Technology, Tianjin, 300222, China
*E-mail:*daifz@tust.edu.cn*
www.tust.edu.cn

Abstract

Aiming at the shortage of the existing state of deep learning to test people fatigue, this paper presents a new idea to detect fatigue. First, a video image format is saved to frames, and then the YOLOV3-Tiny algorithm is used to detect faces in images. Compared with traditional OpenCV image process and other deep learning face recognition methods, YOLOV3-Tiny's advantages are mainly that the network is simple, the computation is small, and it can run on the mobile side or the device side. After face recognition, the recognized face is separate out, and then use OpenCV to process the face, roughly divide the face into three areas: the left eye area, right eye area and mouth area. Finally, YOLOV3 is used to identify the state of each region, and the improved PERCLOS algorithm is used to judge fatigue.

Keywords: deep learning, fatigue testing, convolutional neural network, face recognition network

1. Introduction

Accidents caused by fatigue work are a serious problem and are considered to be direct or indirect causes of accidents. The United States (National Sleep Foundation) statistics [1]: driver fatigue is an important cause of traffic accidents, causing nearly 80,000 casualties each year, the severity of the accident is extremely serious. In-depth research on the causes of road traffic accidents found that in the driver's actions unrelated to driving, the perception and judgment decision-making mistakes are the root cause of traffic accidents.

Therefore, how to effectively detect whether a person is in a state of fatigue is the main research direction for solving fatigue problems. At present, there are three mainstream research directions. The first one is judged by human bioelectrical signals [2], but because of the harsh environment and equipment required, it has been in the state of theoretical research; the second method is to predict people's long-term living habits [2, 3], but due to it takes a long time, this method is not real-time. The third is to predict the changes in the look of people through images [4], which is the main method of current research.

2. Main research method

In this paper, deep learning is used to train the human face and the eye, first break the video into frames by frame, then the face in the captured image is first identified and extract the face separately, and then the eye in the facial image is recognized. Through the network of training to identify the state of the eye and the mouth, and finally judge the fatigue state of the person according to the improved PERCLOS algorithm.

2.1 Data set processing

The data set used in this paper is the video taken by the experimenter in driving. The video is decomposed into pictures by python, and extracted every 25 frames to ensure the recognition frequency to improve the accuracy. Then, the face of the decomposed picture is marked, the marked picture is converted into an XML format, and then the XML format is converted into a TXT format, and a corresponding training set and test set are established.

2.2 Net selection

The detection algorithm selected in this paper is YOLOV3 algorithm, and the algorithm structure is shown in Figure 1.

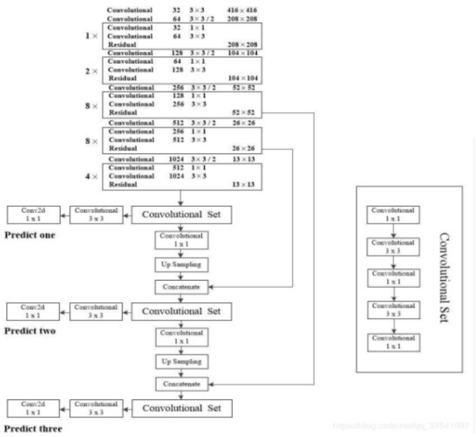


Fig.1. YOLOV3 network structure

The main reason for using YOLOV3 in this paper is that it does all the prediction and detection tasks in one go, which is impossible in the rcnn structure. It is mainly due to the logical design of the loss.

When calculating the loss, the real value and the predicted value are designed in the form of batch size * grid * grid * num_anchor * 25 (or 85), where the grid represents the size of the three-layer output, and in the data feed phase, the annotation value is indirectly convert to the grid dimension and participate in the calculation. This step is ready for the target prediction of the corresponding layer, and the relationship between "grid" and "absolute position" is solved. Num_anchor directly represents which anchor the output layer should use (the first layer 6, 7, 8, the second layer uses 3, 4, 5, the third layer uses 0, 1, 2), also known as multi-scale prediction.

YOLOV3 is a priori detection system which can reuses the classifier or locator for performing inspection tasks. They apply the model to multiple locations and scales of the image. Those areas with higher scores can be considered as test results.

We apply a single neural network to the entire image, which divides the image into different regions, thus predicting the bounding box and probability of each region, which are weighted by the predicted probability. Our model has some advantages over a classifier-based system. It looks at the entire image as it is tested, so its predictions take advantage of the global information in the image. Unlike R-CNN, which requires thousands of single-target images, it is predicted by a single network assessment. This makes YOLOV3 very fast, generally 1000 times faster than R-CNN and 100 times faster than Fast R-CNN. Figure 2 is a comparison of the YOLOV3 algorithm with other algorithms [5].

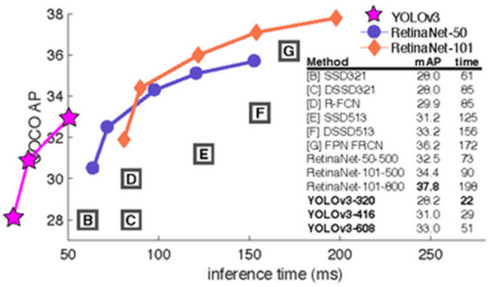


Fig.2. YOLOV3 algorithm comparison

2.3 Identify face areas

In this paper, the YOLOV3-Tiny algorithm is used to detect and extract the position of the face in the picture. YOLOV3-Tiny is a simplified version of YOLOV3. The main difference is that the main network adopts a 7-layer conv+max network extraction feature (similar to darknet19), the grafting network adopts the resolution detection network of 13*13 and 26*26, and the structure is shown in Figure 3. The main advantage of YOLOV3-Tiny is that the network is simple, the amount of calculation is small, and it can be run on the mobile or device side.

YOLOV3-Tiny has only 24 network layers in total, which is much smaller than the yolo3 107 layer. There are only two YOLO layers, yolo16 and yolo23, which are 13x13 and 26x26 respectively. In addition, each YOLO layer also has 3 anchors, for a total of 6 anchors. The front of the YOLO layer is a 1x1 convolutional layer whose input and output sections remain the same in width, height and channels. In darknet, the convolution formula is: output = (input + padding - kernel_size) / stride + 1. Here padding = 0.

The disadvantage is that the accuracy is relatvelly low (candidate frame and classification accuracy are relatively low), but since we only detect one class, the recognition accuracy will be greatly enhanced. Figure 3 is the structural network of YOLOV3-Tiny.

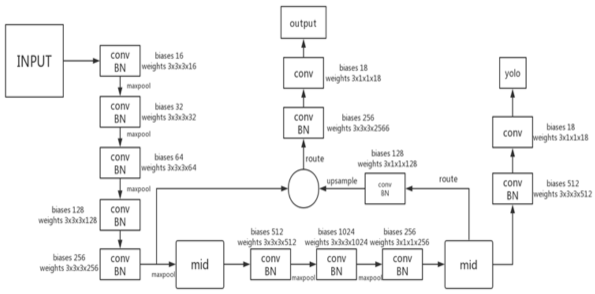


Fig.3. YOLOV3-Tiny network structure

Before training, we first add a crop save function to the darknet image.c, so that we can intercept the target image from the original image after the detection.

2.4 Facial image processing

In this paper, the obtained face is segmented by OpenCV tool, and it is divided into left eye area, eye area and mouth area. After the segmentation completed, each part of the image is marked separately, and the specific target is blinking eyes and yawning mouth. Since the image processed by OpenCV is enlarged by the division, the definition has a certain decline. In order to ensure the accuracy of the recognition, the YOLOV3-spp network is used to train the target separately. YOLOV3-spp is an improvement of YOLOV3 plus 'spp' module, which is more accurate than the YOLOV3 [6]. The difference with the yolo network is shown in Figure 6.

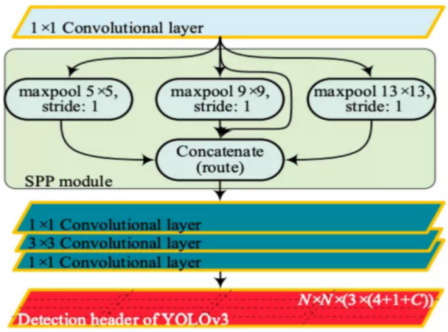


Fig.4. YOLOV3-spp

3. Improved PERCLOS algorithm

PERCLOS (Percentage of Eyelid Closure Over the Pupil over time) is a set of criteria commonly used in fatigue driving research, according to a study by the Carnegie Mellon Institute in the 1970s, within a specified time frame. The percentage of the total length of time in which the eye is in the closed state is highly correlated with the degree of fatigue of the person. The ratio of the length of time and the total length of time in which the human eye is in a closed state for a period of time can determine whether the person is fatigued or not, because the image acquisition frame rate is constant. The duration ratio can be approximated simply by the ratio of the number of image frames. At the same time, we also added the degree of mouth opening parameter A to the criteria for judging fatigue. The formula is shown①and②.

$$\mathbf{P} = \frac{\textbf{Frame Number of Closed Eyes}}{\textbf{All Frames in a Period}} + \mathbf{A} \qquad ①$$

$$\mathbf{A} = \log\left(1 + \frac{\textbf{2*Frame Number of open mouth}}{\textbf{All Frames in a Period}}\right) \qquad ②$$

We choose 30 pictures for one period. Because the time of blinking in normal human state (the time when the eyes are closed) is less than 1s, about 0.2 to 0.4s. The average person blinks a dozen times in one minute, and we extract the video every 25 frames (average 1s to capture a picture). That is to say, without considering the state of the mouth, other factors such as errors are taken into account, and the frequency of closing the eyes of a normal person in one cycle should be 0-7 pictures, so we specify the value of P in the fatigue state. For [0.6, 1.4], the value of P in the awake state is [0, 0.4].

4. Conclusion

In this paper, the training parameters of YOLOV3 are modified, and the maximum number of iterations is set to 22000 times. When iteration is 15000 times, the learning rate is attenuated ten times. At 20000 iterations, the learning rate is again a decade-fold based on the previous learning rate. In order to increase the training sample, the rotation angle was set to 20, the saturation was adjusted to 1.4, and the exposure was 1.6. The learning rate is set to 0.002. Figure 5 and Figure 6 show the change curves of loss and IOU during training. Figure 7 shows the probability of detecting the target.

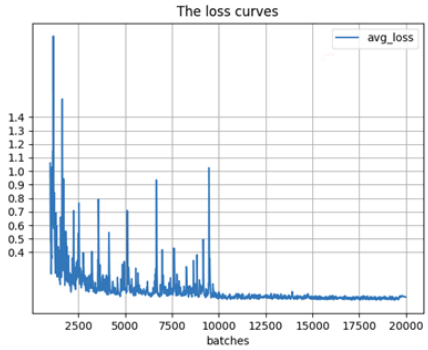


Fig.5. YOLOV3-Tiny loss curves

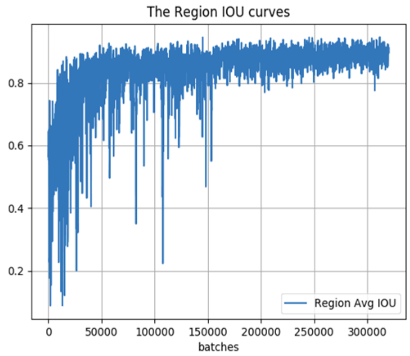


Fig.6. YOLOV3-Tiny IOU curves

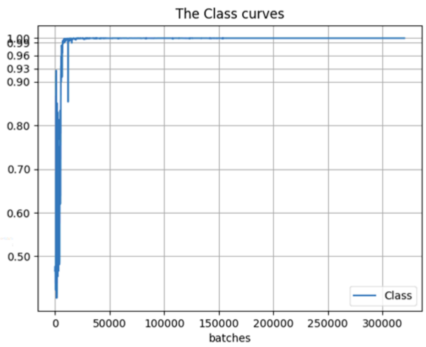


Fig.7. YOLOV3-Tiny Class curves

The verification set uses several tens of seconds of video to frame the video into images for more intuitive observation of accuracy and recognition speed. The average accuracy and recognition speed of each picture are shown in Table 1, and the fatigue state detection results are shown in Table 2.

Tab.1. Network comparison

Net	yolov3-tiny	yolov3-spp
Average-accuracy	94.711%	87.759%
Average-recognition-speed	0.00902s/P	0.06236s/P
Average-detection-frame-rate	69.7/s	16.3/s

Tab.2. Fatigue state detection

Consequence	Awake frame	Fatigue frame
Total pictures	200	50
Awake	177	9
Fatigue	23	41
Accuracy	88.5%	82%

Acknowledgements

The research is partly supported by the Project of Tianjin Enterprise Science and Technology Commissioner to Tianjin Tianke Intelligent and Manufacture Technology Co., Ltd (19JCTPJC53700). It is also supported by the Industry-University Cooperation and Education Project (201802286009) from Ministry of Education, China.

References

1. State of the States Report on Drowsy Driving, National Sleep Foundation, November 2007, U.S.A.
2. Craig A, Tran Y, Wijesuriya N, et al. A controlled investigation into the psychological determinants of fatigue. *Biological Psychology*, 2006, 72(1): 78-87.
3. Marshall N S, Barnes M, Travier N, et al. Continuous positive airway pressure reduces daytime sleepiness in mild to moderate obstructive sleep apnoea: a meta-analysis. *Thorax*, 2006, 61(5):430-434.
4. Donald Sussman, Michael CoPlen. Fatigue and alertness in the United States railroad industry. *Transportation Research Part: F*, 2000, 3:211-220.
5. Tey F, Lin S T, Tan Y Y, et al. *Novel Tools for Driving Fatigue Predicting:(1) Dry Eeg Sensor and (2) Eye Tracker.* International Conference on Augmented Cognition. Springer Berlin Heidelberg, 2013: 618-627.
6. J. Redmon and A. Farhadi. *Yolov3: An incremental improvement*. arXiv, 2018
7. Pengyi Zhang, Yunxin Zhong, Xiaoqiong Li, *SlimYOLOv3: Narrower, Faster and Better for Real-Time UAV Applications*, arXiv, 2018.

Fatigue driving monitoring system based on the EEG

Yuxuan Zhu [1,3], Fengzhi Dai [1,2,3]*, Di Yin [1], Yasheng Yuan [1]

[1] *Tianjin University of Science and Technology, China;*

[2] *Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China;*
[3] *Advanced Structural Integrity International Joint Research Centre,*
Tianjin University of Science and Technology, Tianjin, 300222, China
*China E-mail: * daifz@tust.edu.cn*
www.tust.edu.cn

Abstract

This paper analyzes how to acquire the EEG signal, how to extract and analyze the EEG signal characteristic rhythm and how to estimate the fatigue degree and provide the voice reminder. Firstly, four characteristic rhythms of EEG signals are extracted by wavelet packet decomposition, and the characteristics are analyzed by relative energy frequency spectrum of rhythms. Then, according to the classification of fatigue degree, the energy value of δ wave in F3, F4 and C3 channels is selected as the basis of judging driver fatigue to classify and estimate the fatigue degree of EEG signals. Finally, when the threshold reaches 0.4, the sound card of the computer is called to prompt the tired driver.

Keywords: EEG, fatigue degree, signals, tired driver, working principle

1. Introduction

In recent years, more and more research groups are engaged in this field. With regard to Driver fatigue monitoring technology, the Drowsy Driver Detection system developed by the United States adopts doppler radar and complex signal processing methods to obtain Driver's fidget mood fluctuation, blink frequency, duration and other fatigue data, so as to determine whether the Driver is Drowsy or asleep [1]. At the university of Tokyo in Japan developed the fatigue tester, can take on the driver's wrist, the test instrument with a miniaturized oxygen battery electrodes within measurable driver lactic acid, ammonia and alcohol content in sweat, and then through a small radio transmitter sends the data to the research center, research center, through the computer analysis, determine the level of fatigue of the driver and timely warning to the driver [1, 2]; DAS2000 Road Alert System , developed by American laboratory, is an infrared monitoring device controlled by computer on the highway, which will give a warning to the driver when the vehicle passes the middle line or shoulder of the Road. The Psychomotor Vigilance Tes tests the response time of the driver according to his response to the random light spot on the instrument screen, so as to judge the fatigue degree [3]. The indicator used by Spanish anti-fatigue system is steering pressure of the steering wheel. Once the steering pressure signal of the driver is detected to be abnormal, the driver and surrounding vehicles will be warned by the flashing of the headlight and sound alarm, and the fuel supply will be cut off automatically, forcing the car to stop.

2. EEG acquisition system

Electroencephalic signals are generated by the synchronous activation and inhibition of the postsynaptic potential in the cerebral cortex. These potentials accumulate in the cerebral cortex and pass through the brain's covering to the scalp, a random signal with a fairly complex mechanism. The human brain can be divided into four parts: telencephalon, cerebellum, brainstem and hypothalamus. The cerebral cortex, which underlies much of the higher neural activity, is most involved in the study of electrical signals. The cerebral cortex can be divided into many different regions, each with a different function.

2.1. Acquisition equipment

According to the device size, the number of electrodes and the communication technology, the common signal acquisition equipment is divided into two categories: non-portable and portable. Non-portable EEG acquisition devices typically have 32, 64, 96, 128 or 256 electrodes. Most of these devices are micro-amplitude precision electronic devices with a magnification of ten thousand times. They usually transmit signals without wires, and have a large number of electrodes, so they can obtain rich and accurate EEG signals, which are suitable for scientific research and clinical medicine [4]. Due to its large size, high price, complicated operation, strict requirements on the use environment and condition setting, and making the subjects nervous and other shortcomings, the application of this device in real life is limited.

2.2. Electrode position

There are a number of placement methods for the placement of scalp electrodes. At present, there are a whole set of international standard methods for positioning electrodes on the scalp, and the international 10-20 system electrode placement method is generally adopted at home and abroad [4]. This is the standard electrode placement method recommended by the international electroencephalography society. To distinguish the relationship between the two hemispheres and the electrodes, an even number is usually used on the right side and an odd number on the left. The electrode positions used in this project are shown in fig1.

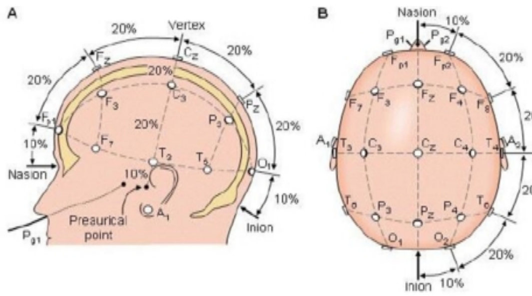


Fig.1.Side and top view of international 10-20

To sum up, the brain electrical signal acquisition using national unification standard of 10-20 system, several important unknown to choose common placement of electrodes, including: the frontal area, occipital area, area and the central, acquisition are: the location of the point of Fz, F3, F4 and F7, F8, T3, T4, C3, C4 and T5, T6, P3, P4, O1, O2, Pz 16 channels of EEG data as the experimental data.

3. EEG signal preprocessing

In this paper, EEGLAB, a toolbox of MATLAB, is used for EEG signal preprocessing. The EEG data from EEG acquisition device, the device has 16 electrodes, according to the national standard of 10 to 20 lead placement, collecting electrodes respectively: Fz, F3, F4 and F7, F8, T3, T4, C3, C4 and T5, T6, P3, P4, O1, O2, Pz, collecting electrode distribution as shown in figure 3 to 5, the subjects in a state of calm him only. During the test, subjects were asked to remove distractions and relax as much as possible. The sampling frequency is 250Hz, the sampling time is 180min, and every 30min is used as the sample of independent component analysis.

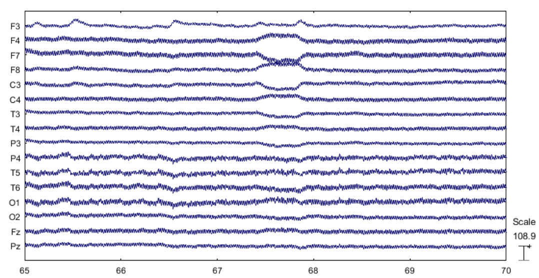


Fig.2 Primitive EEG signals

Fig2 shows the original EEG signal with a length of 5s, and the number of signals is 76,000. It can be seen from the figure that the eye electrical artifact, such as F4, F7, F8, C3, C4, T3, T4 and P3, blinks twice between 67 and 68s. F3 has multiple single blinks between 65 and 68s. The artifact of eye electrical signals will bring great difficulties to the subsequent analysis and processing of EEG signals. In theory, it is believed that the interference signals in EEG signals such as cardiac, eye movement artifacts, EEG signals and other interference sources are all generated by independent sources. By ICA decomposition, useful EEG signals can be extracted.

In this paper, the extended Infomax algorithm is used for independent component analysis, and the first 30 minutes of signal collection are taken as samples of the independent component analysis method. The decomposition results are shown in fig3. It can be clearly seen from the figure that components 5, 11 and 12 are obvious EEG signals, because the frequency of these signals is relatively large.

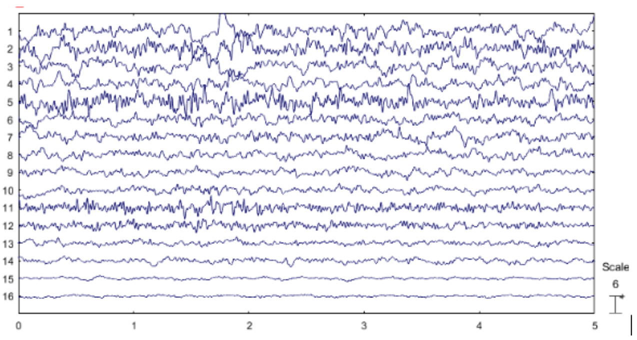


Fig.3 Each component diagram

It can be seen from fig4 that IC6 is an eye electrical artifact. Since the energy of IC6 component is mainly concentrated on the right side of the forehead, it indicates that the artifact is caused by eye movement. The energy of the independent components IC15 and IC16 is mainly concentrated in the temporal region and near the ear root, which is consistent with the typical electromyographic artifact interfering with the EEG topographic map.

4. Analysis of characteristic rhythm of EEG signals

In order to display the characteristics of EEG signal changes with fatigue obviously, the method of calculating

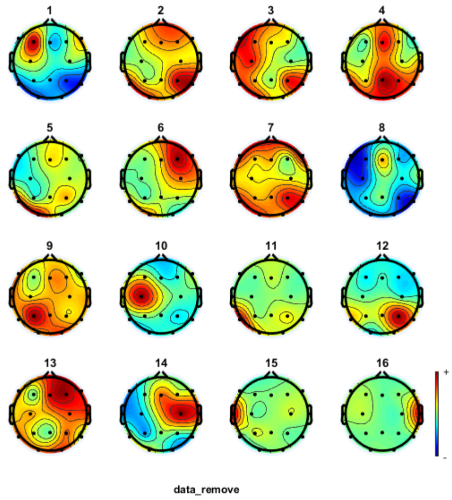


Fig.4 Topographic map of the individual components

the relative energy value of EEG signal in each rhythm was selected to analyze the characteristic changes of EEG signal in fatigue driving. The db1 of the dau-bechies system was used for EEG signals, which were decomposed by 7 layers of multi-scale wavelet packets, to re-construct four characteristic rhythm waves (0~4Hz), waves (4~8Hz), waves (8~13Hz) and waves (14~30Hz). Since wavelet packet decomposition can meet the principle of constant signal energy, the energy values of each rhythm and its relative energy were used as measurement indexes.

In the experiment, the sampling frequency was 250Hz, the sampling time was 3h, and every 30min was a segment, which was divided into 1~6 segments. The first segment was 0~30min, the second segment was 30~60min, the third segment was 60~90min, the fourth segment was 90~120min, the fifth segment was 120~150min, and the sixth segment was 150~180min. Since the collected EEG signal data volume is too large, and the EEG signal variation rules of Fz, F3, F4, F7, F8, T3, T4, C3, C4, T5, T6, P3, P4, O1, O2 and Pz are roughly the same, this paper only takes Fz conductance as an example for analysis.

Based on the above principles, the subband energy extraction method for EEG signals can be summarized as follows: first, the EEG signals preprocessed in the previous chapter are decomposed into wavelet packet 7 layers by using "db1" function. Then, the corresponding frequency bands of wave, wave, delta wave and delta wave are reconstructed. The reconstructed signal is then calculated for each frequency band. Finally, the proportion of the four characteristic rhythms in the total energy is calculated. The relative energy values of Fz dot in 6 time periods are shown in Table 1.

Table1 The relative energy value of characteristic rhythm

Time	The energy value			
	δ	θ	α	β
0-30min	0.363	0.275	0.173	0.166
30-60min	0.376	0.245	0.167	0.212
60-90min	0.384	0.236	0.160	0.223
90-120min	0.413	0.220	0.150	0.243
120-150min	0.382	0.242	0.163	0.217
150-180min	0.367	0.272	0.171	0.180

As can be seen from the table, the relative energy value of and rhythm increases gradually in the first to the fourth segment, weakens in the fifth segment, and increases again in the sixth segment. The changes in the mean value of the relative power spectrum of the circadian rhythm are the opposite of the delta and delta rhythms. The relative power spectrum of circadian rhythm is gradually weakened in segments 1 to 4, and then enhanced in segments 5 and 6. Theoretically speaking, the main waveforms in the excited state of the cerebral cortex are the delta and delta waves, and the slow waves are the main manifestations of the electrical activity in the suppressed state of the cerebral cortex [5].

The analysis of the test results showed that: during the first and second period of time, the changes in and rhythm were relatively small, so it was approximately believed that the driver's brain was active at this time and there were no symptoms of fatigue. In the third period, the mean value of relative power spectrum of and rhythm increased significantly, while the relative energy value of cycle and delta rhythm decreased significantly, indicating that the driver had developed fatigue symptoms at this time [6]. Compared with the first three paragraphs, the changes of each rhythm in the fourth paragraph were more obvious. The relative energy values of and rhythms were significantly enhanced, while the relative energy values of and rhythms were just the opposite, indicating that the driver's brain was greatly inhibited and the driver was in a state of fatigue. In the fifth period, delta and theta rhythms of relative energy appears more obviously weakened, and beta rhythm relative energy appear obvious enhancement, the relative energy of alpha rhythm also increased, this is because the driver fatigue to a certain extent, can appear sleepy phenomenon, the nerve center of the brain at high inhibition, and the driver drowsiness after a short period of time, the brain's neural inhibition of ease, and relieve fatigue, alpha and beta waves will increase, so the relative energy increase, the corresponding delta and theta will be reduced, thus weakened the relative energy value. In the sixth period, the relative energy value of circadian rhythm was significantly decreased again, and the value was the smallest in all time periods. The relative energy value of, and circadian rhythm was increased to different degrees, indicating that the driver was in the most fatigued state at this time. Because delta, delta, and delta waves increase when you're at your most fatigued.

Acknowledgements

The research is partly supported by the Project of Tianjin Enterprise Science and Technology Commissioner to Tianjin Tianke Intelligent and Manufacture Technology Co., Ltd (19JCTPJC53700). It is also supported by the Industry-University Cooperation and Education Project (201802286009) from Ministry of Education, China.

References

1. Arnaud Delorme, Scott Makeig. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 2003, 134(1).
2. Ozcan Ahmet Remzi, Erturk Sarp. Seizure Prediction in Scalp EEG Using 3D Convolutional Neural Networks With an Image-Based Approach. *IEEE transactions on neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society*, 2019, 27(11).
3. Mayo Morimoto, Misako Imono, Seiji Tsuchiya, Hirokazu Watabe. Emotion Judgment System by EEG Based on Concept Base of EEG Features. *ZTE Communications.*
4. Jin Jing, Emile Dangremonta, Senan Ebrahim, Mohammad Ghassemi, Eric Rosenthal, Sahar Zafar, Brandon Westover .Automatic clustering of EEG data from ICU patients. 2018(01).
5. Jean-Christophe Cuilliere, Vincent Francois, Remi LacroisLi, A new approach to automatic and a priori mesh adaptation around circular holes for finite element analysis, *Computer-Aided Design*, 2016.
6. Qingxia Li, Wenhong Wei, Study on the structure of an extended OTIS-Mesh network, *Computer Science*, 2012, z3: P90 - P92.

Research of the Control Strategy of Vienna Rectifier Circuit based on the Vector Control

Yuxuan Zhu [1,3], Fengzhi Dai [1,2,3]*, Di Yin [1], Yasheng Yuan [1]

[1] *Tianjin University of Science and Technology, China;*

[2] *Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China;*
[3] *Advanced Structural Integrity International Joint Research Centre,*
Tianjin University of Science and Technology, Tianjin, 300222, China
*China E-mail: * daifz@tust.edu.cn*
www.tust.edu.cn

Abstract

This paper analyzes the working principle of VIENNA rectifier and establishes the transformation model of rectifier in three different mathematical coordinate systems. First is to obtain the control structure of the converter, which is a double closed loop control structure (the voltage is the outer loop control and the current is the inner loop control). By using a feed forward control strategy to solve the problem of phase-to-phase coupling. By adding the voltage equalization loop and using the midpoint balance algorithm to solve the problem of the voltage imbalance caused by the rectifier under load. Finally, uses MATLAB to build the model, simulates and verifies the established model system.

Keywords: VIENNA rectifier, model systems, voltage imbalance, simulate, working principle

1. Introduction

Power electronic devices can be seen everywhere in our life. AC-DC rectifiers are common. Most rectifying devices adopt traditional rectifying methods, such as non-controlled rectifying circuit with diode and semi-controlled rectifying circuit with thyristor [エラー! 参照元が見つかりません。]. These topology structures generally have some shortcomings such as high current harmonics and low power factor. Due to the non-linear rectifying characteristics of rectifiers, a large number of harmonics and reactive power are injected into the power grid, causing serious interference to the power grid environment and reducing the power factor.

Therefore, how to solve the problem of harmonic and reactive power has attracted people's attention. Reducing harmonic pollution to the grid environment is what every engineer and government want, so many industry standards have been developed. The most widely used are ieee519-1992, iec555-2 and iec1000-3-2.

The application of PWM technology to control the rectifier, can make the rectifier control simple, excellent performance, at the same time the power grid side current sine and voltage in phase to ensure that the equipment in the unit power factor operation. VIENNA rectifier two level relative to the rectifier, VIENNA rectifier output level increase in the number, therefore, under the same bus bar voltage, low switch voltage stress, high power factor, low input current harmonics, high reliability, good features, become the hot research topic in the rectifier, the merits of the control strategy determines the performance of the rectifier, the VIENNA rectifier control strategy research has the important value of engineering application [2].

2. Working principle

The VIENNA rectifier is shown in Fig.1 with three-phase symmetrical ac input power supply and equivalent inductance on ac side. It plays the role of energy transfer in the rectifier, and can effectively suppress the high-order harmonics generated in the switching process and balance the voltage of each bridge arm and the grid input voltage. The capacitor on the output side of the dc bus with two identical parameters constitutes the tri-level structure. The rectifier has two fast recovery power diodes per bridge arm, so the rectifier needs six fast recovery power diodes. The rectifier also has three sets of bidirectional power switches. Because of the existence of the middle line, the three phases are independent of each other, and each phase can form an independent single-phase tri-level structure. In this paper, the three - phase three - wire topology is studied.

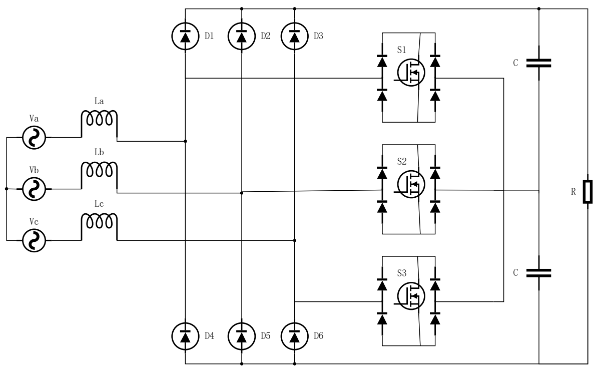


Fig.1 VIENNA circuit structure schematic diagram

3. Mathematical model of VIENNA rectifier

3.1. Mathematical model in three phase coordinate system

In order to achieve the best control performance, the circuit needs to be idealized. In the process of establishing the mathematical model of VIENNA rectifier, the following assumptions are made[3]:

1) the sinusoidal voltage waveform of the three-phase ac input side of the rectifier is symmetrical;

2) the power switching devices in the analyzed system are all set as ideal components;

3) the switching frequency of the rectifier is much higher than the fundamental frequency of the ac input side;

4) the upper and lower capacitance parameters of the dc bus side are the same.

Redefine the switching function, and set Sj (j= a,b,c) as the switching function of the JTH phase, which can be expressed as equation. The equivalent circuit diagram is shown in Fig.2.

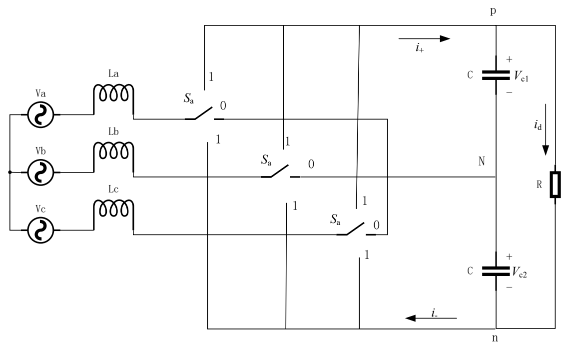


Fig.2.VIENNA circuit diagram

In order to meet the three-phase power grid voltage symmetry, the mathematical model expression of VIENNA rectifier in the three-phase stationary coordinate system of ABC can be obtained. It is known that VIENNA rectifier is a multivariable, strongly coupled, high-order nonlinear system. The state equation of VIENNA circuit in three-phase ac coordinate system can be expressed as:

$$Z\frac{dy}{dx} = AX + BV \tag{1}$$

3.2. Mathematical model in two phase coordinate system

The system is not easy to control and the calculation is difficult. In order to simplify the control structure of the system, coordinate transformation is carried out to obtain the synchronous rotating coordinate system. In this coordinate system, it is easy to calculate the direct flow. Let d axis in d-q coordinate system and a axis in ABC three-phase coordinate system be in the same direction, so the included Angle is 0. The transformation of coordinate system is divided into two stages. The first stage is the transformation of $\mathrm{abc}\text{-}\alpha\beta$ coordinates. The second stage is the $\alpha\beta\text{-}\mathrm{dq}$ coordinate transformation. Fig.3 shows the transformation matrix from three-phase ac static coordinate system to two-phase ac static coordinate system, and the change of constant amplitude [4].

abc - $\alpha\beta$ can be expressed as:

$$T_{3s/2s} = \frac{2}{3}\begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{pmatrix} \quad (2)$$

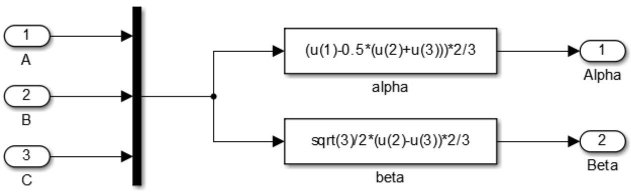


Fig.3 Clark transformation matrix

3.3. Mathematical model in d-q coordinate system

$\alpha\beta$ - dq can be expressed as:

$$T_{2s/2r} = \begin{pmatrix} sinwt & -coswt \\ coswt & sinwt \end{pmatrix} \quad (3)$$

In the formula, wt is the output phase.

The $\alpha\beta$ - dq is simulation model is shown in Fig.4.

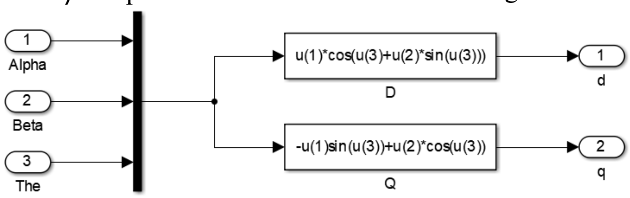


Fig.4 $\alpha\beta$ - dq is simulation model

4. System simulation of VIENNA rectifier

4.1 open-loop simulation of VIENNA rectifier

In the process of designing the open loop control system, the system controller and the controlled object should be considered. The controller generates the control signal to the controlled object to control it to achieve the expected function. The open-loop system is a relatively basic control system. Firstly, the open-loop circuit of VIENNA rectifier is built to verify the feasibility and anti-interference ability of the system, so as to draw the difference between open-loop control system and closed-loop control system. As can be seen from Fig.5-1 and Fig.5-2 below, stable voltage and current output can be achieved when there is no interference in the system. When there is interference in the system, the output of the open-loop system with sudden load reduction has a big jump, and there is no way to restore to the set value, so the anti-disturbance ability of the open-loop system is poor.

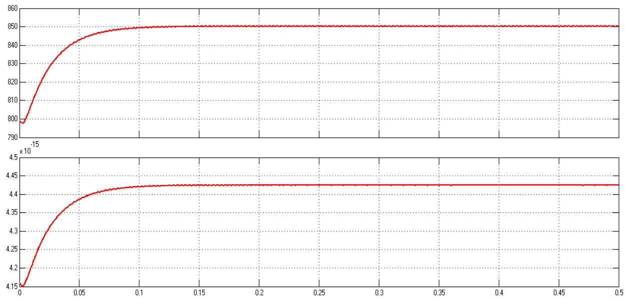


Fig.5-1 VIENNA rectifier open-loop steady state voltage and current output

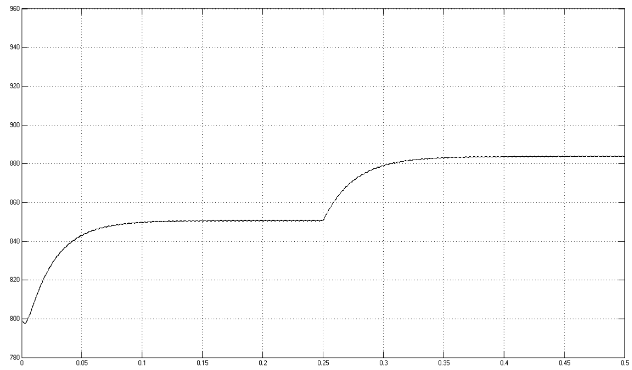


Fig.5-2 perturbation in an open-loop system

4.2 Closed-loop simulation of VIENNA rectifier

The control scheme is verified by simulation and experiment. As long as the control parameters are calculated correctly, the circuit performance can meet the specified requirements.

Simulation parameters of the main circuit system are as follows: switching frequency fs=10KHz; Inductance value La=Lb=Lc=10mH; Stabilized voltage filter capacitor C1=C2=2200μF. Fig6-1 and Fig6-2 are respectively the dc side voltage simulation waveform and the three-phase input voltage and input current simulation waveform of the VIENNA rectifier in steady-state operation. The ac input voltage Va=Vb=Vc =220V, rectifier at full load, load impedance load for R =72 Ω, Vdc+=Vdc-=425V.

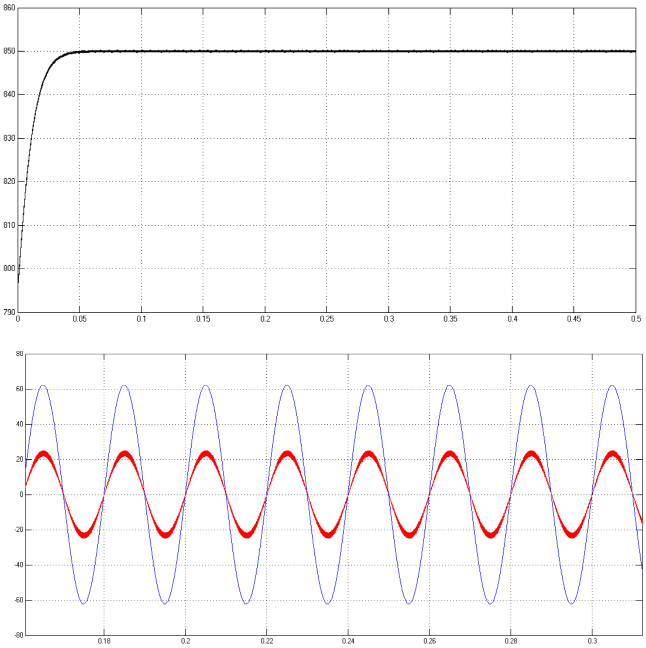


Fig.6-2 The steady-state waveform

4.3 simulation of mutation load

In order to verify the reliability of the system of the system and the ability to suppress the load variation, the simulation of load abrupt addition and abrupt subtraction was carried out in the built simulation model, and the waveform change of voltage and current on the ac side and voltage on the dc side was observed.

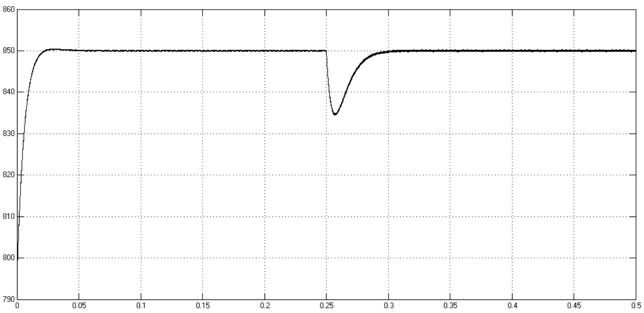


Fig.7-1 Output voltage waveform applied by load

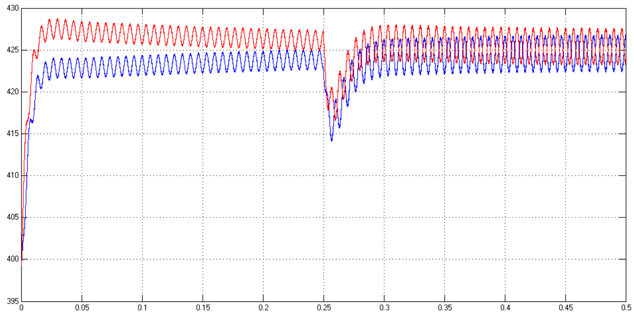


Fig.7-2 Output capacitance voltage waveform applied by load

Acknowledgements

The research is partly supported by the Project of Tianjin Enterprise Science and Technology Commissioner to Tianjin Tianke Intelligent and Manufacture Technology Co., Ltd (19JCTPJC53700). It is also supported by the Industry-University Cooperation and Education Project (201802286009) from Ministry of Education, China.

References

1. Kolar J W, Friedli T. The essence of three-phase PFC rectifier systems—Part I. *Power Electronics, IEEE Transactions on*, 2013, 28(1): 176-198.
2. Soeiro.T. Three-Phase High Power Factor Mains Interface Concepts for Electric Vehicle Battery Charging Systems. *APEC*, 2012.
3. Mengdi Li, Xuan Wu, Shoudao Huang, Ge Liang. Model predictive direct power control using optimal section selection for PWM rectifier with reduced calculation burden. *International Journal of Electrical Power and Energy Systems*, 2020, 116.
4. Angela Osborne. Dynapower's Multiunit Rectifier Controller Consolidates Power Controls. *Products Finishing*, 2019, 84(2).
5. Sodikdjon A. Kodirov, Vladimir L. Zhuravlev, Johannes Brachmann. Prevailing Effects of Ibutilide on Fast Delayed Rectifier K+ Channel. *The Journal of Membrane Biology*, 2019, 252(6).

Research on emotion classification based on EEG

Di Yin [1,4], Fengzhi Dai [1, 3,4]*, Mengqi Yin[2], Jichao Zhao[1]

1 Tianjin University of Science and Technology, China;
2 Hebei University of Chinese Medicine, China,
[3] Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China,
[4] Advanced Structural Integrity International Joint Research Centre, Tianjin University of Science and Technology, Tianjin, 300222, China
*China E-mail: * daifz@tust.edu.cn*
www.tust.edu.cn

Abstract

Research shows that human emotion production is closely related to the activity correlation of cerebral cortex, so the research of emotion classification by EEG provides a reliable basis. The feature extraction and classification application of EEG have made rapid development, so we combine EEG with emotion to study emotion classification. However, there are differences between EEG signals of different subjects, which have a certain impact on emotion classification. How to ensure the high accuracy and robustness of recognition is a problem. In view of this problem, the spectrum analysis method is used to extract features to study different subjects in different states. The extracted features are classified into emotion by discriminant analysis algorithm, and the classification effect is satisfactory. There are many methods involved in feature extraction and the space is long, different feature extraction methods will be compared later, so as to improve the robustness and efficiency of emotional classification of EEG signals.

Keywords: EEG; Feature extraction; Channel selection; Spectrum analysis; Sentiment classification

1. Introduction

The emotional state of a person has a certain influence on the body's own cognitive and behavioral aspects. At present, the well-known emotion research mainly focuses on the external analysis of speech tones, word language and facial expressions. Human emotions are mainly generated by their own physiological and psychological information. This information alone cannot accurately reflect human emotion changes. The field of brain research has become an important research field. With the development of neuroscience and brain science, emotions are no longer unpredictable. Using EEG information data to classify and identify different emotions, aiming at improving the classification and recognition effect of emotions and exploring the adjustment mechanism of different emotions. For emotional exploration, it not only helps the treatment of mental illnesses such as depression, but also has great significance in the field of brain-computer interaction.

In emotion research, researchers designed different experimental models to detect emotion through different signals and different stimulus materials. Koelstra [1] et al. Used 40 audio and video clips as stimulus materials to induce EEG signals and peripheral physiological signals of subjects. The later research of deap emotion analysis database created by the experiment made great contributions. Lin [2] et al. extracted the differential laterality (DLAT) features of the original EEG signals, so as to link the spectral pattern of EEG space with the hidden emotional state, and explore the feasibility of improving the emotional classification performance by using the multi day EEG data of each person [3].

There are differences between the EEG signals of different subjects, which have a certain impact on the emotional classification. How to ensure high accuracy and robustness of identification, this paper uses spectrum analysis method to extract features and discriminate and analyze EEG data.

2. Principle of emotion recognition

2.1 EEG data emotion recognition principle

The complete EEG emotional recognition system consists of emotional EEG data sets, preprocessing, feature extraction and sentiment discrimination classification, which is shown in Figure 1.

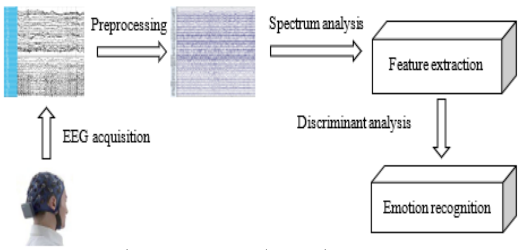


Fig1. System schematic

2.2 Emotional model

The most famous emotional dimension in emotional research is the Valence-Arousal model, but the model does not accurately map all human emotions. Therefore, a three-dimensional model appears. As shown in Figure 2, each dimension represents the degree of unpleasantness and happiness. The degree of excitement and the degree of relaxation to tension provide a strong support for emotional research [4].

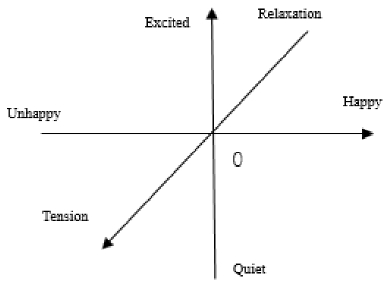


Fig2. Three-dimensional emotion model

3. Experiment

3.1 Experimental data acquisition

The EEG acquisition uses a Neuroscan Synamps 2 brain electrical amplifier with a sampling frequency set to 1000 Hz. The electrode is set according to the international 10-20 system, and the top of the head is selected as the reference electrode. The 32-lead electrode cap is used for recording, and all electrode impedances are lower than 10KΩ as shown in Figure 3.

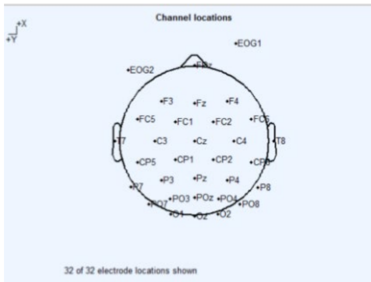


Fig3. Schematic diagram of the 10-20 system recording electrodes distribution of 32 channels

3.2 Data source and preprocessing Clustering results

Five subjects were recorded according to the five emotions of pleasure, relaxation, excitement, nervousness and calmness in the five audio and video states of funny, relaxed music, games, horror and learning. Each audio and video lasted for 1 minute. Make a record. The EEG data preprocessing process is as follows: using a 1-50 Hz bandpass filter to remove low frequency drift, ICA analysis, eliminating A1, A2 useless electrodes, and remaining 30 channels of EEG data. Divide each piece of data into labels every 2S, so each subject's EEG data is 30 segments.

4. Feature extraction

The feature extraction of EEG is mainly for de-evolving, dimensionality reduction and decorrelation. The commonly used feature extraction methods are divided into three categories: time domain frequency domain analysis, spatial domain analysis and nonlinear dynamic analysis. In this paper, spectrum analysis is used. In the future research, the other two types will be used, and will not be described here.

The EEG data is a time series signal. Generally speaking, the time domain representation is more vivid, the frequency domain analysis is more concise, and the analysis problem is more profound. Taking the frequency domain as the coordinate of various physical quantity lines and curves, various amplitude spectrum, phase spectrum, power and various spectral densities can be obtained, and the EEG signals and different waves of different rhythms can be more intuitively distinguished [5]. Continuous-time Fourier transform is suitable for the theoretical analysis of time-continuous signals. Since the function $f(t)$, $F(\omega)$ on both sides of the transformation is a continuous function, engineering applications often need to perform Fourier analysis on the sampled data. The numerical calculation method of the Fourier transform [6].

If the main value interval of $f(t)$ is $[t_1, t_2]$, define $T=t_2-t_1$ as the interval length. Sampling N points in this interval, the sampling interval is Δt=T/N, then there are:

$$F(\omega)=\sum_{n=0}^{N-1} f(t_1+n\Delta t)e^{-jw(t_1+n\Delta t)}\Delta t=\Delta t\cdot\sum_{n=0}^{N-1} f(t_1+n\Delta t)e^{-jw(t_1+n\Delta t)} \quad (1)$$

The above equation can calculate the Fourier transform value of any frequency point. If the main value interval of $F(\omega)$ is $[w_1, w_2]$, to calculate the k values of uniform sampling between them, there are:

$$F(w_1+k\Delta w)=\Delta t\cdot\sum_{n=0}^{N-1} f(t_1+n\Delta t)e^{-j(w_1+k\Delta w)(t_1+n\Delta t)} \quad (2)$$

Where $\Delta w=(w_2-w_1)/k$ is the frequency domain sampling interval.

After multiple screening, filtering and channel selection, the following spectrum analysis charts in different states are obtained. Figure 4 is the filtered EEG data image, and figure 5 is the analyzed spectrum result. It can be seen from Fig. 4 that the time domain diagram under different states can simply see different points, but the amplitude of each state is not obvious. Thus, the frequency-domain results in Figure 5 show that the amplitude is different in different states. In the pleasant state, the amplitude points are more fluctuating, the amplitude is larger under the excitement, the amplitude is second in the tension state, and the amplitude is smaller in the relaxed and quiet state.

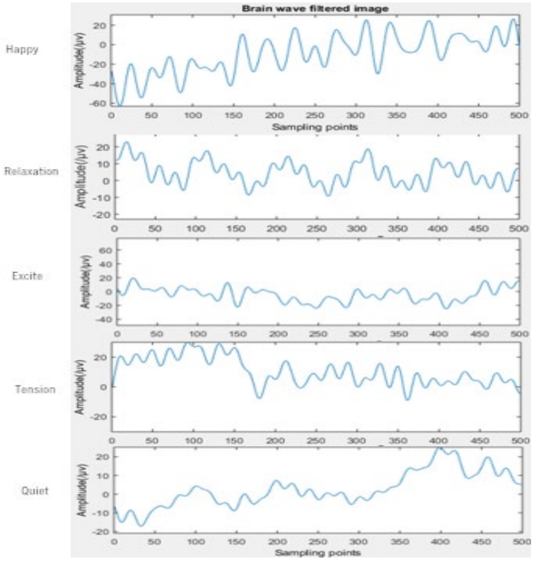


Fig4. Filtered EEG image

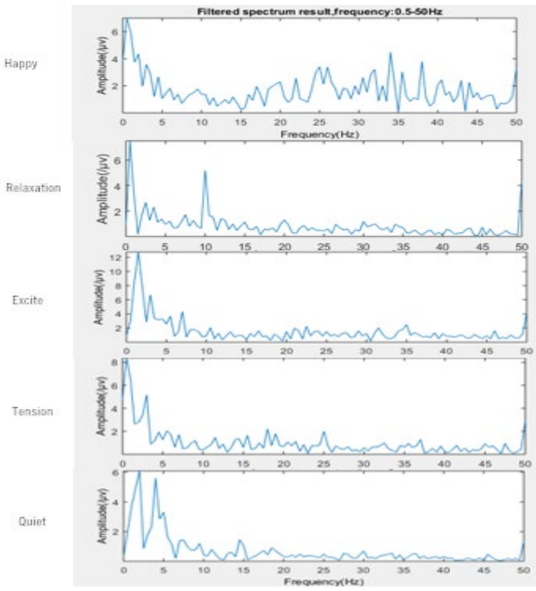


Fig5. Results of spectrum analysis

5. Classification

LDA can transform high-dimensional EEG data into low-dimensional space by means of mapping. In this low-dimensional space, the separability of data is the best, so that the distance between two types of data is maximized. The distance is minimized as much as possible to achieve dimensionality reduction and classification [7].

The general formulation of the problem is to have k totals $G_1, G_2, \cdots, G_k$ and the known sample X is taken from the K populations, and is one of them. It is now required to determine which total X belongs to. The basic idea of discriminant analysis to solve this problem is to first master the known knowledge of K populations and the observed values of indicators with discriminant significance of discriminant samples, and then find out the statistical path between the observed values of some indicators of sample x to be discriminated and the known knowledge according to the comparative analysis of the observed values and the known knowledge of populations, so as to determine the belonging population of sample X [8].

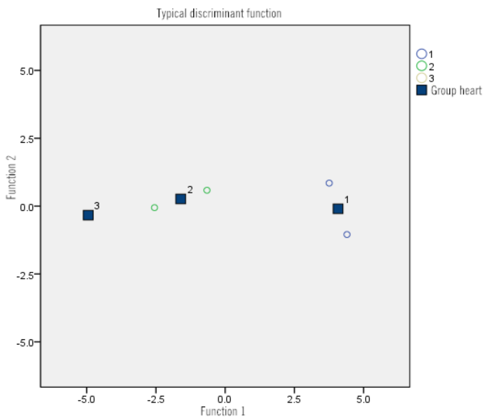


Fig6. Classification scatter plot

The five different emotions are divided into positive, intermediate, and negative levels, which are represented by 1, 2, and 3 respectively. According to the scatter plot made by the discriminant score in Fig. 6, it can be seen from the figure that the distinction between the three groups is still obvious.

6. Conclusion

Through the data collection of 5 subjects in 5 different states, through the spectrum analysis and LDA classifier classification verification, a satisfactory classification effect was obtained. On the whole, spectrum analysis is only applicable to single-channel data. Researchers are faced with multiple channel selections, which increases complex calculations and data processing; LDA classifiers are only for linear classification. In the later research, the multi-channel feature extraction, as well as the classification and screening comparison of multiple classifiers, improve the robustness and efficiency of EEG emotional classification.

Acknowledgements

The research is partly supported by the Project of Tianjin Enterprise Science and Technology Commissioner to Tianjin Tianke Intelligent and Manufacture Technology Co., Ltd (19JCTPJC53700). It is also supported by the Industry-University Cooperation and Education Project (201802286009) from Ministry of Education, China.

References

1. Koelstra, Sander, et al., Deap: A database for emotion analysis; using physiological signals, *IEEE transactions on affective computing* 3.1 (2011): pp.18-31.
2. Lin, Yuan-Pin, et al., EEG-based emotion recognition in music listening, *IEEE Transactions on Biomedical Engineering*, (2010): 1798-1806.
3. Bos, Danny Oude, EEG-based emotion recognition, *The Influence of Visual and Auditory Stimuli*, (2006): pp.1-17.
4. Nie, Dan, et al., *EEG-based emotion recognition during watching movies*, 2011 5th International IEEE/EMBS Conference on Neural Engineering. IEEE, 2011.
5. Akin, Mehmet, Comparison of wavelet transform and FFT methods in the analysis of EEG signals, *Journal of medical systems*, (2002): pp.241-247.
6. Yazdani A, Lee J S, Ebrahimi T, *Implicit emotional tagging of multimedia using EEG signals and brain computer interface*, Proceedings of the first SIGMM workshop on Social media. ACM, 2009: pp.81-88.
7. Subasi, Abdulhamit, and M. Ismail Gursoy, EEG signal classification using PCA, ICA, LDA and support vector machines, *Expert systems with applications*, (2010): pp.8659-8666.
8. Murugappan, Nagarajan Ramachandran, Yaacob Sazali, Classification of human emotion from EEG using discrete wavelet transform, *Journal of biomedical science and engineering*, (2010): pp.390.

Identification of Synthetic Pigment based on Fluorescence Spectroscopy Combined with RBF Neural Network

Di Yin [1,3], Fengzhi Dai [1,2,3]*, Yuxuan Zhu [1], Yasheng Yuan [1]

[1] *Tianjin University of Science and Technology, China;*

[2] *Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China*
[3] *Advanced Structural Integrity International Joint Research Center, Tianjin University of Science and Technology, Tianjin, 300222, China*
*E-mail: * daifz@tust.edu.cn*
www.tust.edu.cn

Abstract

Compared with natural pigments, the synthetic pigment is cheap as an important component of food additives, which has good stability, strong coloring ability, etc. But the consumption of excessive synthetic pigment will cause harm to the human body, so an effective method of detecting the pigment is needed. In this paper, 12 samples were selected from 1ug/ml to 100ug/ml different concentrations of carmine and amaranth, the fluorescence spectra of all the samples were measured by the LS-55 fluorescence spectrophotometer. The results showed that the fluorescence peaks were respectively around 450nm and 420nm, redshifts and the fluorescence intensity decreased with the concentration. Two kinds of pigment samples were predicted by RBF neural network model, and the results showed the accuracy of the samples classification was 100% and the error of the sample concentrations was very small, the RBF neural network provides a method for detecting the pigment of foods.

Keywords: Fluorescence spectroscopy; RBF neural network; Carmine; Amaranth red

1. Introduction

At present, there are many kinds of additives in food. Synthetic pigment is the most common one. It can reduce the price and adjust the bright color of food. It is called food dye or colorant, and has been widely used in food [1]. However, it will be absorbed by the human body after a large number of consumption of synthetic pigments, because most of them contain R-N = R-N 'bond and benzene ring, which will cause cancer in the human body for a long time [2]. Many countries in the world regard food safety as their own public safety, and have promulgated a series of laws to restrict the use of synthetic pigments.

In this paper, the fluorescence spectrum data of synthetic pigment is extracted to simulate and analyze the fluorescence spectrum of synthetic pigment, and then combined with RBF neural network for qualitative and quantitative analysis. The results can accurately identify and predict the types and concentrations of synthetic pigment of carmine and amaranth. This method can effectively reduce the cost, improve the accuracy of measurement data and shorten the time of analysis of characteristics, which is of great significance for food safety research.

The content of this paper is to produce fluorescence spectrogram of carmine and amaranth solution with different concentration under the optimal excitation wavelength by using fluorescence spectrograph, then use origin to process the graphs and analyze the fluorescence spectrograms corresponding to each concentration. RBF neural network was used for pretreatment, training and modeling to predict the type and concentration of synthetic pigments [3].

2. Experiment

2.1 Experimental device

This experiment uses the LS-55 spectrofluorimeter of PerkinElmer company of America, and the light source is pulsed xenon lamp. The cuvette is a 1cm quartz liquid cell with stopper. This consists of 5 units, namely, light source, monochromator, sample cell, detector and display.

2.2 Experimental method and sample configuration

In this experiment, GBW(E)100161 standard substances of rouge and amaranth solution (1.00ug/ml) and ultrapure water were used. Allocation carmine and amaranth solutions of 5ug/ml, 10ug/ml, 15ug/ml, 20ug/ml, 30ug/ml, 40ug/ml, 50ug/ml, 60ug/ml, 70ug/ml, 80ug/ml, 90ug/ml, 100ug/ml.

Determine the best excitation wavelength of cochineal red and amaranth red solution and set appropriate scanning parameters. In this paper, the excitation wavelength was set from 250nm to 350nm at a 5nm interval. The emission wavelength ranged from 200nm-700nm, the scanning speed was 1000nm/min, and the fluorescence spectra of the two solutions were obtained under this condition.

2.3 Data processing and analysis

Fluorescence spectra of the two solutions were measured and the characteristic parameters (excitation wavelength, emission spectrum, fluorescence intensity and solution concentration) of carmine and amaranth solutions with different concentrations were analyzed[4].

2.4 Fluorescence spectra of carmine red and amaranth red solutions

2.4.1 The fluorescence spectrum of cochineal

A fluorescence spectrophotometer was used to scan the prepared 5ug/ml carmine solution at a wavelength of 450nm. It can be seen from figure 1 that there are three fluorescence peaks, with wavelengths of 236nm, 240nm and 310nm respectively. The fluorescence peak is the highest at the wavelength of 236nm, that is, the optimal excitation wavelength of cochineal pigment is 236nm.

Cochineal pigment has characteristic fluorescence peak at wavelength of 440nm and fluorescence intensity of 660nm.

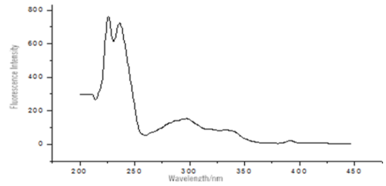


Fig1 Excitation spectrum of carmine

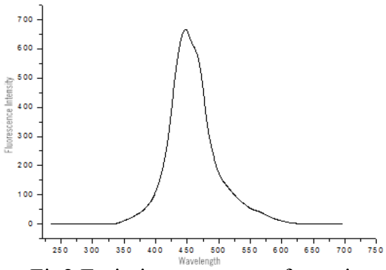
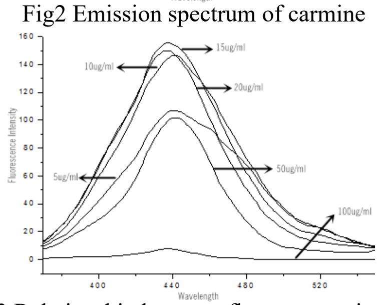


Fig2 Emission spectrum of carmine



Fir3 Relationship between fluorescence intensity and concentration

As shown in figure 3, there is a non-linear function between cochineal red solutions with different concentrations and fluorescence peak. To detect cochineal pigment, RBF neural network is required for modeling.

2.4.2 The fluorescence spectrum of amaranth

A fluorescence spectrophotometer was used to scan the prepared 5ug/ml carmine solution at a wavelength of 415nm. It can be seen from figure 4 that there are two fluorescence peaks, with wavelengths of 240nm and

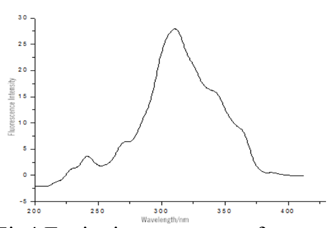


Fig4 Excitation spectrum of amaranth

310nm respectively. The fluorescence peak is the highest at the wavelength of 310nm, that is, the optimal excitation wavelength of amaranth is 310nm.

Amaranth has characteristic fluorescence peak at the wavelength of 340nm and 420nm, and the fluorescence intensity of 43nm.

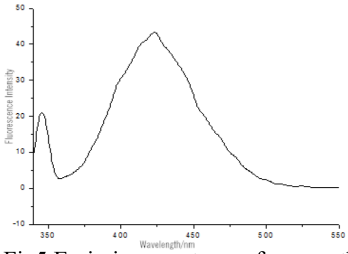
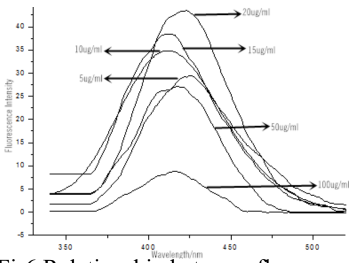


Fig5 Emission spectrum of amaranth



Fir6 Relationship between fluorescence intensity and concentration

It can be seen from Fig. 6 that there is a non-linear function between amaranth solution of different concentrations and fluorescence peak [5]. To detect carmine pigment, RBF neural network model is needed.

3. Pigment classification and concentration prediction based on neural network

3.1 Introduction of RBF neural network

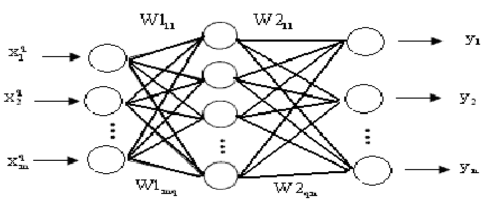


Fig7 RBF neural network structure

The input layer is composed of the nodes of the signal source; the output layer responds to the input mode; and the transfer function of the hidden layer uses the radial basis function (using the Gaussian function as the expression) [6]. The formula is as shown in (1) below:

$$f(x) = e^{-x^2} \tag{1}$$

3.2 Establishment of RBF Neural Network Model

After repeated experiments, the parameters of the RBF neural network are set as follows: The number of neurons in the input layer is set to 261, the number of neurons in the output layer is set to 1, and the precision of the best training of the neural network is 0.01.

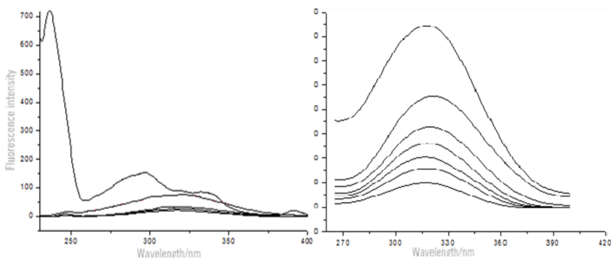
The emission fluorescence wavelengths of the two solutions were from 340~600nm, and one point was taken every 1nm. Further, 12 sets of spectral data processed are used as network feature values, the number of input layer neurons is 261, the number of output layer neurons is 1, the output target vector is the type of test sample, and 0 represents carmine, 1 represents amaranth. The trained RBF radial basis function neural network is used for recognition and prediction [7]. The prediction results are shown in Table 1. The identification accuracy is 100%, which confirms that the neural network has a good identification function.

Table 1 Identification results of carmine red and amaranth red pigment

sample	true value	predicted value	recognition rate
carmine	0	0	
carmine	0	0	100%
amaranth	1	1	

3.3 Prediction of pigment concentration in RBF neural network

As shown in Fig. 8(a), the fluorescence spectra of nine different concentrations of amaranth solution ranged from 200 to 598; therefore, we need to select the band by dividing the entire spectral interval into n. The wide sub-intervals are then subjected to regression processing for each sub-interval. As shown in Fig. 9, the best band is 267-389, which can better reflect the relationship between different concentrations of amaranth solution



(a) Original spectra (b) Processed spectrogram

Fig8 Fluorescence spectrum of amaranth

and fluorescence intensity. As shown in Figure 8(b), using this data can reduce the modeling error.

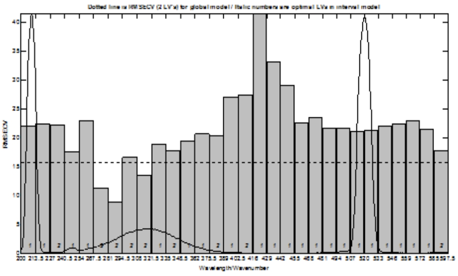


Fig9 Band selection diagram

In this experiment, the neural network toolbox in MATLAB software was used to establish the RBF neural network model to predict the concentration of synthetic pigments. After repeated experiments, the parameters of the RBF neural network were set as follows: Radial basis function The scatter constant (SPREAD) is taken as 1.8, the RBF neural network performance function is set to mean square error (the purpose is to minimize the error) [8], and the neural network training accuracy is 0.001.Two sets of data of 20ug/ml and 60ug/ml of carmine solution were taken as samples and substituted into the trained RBF neural network model for concentration prediction. The experimental results are shown in Table 2.

Table2 Prediction results of amaranth solution (unit:ug/ml)

sample	actual concentration	predicted concentration	error
1	20	26.233	-6.23
2	60	55.096	4.9

The experimental data shows that the concentration prediction function of the RBF neural network is excellent and the accuracy is very high, and its reference value and the predicted value can be in a good linear relationship as shown in Fig10, and the correlation coefficient reaches 0.9997.

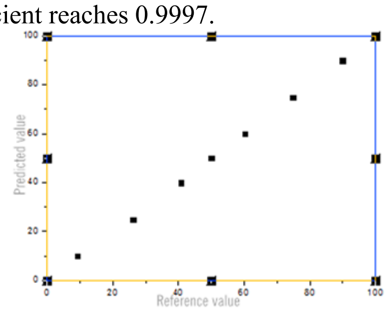


Fig10 Correlation between reference value and prediction value of amaranth solution

4. Conclusion

This study demonstrates that the use of fluorescence spectroscopy combined with RBF radial basis function neural network can accurately, efficiently and quickly identify synthetic pigments. This method is suitable for identifying pigments and other fluorescent substances in various synthetic foods. Species identification and quantitative determination provide a good method for food safety identification.

Acknowledgements

The research is partly supported by the Project of Tianjin Enterprise Science and Technology Commissioner to Tianjin Tianke Intelligent and Manufacture Technology Co., Ltd (19JCTPJC53700). It is also supported by the Industry-University Cooperation and Education Project (201802286009) from Ministry of Education, China.

References

1. Yin Di, Dai Fengzhi, et al, *Cluster Analysis of Wine based on Three-dim Fluorescent Spectra Characteristic*, Proceedings of the 2019 International Conference on Artificial Life and Robotics (ICAROB2019), Oita, Japan, 2019, pp.280-283.
2. Huang X D, Wang C Y, Fan X M, et al. Oil source recognition technology using concentration-synchronous-matrix-fluorescence spectroscopy combined with 2D wavelet packet and probabilistic neural network. *Science of The Total Environment*, 2018, 616: pp.632-638.
3. Sahar, Amna, et al., Monitoring of thermal changes in meat by synchronous fluorescence spectroscopy, *Journal of food engineering* 168 (2016): pp.160-165.
4. Gomez, G. F., Eckert, G. J., Zandona, A. F, Orange/red fluorescence of active caries by retrospective quantitative light-induced fluorescence image analysis. *Caries research*, 2016, 50(3), pp.295-302.
5. Gu, Wei, et al., Black phosphorus quantum dots as the ratiometric fluorescence probe for trace mercury ion detection based on inner filter effect, *ACS sensors*, (2017): pp.576-582.
6. Flyer, Natasha, et al., On the role of polynomials in RBF-FD approximations: I. Interpolation and accuracy, *Journal of Computational Physics*, (2016): pp.21-38.
7. Zhu, Z-W., et al., Discrimination of Brands of Strong Aroma Type Liquors Using Synchronous Fluorescence Spectroscopy and Chemometrics Methods, *Journal of Applied Spectroscopy*, (2019): pp.1101-1106.
8. Shu-tao, Wang, et al., Concentration Detection of Polycyclic Aromatic Hydrocarbon Combining Three-dimensional Fluorescence Spectroscopy with HGA-RBF Neural Network, *ACTA PHOTONICA SINICA* , 2017.

Simulation study of 3D reconstruction in Electromagnetic Tomography with Two-layer Sensors Array

Yuanli Yue, Ze Liu, Yu Miao*

School of Electronic and Information Engineering,
Beijing Jiaotong University, China

*E-mail: *ymiao@bjtu.edu.cn*
www.bjtu.edu.cn

Abstract

Electromagnetic tomography (EMT) is an emerging technology of non-destructive testing, which is used in industrial process monitoring and biomedical detection due to its invasive and non-contacting nature. A volume imaging using two layers sensor array, namely, 3D-EMT, has been developed in this paper. The paper provides the simulation study of 3D reconstruction in electromagnetic tomography with two-layer sensors array. Various models of the sensor array with different layers and different numbers of coils are established in FEM software—COMSOL Multi-Physics. The simulation results show that 3D images have advantages in position, shape, and size than 2D images. The results confirm the feasibility of 3D reconstruction in electromagnetic tomography with Multi-layer Sensors Array.

Keywords: electromagnetic tomography; multi-layer sensors array; FEM software; 3D reconstruction;

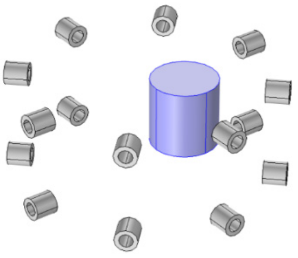
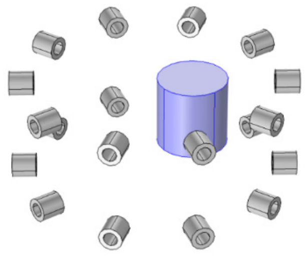
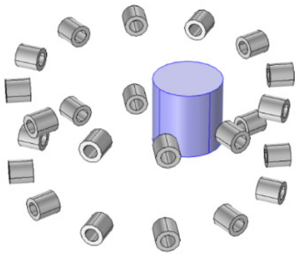
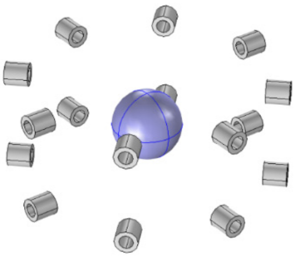
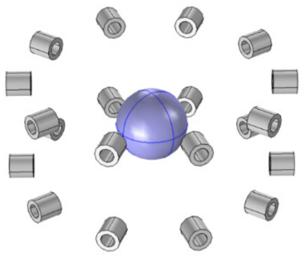
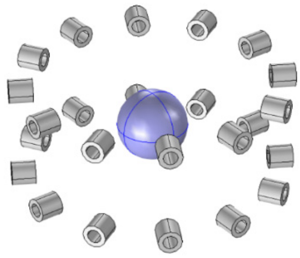
1. Introduction

Progress tomography (PT) is a novel type of Non-Detection Technology, which is widely used in many industrial and biomedical areas, such as industrial progress monitoring, mining and medical tomography. PT mainly includes Electromagnetic Tomography (EMT)[1,2], Electrical Impedance Tomography (EIT)[3] Electrical Resistance Tomography (ERT)[4] and Electrical Capacitance Tomography (ECT)[5,6]. EMT is a relatively new member of the progress tomography family, which works based the eddy current in conductive objects.[7] Among these tomography techniques, EMT is considered the most promising technique for non-detection imaging, because of its contactless with the object.

Presently, most of the EMT systems are still in two-dimensional scale. The 3D reconstruction was researched in ECT and in EIT firstly, then EMT is developed later.[8] A method about time-difference imaging of magnetic induction tomography in a three-layer brain physical phantom is proposed, it is able to simulate the conductivity distribution of the main structures of a human brain. It is a step forward towards a realistic brain phantom for MIT.[9] A method of volumetric (3D) imaging of passive electromagnetic properties using low frequency electrical and electromagnetic imaging, and the standard Tikhonov regularization method have been used in this technique.[10] By using this method, volumetric imaging can be constructed and hence can improve the spatial resolution of the reconstructed images.[11]

This paper presents the results of 3D simulations with multi-layer sensors array. Compared with 2D images, 3D images can get more information about the distribution of electromagnetic properties; conductivity, permittivity and permeability. It can display the space structure of objects with more information in the z-direction than cross-section with 2D images.

Table 1. Different models and different distributions with different number of coils using COMSOL Multi-physics.

	16-coil model	20-coil model	24-coil model
Distribution 1			
Distribution 2			

2. Forward problem

The forward problem in EMT is based on the eddy current, which is a multiple coil eddy current method.[12] The sinusoidal signal is applied to the coil array, producing a primary magnetic field B. Then the magnetic field interacts with conductive and permeable objects, which can produce eddy currents, and produce a secondary magnetic field ΔB. Through the mixture of the primary magnetic and secondary magnetic fields (B+ΔB), the induced voltages are acquired through other coils.

The fundamental formulation for solving the forward problem is the Maxwell equation group (1).

$$\begin{cases} \nabla \times H = J_e + J_s \\ \nabla \times B = -j\omega B \\ \nabla \cdot B = 0 \\ \nabla \cdot D = 0 \end{cases} \quad (1)$$

In this equation, J_e is the eddy current density, J_s is the source current density, H is the magnetic field strength, B is the magnetic flux density; and D is the electrical displacement vector.

As for the object in the detection field, this equation group is established as (2).

$$\begin{cases} D = \varepsilon E \\ B = \mu H \\ J_e = (\sigma + j\omega)E \end{cases} \quad (2)$$

Where ε is the permittivity, μ is the permeability and σ is the conductivity and E is the electric field intensity. Combining (1) and (2) we obtain (3).

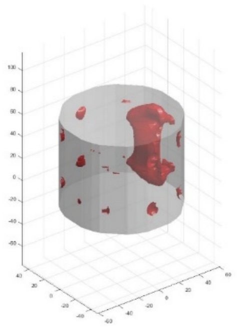
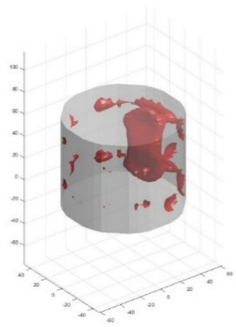
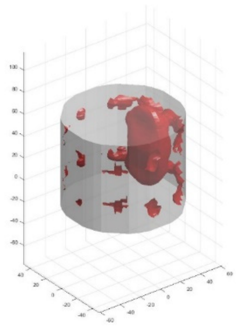
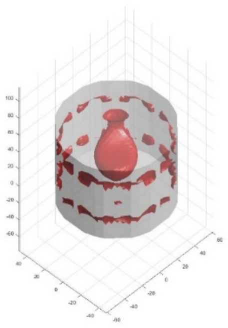
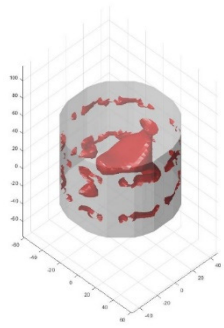
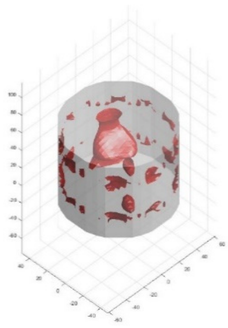
$$\frac{\nabla^2 A}{\mu} - j\omega\sigma A = -J_s \quad (3)$$

A is the magnetic vector potential. The induced voltage of the detection coil can be described as (4).

$$\mathrm{V} = -\mathrm{j}\omega \oint A dl \quad (4)$$

The induced voltage is important to solve the forward problem, which is the most important step to observe the change between the homogeneous field and inhomogeneous field.

Table 2. The constructed images with different number of coils using the conjugate gradient iterative algorithm.

	16-coil reconstruction	20-coil reconstruction	24-coil reconstruction
Distribution 1			
Distribution 2			

In this paper, different kinds of models are stablished. The simulated systems have 16-coil, 20-coil and 24-coil respectively, used for excitation and detecting. Table 1 shows the different type of the sensor arrays and two kinds of distributions. In distribution 1, the diameter of the cylinder is 30 mm, located at (20, 0, 20). In distribution 2, the copper ball with a 30 mm diameter, centered at (0, 0, 20).

In the forward problem, each coil is 200 turns using copper wire of diameter of 0.18mm and have a 6 mm inner and 10mm outer diameter. The region of interest for the imaging is a cylinder with diameter 100mm, centered at (0, 0, 20).

3. Inverse problem

The inverse problem is to acquire the conductivity distribution from the voltage measurements that were obtained from the forward problem, which is a non-linear and ill-conditioned problem.[13] In order to solve the inverse problem, a linear approximation method is implemented by regarding the conductivity distribution as having no relationship with the sensing field function, which is caused by the soft field effect. After discretization, linearization, the linear equation between the induced voltage and conductivity distribution is shown as follows (5):

$$\mathrm{U} = \mathrm{Sg} \tag{5}$$

Where U is the normalized measurement voltage vector, S is the normalized sensitivity matrix and g is the normalized conductivity distribution vector.

For electrical tomography, the inverse problem is ill-posed, as there are less limited independent measurements than unknown voxel values, which cause the inverse matrix to be incalculable. In order to use the conjugate gradient iterative algorithm, equation (6) is multiplying both sides by S^T.

$$S^T\mathrm{U} = S^T S g \tag{6}$$

The conjugate gradient iterative algorithm (CG) is a well-known and effective global iterative method that is widely used in electrical tomography. CG is the most popular iterative method for solving large systems of linear equations. The algorithm is aimed at finding the minimum of the function. It can be described as follows (7):

$$\mathrm{f(g)} = \frac{1}{2} g^T (J^T J) g - (J^T U) g \quad (7)$$

Where J is a Jacobian matrix, U is the voltage measurement and g is the normalized conductivity distribution vector.

In this paper, the modelling was established and study in COMSOL multi-physics, which is a famous software for FEM. Numerical simulation was solved in MATLAB, which is a common software for mathematical analysis, through combining COMSOL multi-physics and MATLAB to get the reconstruction images. The results are presented in 3D imaging of electrical conductivity imaging using multi-layer sensor array. Table 2 shows the 3D reconstruction images with different models – 16 coils, 20 coils and 24 coils. These images are the results of 30 conjugate gradient iterations. The images demonstrate the conductivity distribution in the detection field.

For 3D images, results show the accuracy of positon and shape of objects, as well as the spatial distributions along the z-direction. It can distinguish and compare the difference between the original distribution of the copper ball and the reconstruction images clearly. The 3D images has a more vivid views than 2D images.

4. Conclusion

In this paper, a volume imaging using two-layer sensor array, namely, 3D-EMT, has been developed. For the forward problem solving, different models 16-coil, 20-coil, 24-coil, were established and the 3D imaging technique based on the EMT principle is demonstrated. In the inverse problem, the conjugate gradient iterative algorithm was chosen to reconstruction images with the results of 30 iterations. The 3D EMT images not only can show us the shape and position of the object, but also demonstrated the information about z-direction, which could reflects the conductivity distribution along the z axis.

The result can give a guideline to the practical system in non-detection field. For example, 3D EMT reconstruction can help a doctor judge the brain tumor's position, shape and size, which is a revolution in medical tomography.

Acknowledgements

This work was supported by the National Natural Science Foundation of China under Grant 61771041.

References

1. X. Ma, A. J. Peyton, R. Binns, and S. R. Higson, Electromagnetic techniques for imaging the cross-section distribution of molten steel flow in the continuous casting nozzle, *IEEE Sensors J.*, 2005.pp. 224–232.
2. Yin X, Wu H, Jia J and Yang Y, A micro eit sensor for real-time and non-destructive 3-d cultivated cell imaging, *IEEE Sensors J.*, 2018, pp. 99 1-1
3. Yin X, Wu H, Jia J and Yang Y, A micro eit sensor for real-time and non-destructive 3-d cultivated cell imaging, *IEEE Sensors J.,* 2018, pp. 99 1-1
4. Ren S, Zhao J and Dong F, Dimensionality reduced simultaneous iterative reconstruction technique for electrical resistance tomography Flow Measurement and Instrumentation 2015, pp. 284-291.
5. L. Peng, J. Ye, G. Lu, W, Yang, Evaluation of effect of number of electrodes in ECT sensors on image quality, *IEEE Sensors J.*, 2012, pp. 1554-1564.
6. Wang A, Marashdeh Q, MOTIL B J, FAN L S Electrical capacitance volume tomography for imaging of pulsating flows in a trickle bed, *Chemical Engineering Science* 2014, pp. 77-87.
7. Biro O and Preis J, On the use of the magnetic vector potential in the finite element analysis of three-dimensional eddy currents, *IEEE Trans. Magn.*,1989, pp. 3145–59.
8. Soleimani, M., Sensitivity maps in three-dimensional magnetic induction tomography, Insight, 2006, pp 39-44.
9. Ruigang Liu, Ye Li, Feng Fu, et al. Time-difference imaging of magnetic induction tomography in a three-layer brain physical phantom. *Meas. Sci. Technol.*, 2014, pp. 788-92.
10. Wei H, Yang C, Ma L, et al. Three-dimensional industrial process tomography using electrical and electromagnetic tomography : recent developments,2012.
11. Wei, H.-Y., L. Ma, and M. Soleimani, Volumetric magnetic induction tomography, *Meas. Sci. Technol.*, 2012, pp. 055401-55409.
12. Tan C, Wu Y R, Xiao Z L and Dong F Optimization of dual frequency-difference MIT sensor array based on sensitivity and resolution analysis, *IEEE ACCESS*, 2018, pp. 34911-34920.
13. Xiao Z, Tan C, Dong F, Effect of inter-tissue inductive coupling on multi-frequency imaging of intracranial hemorrhage by magnetic induction tomography. *Meas. Sci. & Technol.*, 2017, 084001.

Visualized the knowledge map in children's minds: A study on cognitive structure measurement

Qiang Wei[1*], Hua Dong, Yitong Zhang, Aonan Zhang

School of Education, Jianghan University, China;

*E-mail: * weiqiang@jhun.edu.cn*

Abstract

How were concepts organized in student's minds? The aim of this study was to explore a measurement based on the spreading activation model (SAM) and multi-dimensional scaling (MDS) to visualize student's cognitive structure, which helps to accurately locate the personalized model of learners and promote personalized learning, and helps teachers to effectively teach. 120 junior high school students (60 from Grade 7; 60 from Grade 9) in this study. Based on ranking of physics achievements, students were divided into three groups (top, average, poor). Fifteen concepts were selected from the junior high school physics textbook, to compare the cognitive structure maps of different grades and academic levels. The results indicate that the cognitive structure of Grade 9 is more closely related to that of Grade 7, there are significant differences in the mastery of kinematic knowledge points between the top group and poor group.

Keywords: cognitive structure; visualization; multidimensional scale analysis; personalized learning

1. Introduction

With the explosive growth of data scale, the world has entered the era of big data[1]. AI + education is the application of artificial intelligence in education, and it has become a research hotspot. The application of various educational robots promotes the development of science and Engineering in teaching[2].

Many researches show that cognitive structure plays an important role in students' understanding of knowledge[3, 4]. Cognitive structure is the basis of meaningful learning, and it is helpful for teachers to organize materials by identifying learners' cognitive structure[5]. Appropriate materials can rationalize the teaching process, optimize the cognitive structure of students, and promote the development of them[6]. The study can help teachers better understand students' mastery of knowledge points, while help students to rationalize their knowledge structure. Intelligent Tutoring Systems (ITSs) can simulate teachers to realize one-to-one intelligent teaching, which is a typical application in the field of AI + education[7]. According to the characteristics of each student, this study can provide a theoretical basis for the realization of personalized learning[8].

The cognitive ability of human brain is realized by concepts and the connection between them[9]. In junior high school, the concept of physics is abstract and some laws are not verifiable, which is considered as a difficult subject for students to study. It is very important for middle school students to master physical concepts.， and the in-depth study of the concept is useful to improve their academic performance and motivation. The physical cognitive structure in students' mind is a subjective experience system accumulated in the learning process, reflecting the learners' understanding of physical knowledge. At present, the paper and pencil test is a main way for school to evaluate students' knowledge mastery level, but it can not directly and comprehensively reflect the cognitive structure in students' minds.

As an exploratory data analysis method, multi-dimensional scaling(MDS) can reduce the dimension of the subjective qualitative data of the input evaluator, and output the visualized three-dimensional spatial structure chart, so as to explore the potential laws contained in the data intuitively[10]. Through the semantic features stored in the long-term memory, the participants scored the similarity between the two concepts, which indicated the distance between the two concepts in the participants' mind.

In the spreading activation model (SAM), the concept of words is represented by nodes, while the distance between two nodes indicates the degree of connection between them.

Based on the theoretical basis of MSD and SAM, it is the goal of this study to visualize the intelligent education system, which helps to accurately locate the personalized model of learners, and helps teachers to effectively teach and improve the learning effect of students.

2. Experiment

2.1. Stimuli

Total 15 basic concepts of electricity, mechanics and kinematics in junior high school physics textbook are used. They are: charge, current, voltage, resistance, work, power, gravity, friction, mass, Newton, pressure, distance, time, velocity, acceleration.

2.2. Participants

120 students（including 60 participants in Grade 7, 60 participants in Grade 9）and 2 physics teachers from the Middle School in Wuhan were participated the experiment. According to the ranking of physics achievements, 60 seniors were divided into three groups : 20 in top group; 20 in average group; 20 in poor group.

2.3. Procedure and data acquisition

The experimental materials include presentation of PowerPoint and answer sheet, making up 105 pairs of "stimulus pairs" which are arranged in random order, with instructions and 105 5-point scales printed on the answer sheet.

Class as a unit, and participants were measured in group. The researchers distributed the answer sheet to the participants, explained the experimental requirements to them, and then broadcast the experimental materials through PowerPoint.

The participants were required to judge the degree of connection between the two words presented on the screen. 1-5 means that the degree of connection gradually weakened. In the experiment, the participants could rest for 5 minutes. Take back the answer sheet on the spot after the test.

2.4. Data analyses

Input 105 scores on the answer sheet into Excel, and then establish the 15 * 15 dissimilarity matrix of each subject. According to different levels of independent variables, the multi-dimensional scale analysis was used for statistical processing to obtain the spatial structure chart.

Average nearest neighbor distance method: calculate the distance between each point and its nearest neighbor, and then calculate the average value of the nearest neighbor distance of all points. For each point, the nearest neighbor is determined according to its Euclidean distance.

The average nearest neighbor distance reflects the distribution of points in space. The smaller the nearest neighbor distance is, the denser the spatial distribution of the points is, on the contrary, the more discrete the points are. Spss21.0 was used to test t-test and analysis of variance for participants of different grades and academic levels.

MDS usually uses Euclidean distance to express the degree of connection between the two concepts, and the smaller the value is, the closer the connection between the two concepts is. According to the Euclidean distance between 105 pairs of concepts output by MDS, spss21.0 was used to conduct t-test and variance analysis for participants of different grades and learning levels.

3. Results

3.1. Spatial cognitive structure in grade 7 and grad 9

The spatial cognitive structure results of MDS showed difference between the grad 7 and the grad 9 participants (Fig.1, Fig.2). The fitting Stress of the grad 7 and the grad 9 models is 0.1789, 0.1528, and the goodness of fit RSQ values are 0.968, 0.977, shows well model fitting.

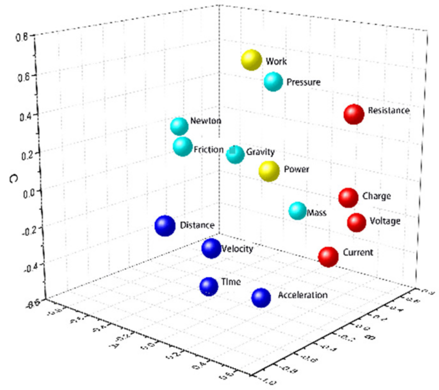


Fig .1 the spatial structure of the Grade 7 students

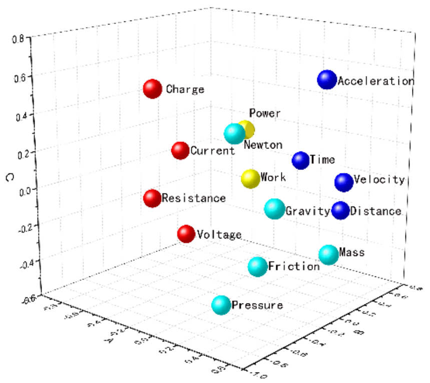


Fig.2 the spatial structure of the Grade 9 students

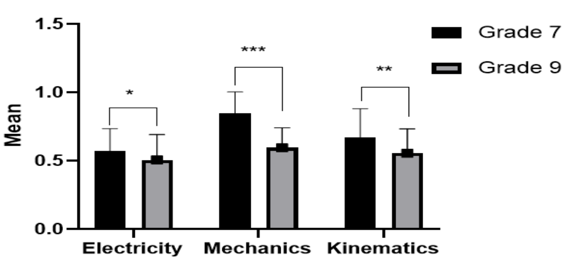


Fig.3 Significant difference of mean nearest neighbor distance between grad 7 and grad 9 ($p<0.05$ *, $p<0.01$ **, $p<0.001$ ***)

The average nearest neighbor distance of participants in different grades and knowledge points (electricity, mechanics, kinematics) was tested by independent sample t-test (Fig.3).

105 pairs of concepts in different grades were tested by independent sample t-test, and the results showed that (Table 1): only the concept pairs with significant differences were selected ($p < 0.05$).

Table 1 The t-values for Euclidean distance of concept pairs between grad 7 and grad 9

Concept pair	t	Concept pair	t
Current-voltage	3.028*	**Friction-mass**	8.157***
Current-resistance	3.603***	**Work-power**	5.024***
Voltage-resistance	3.815***	**Friction-pressure**	3.718***
Charge-work	5.883***	**Friction-Newton**	8.765***
Current-work	5.891***	**Mass-Newton**	9.46***
Voltage-work	2.14*	**Newton-pressure**	4.997***
Resistance-work	6.36***	**Distance-velocity**	3.269**
Charge-power	2.333*	**Time-velocity**	3.447**
Current-power	2.37*	**Distance-acceleration**	3.597***
Resistance-power	2.139*	**velocity-acceleration**	3.183**

*p<.05, **p<.01, ***p<.001

3.2. Spatial structure between academic levels

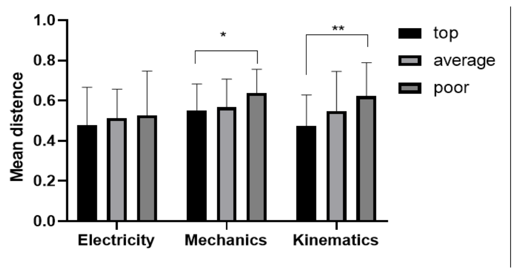


Fig.4 mean nearest neighbor distance with different academic levels ($p<0.05$ *, $p<0.01$ **)

One-way ANOVA was used to analyze the average nearest neighbor distance of participants at different academic levels (Fig.4). According to the multiple comparison of variance analysis results, the probability of the difference between the students with excellent learning ability and those with poor learning ability in general, mechanics and kinematics knowledge points are 0.014, 0.042 and 0.009 respectively, shows significant difference.

One-way ANOVA was used to analyze the Euclidean distance between 105 pairs of participants with different academic levels. The results showed that only the concept pairs with significant differences were selected (Table 2).

Table 2 Euclidean distance of concept pair with different academic levels

Concept pair	Groups	Mean	SD	F
Work-power	top	0.591	0.164	6.583**
	average	0.664	0.165	
	poor	0.79	0.195	
Friction-pressure	top	0.66	0.22	5.22**
	average	0.715	0.239	
	poor	0.49	0.23	
Distance-velocity	top	0.396	0.202	5.08**
	average	0.503	0.286	
	poor	0.629	0.198	
Current-work	top	0.654	0.181	6.267**
	average	0.708	0.218	
	poor	0.876	0.219	
Voltage-work	top	0.768	0.216	8.006**
	average	0.771	0.24	
	poor	1.029	0.25	

*p<.05, **p<.01, ***p<.001

According to the multiple comparison of variance analysis results, there are significant differences between poor learning, top learning and average learning groups in the concept of “work-power”, “work-gravity”, “friction-pressure”, “voltage-work”, “current-work”. While there is a significant difference between the poor students and the top students in the concept of distance-velocity, but there is no significant difference between them.

4. Discussion

4.1. Analysis results of participants in different grades

According to the nearest neighbor distance, the concepts are classified into four categories: electricity (charge, resistance, voltage, current), work-power, mechanics (gravity, friction, mass, Newton, pressure), kinematics (distance, time, velocity, acceleration).

It can be seen that the knowledge structure of the Grade 9 is clearly classified, and the overall structure of the Grade 7 is relatively loose. The calculation results show that the average nearest neighbor distance of the Grade 7 in the whole and part is significantly greater than that of Grade 9. This also shows that the knowledge structure of junior three is more intensive than that of junior one. This should be that the learning of the concept of junior one is still in the initial stage, and the understanding of the concept is not enough, and the relationship between the concepts is still confused.

According to the results of Euclidean distance calculation, the semantic distance between the electricity and kinematics of Grade 7 is significantly greater than that of Grade 9. The result indicated that the participants can roughly classify based on understanding, but do not understand how to transform and build concepts.

4.2. Analysis results of participants with different academic levels

According to the nearest neighbor distance, the concepts can be classified into four categories: electricity (charge, resistance, voltage, current), work-power, mechanics (gravity, friction, mass, Newton, pressure), kinematics (distance, time, velocity, acceleration).

From the spatial structure diagram, we can see that the knowledge structure of the students with excellent, average and poor learning is similar, and the spatial distribution of each concept in different knowledge points is different. According to the results of the spatial dispersion calculated by the average nearest neighbor distance, it shows that there are significant differences in the overall, mechanical and kinematic knowledge points between the students with excellent learning ability and the students with poor learning ability. This may be because the kinematics is at the knowledge points learned in the second year of junior high school, and the students with poor learning ability do not review in time and forget them, resulting in unclear concept understanding, while the mechanical part is the key and difficult part of physical learning, so it is necessary to focus on the mechanics and kinematics.

According to Euclidean distance, the results of 105 pairs of concepts of different academic level participants show that: The distance between the concepts of "work-power", "friction-pressure", "distance-velocity", "current-work" and "voltage-work" of the students with poor academic performance is significantly farther than that between the students with excellent academic performance and the students with average academic performance, which may be due to the lack of practice and the difficulty in transforming concepts and formulas.

5. Conclusion

The cognitive structure shows spatial difference between grad 7 students and grad 9 students. The physical cognitive structure of the students in Grade 9 is hierarchical, clear and integrated, which proves the feasibility of MDS analysis in the measurement of students' implicit cognitive structure.

There are significant differences in the dispersion of spatial structure map in grades and academic levels, and the cognitive differences in different knowledge points can also be analyzed intuitively through spatial structure map.

The test results of the connection degree between the two concepts show that there are significant differences in the understanding of the concepts of work-power, work-gravity, friction-pressure, distance-velocity, current-work and voltage-work between the top students and the poor students, so that teachers can carry out targeted teaching for different types of students according to specific concepts and improve teaching efficiency.

This study provides the possibility for visualizing the intelligent education system to realize the personalized learning, which helps teachers with targeted teaching.

References

[1] TIMES N Y. The Age of Big Data. 2012,
[2] BENITTI F B V, SPOLAôR N. How have robots supported STEM teaching?. 2017.
[3] Dozois D J A , Frewen P A . Specificity of cognitive structure in depression and social phobia: A comparison of interpersonal and achievement content. *Journal of Afferctive Disorders*, 2006, 90(2-3):101-109.
[4] Heinz A J , Veilleux J C , Kassel J D . The role of cognitive structure in college student problem drinking. *Addictive Behaviors*, 2009, 34(2):212-218.
[5] Ifenthaler D , Masduki I , Seel N M . The mystery of cognitive structure and how we can detect it: tracking the development of cognitive structures over time. *Instructional Science*, 2011, 39(1):41-61.
[6] NAKIBOGLU C. Using word associations for assessing non major science students'knowledge structure before and after general chemistry instruction: the case of atomic structure. *Chemistry Education Research and Practice*, 2008, 9(4):309.
[7] Hwang G J , Kuo F R , Yin P Y , et al. A Heuristic Algorithm for planning personalized learning paths for context-aware ubiquitous learning. *Computers & Education*, 2010, 54(2):404-415.
[8] ASWANI KUMAR C, ISHWARYA M S, LOO C K. Formal concept analysis approach to cognitive functionalities of bidirectional associative memory. *Biologically Inspired Cognitive Architectures*, 2015, 12:20-33..
[9] Hwang, Gwo-Jen. Definition, framework and research issues of smart learning environments - a context-aware ubiquitous learning perspective. *Smart Learning Environments*, 2014, 1(1):4.
[10] SAEED N, NAM H, HAQ M I U, et al. A Survey on Multidimensional Scaling. *Acm Computing Surveys*, 2018,51(3): 1-25.

Escape route in subway under fire emergency: An experimental study in virtual reality environment

Hua Dong[1*]; Qiang Wei[1]; Qingqing Zhang[1]; Lanlan Fang[1]; Weiran Cui[1]

[1] *School of Education, Jianghan University, China;*

E-mail: angela-dong@jhun.edu.cn

Abstract

The goal of present study was to investigate gender differences stress level and indicator searching behavior in the fire emergency in subway station. Sixty undergraduates were selected as participants and divided into two groups based on gender (21 in male group, 39 in female group). In the experiment, virtual environment was used to simulate a fire emergency to arouse participants' stress response. Participants were required to escape from the fire as soon as possible. Eye tracking was used to assess participants' indicator searching behavior during escape. The results showed that female group had significantly higher stress level than male group. Female also took longer time to look for helpful indicator from surroundings in order to find an escape route, while male tended to be more focused on indicator upside. The results suggest that fire evacuation system for subways (or other seal off environment) should be designed by taking into consideration the stress response characteristics of different genders.

Keywords: Stress Level; Indicator Searching; Fire Emergency; Virtual Environment

1. Introduction

Way-finding is the process of determine the route from a starting point to a successfully arriving at the destination. In emergencies, it is important for people to find an exit quickly and correctly.

Under stress, people's range of consciousness will change, which will affect their choices in emergency situations. However, the study on fire emergency research faced many difficulties. On the one hand, it is impossible to conduct experiments in real situation because of safety and ethics; on the other hand, it is limited to bring real experience to the participates under the laboratory condition. After considering the limitations to restore people's real stress-level under laboratory condition, a virtual environment (VE) technique is used as an effective and safe tool to simulate a real fire situation. The efficacy of the VE in the way-finding-related experiments has been confirmed by empirical studies. It was showed that the VE was suitable for emergency simulations and could be used as an effective training tool for fire escape training.

In engineering design, the restraint of human factors should always be considered. Such as the speed of visual information search, scope and its rapidity. Influenced by the surrounding environment, people's use of signs in emergencies will also affect their choices. Another difficulty in this study is to estimate sign searching behavior since it is hard to recognize fixation point and track eye movement under traditional experiments. Based on these, the study used eye tracker to record eye movement trajectory. Eye movement index can be used to detect human visual information extraction and visual control problems. Through the eye movement research on the sign searching behavior of people in emergency like fire, reasonable suggestion can be provided for the sign setting of subway station, so that people can escape from the disaster scene more easily and effectively in emergency situations.

Above all, the process of finding is influenced by stress-level and selective attention. According to previous researches, both stress-level and selective attention has significant difference between male and female[1,2]. Therefore, researchers aim at estimating gender difference of stress-level and sign searching behavior in a fire emergency and figuring out the interaction between them.

2. Method

2.1. Participants

Sixty participants (21 males and 39 females) were recruited from Jianghan University to take part in the study. The participants had normal visual acuity or corrected visual acuity, and all other physiological indicators were normal.

2.2. Materials

In the experiment, virtual subway station in fire emergency environment was set up by Unity. This VE consisted of multi-sensory stimuli including virtual fire and smoke on monitor, fire alarm and hot air. They were to create a relatively high-fidelity virtual fire environment and provide a stress to participants (Fig.1). The eye-tracking data of participants were recorded with SensoMotoroc Instruments (SMI), Teltow, Germany, ETG 2.7 eye-tracking device. (SensoMotoroc Instruments, Teltow, Germany). A questionnaire on a five-level Likert scale from 1(strongly disagree) to 5 (strongly agree) which contained 11 items aimed to describe participants' decision-making process during escaping (e.g. "I chose the escape route depended on indicator information around"; "I made decision on intuition").

Fig. 1. Fire emergency in VE subway station

2.3. Procedure

Upon arrival, participants were informed the notes of escaping task. In the warm-up stage, the participants were instructed to become familiar with the control methods in the Unity system. The participants were allowed to practice until they could smoothly control their movements. They were required to complete escaping task according to the notes they'd read before. The eye-tracking device was set up to record the participants' eye movement data during the formal experiment.

In the formal experiment, participants were initially set in a random location to start escaping task, they were required to escape in finite time. Finally, participants completed the questionnaire to describe their decision-making process during escaping task.

After finishing the escape experiment, all participants were required to answer the questionnaire about decision-making process during escaping.

3. Results

3.1. Questionnaire results of decision-making

The pressure respond level was calculated and compared between male group and female group by a independent sample t-test. The results showed that the pressure respond level in female group (mean=3.9, SD=1.071) was significantly higher compared to that in male group (mean=2.76, SD=1.446), $p<0.05$ (Table 1.).

The results for questionnaire on decision-making were calculated and compared between male group and female group. The independent sample t-test results showed that it was significantly higher in female group (mean=3.52, SD=1.209) compared to that in male group (mean=2.79, SD=1.380), $p<0.05$ (Table 1.).

3.2. Eyes movement data analysis

The distributions of specific eye-tracking fixation points were shown in Fig. 2. The comparison on differences between males' fixation points and female' showed different attention bias in 2 groups. The result showed that males' fixation points were relatively concentrated and mostly on upside. While the distribution of females' fixation points was more widespread, showed a part of the attention scattered around.

Table 1. Psychological factors analysis results of male and female in subway escape situation

	MG, mean (SD)	FG, mean (SD)	t	p
Stress level	2.76(1.446)	3.90(1.071)	3.162	.003*
Systematic analysis of the situation before decision making	2.79(1.380)	3.52(1.209)	2.118	.040*

Notes. M=male group; F=female group.

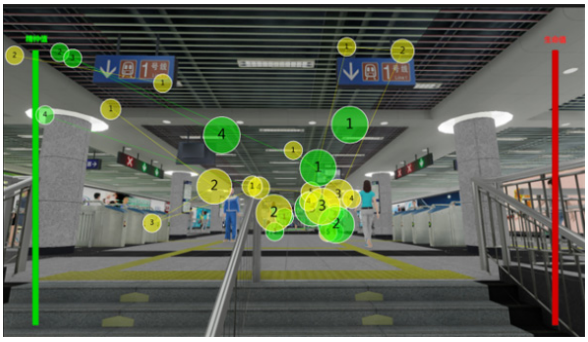


Fig.2. Distribution of fixation points on subjects (male-green; female-yellow).

4. Discussion

In this study, female had higher stress-level in fire emergency than male generally, while other studies found male had higher stress-level than female in daily life. The difference might due to the fact that females are more vulnerable to fire situations, either the fact that males fail to fit themselves into the real situation.

As for indicator searching behavior, females might look for helpful signs from their surroundings according to the fixation point is more scattered. However, males tend to concentrate on the information ahead and upside. Based on these differences, suggestions are proposed for escape instructions in the subway stations:

Fig. 3. Indicator board arrangements based on our experiments.

(1) The evacuation indicator should be placed separately to distinctly guide passengers who were nervous in emergency finding escaping route (Fig.3).

(2) Corresponding emergency escape indicators should be set based on the perceptual habits and characteristics of gender differences.

(3) The exit indicator light belt should be set up on each pillar. In a fire emergency, smoke would trigger alarm, in that case, the escape indicator light belt would help passengers to escape quickly (Fig. 3).

References

1. Ma xuan, Chen long, Pei hua, et al. Relationship between coping style, personality characteristics and stress level, *Chinese journal of clinical psychology*, 2004, 12(1):48-49.

2. Wang liyan, Sui guangyuan, Wang yonghui, et al. Eye movement research of gender differences in visual selective attention. *Psychological science,* 2010, (2):409-411.

3. Fanxing Meng, Wei Zhang, Way-finding during a fire emergency: An experimental study in a virtual environment, *Ergonomics*, 2014, 57(6):816-827.

4. Fujii, Kosuke, Wayfinding evacuation behavior in a large-scale station based on analysis of path choice area - study on the evacuation in railway stations in case of fire part. *Journal of Environmental Engineering*, 2014, 79(704): 829-839.

5. Gamberini, L.P.Cottone, A. Spagnolli, D.Varotto, and Mantovani, Responding to a Fire Emergency in a Virtual Environment: Different Patterns of Action for Different Situations. *Ergonomics*, 2003, 46(8):842-858.

Research on the Smart Home Design based on Single-chip Microcomputer

Hongbo Hao [1], Fengzhi Dai [1, 2]*, Haokang Wen [1], Jichao Zhao [1]

[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

The research of the home is mainly reflected in the indoor environment temperature, humidity, smoke concentration and the human body to monitor. This indoor environment monitoring system takes the STC89C52 monolithic computer as the control core, using MQ-2 smoke sensor, temperature and humidity sensor, infrared sensor, buzzer and other modules, the main function is to detect the indoor environment temperature and humidity, smoke concentration exceeded the predetermined value, immediately for sound and light alarm; and when the body is detected, immediately to the sound and light alarm. This system can be used in shopping malls anti-theft, Warehouse anti-theft, bank anti-theft and other occasions. The device for the current people's daily life has a strong guiding significance.

Key words: 52 single-chip microcomputer, temperature and humidity detection, MQ-2 sensor

1. Introduction

1.1. Research background and significance

In recent years, with the advancement of science and technology, intelligent environmental monitoring has become a topic of common concern to mankind. Smart cities and smart homes bring together control systems such as lighting, fire detection, air conditioning, and security alarms via the Internet to implement intelligent temperature and humidity adjustment, appliance lighting control, fire and burglar alarm control, and environmental detection automatic adjustment functions[1]. Compared with the past and present, the biggest difference between smart homes and traditional homes lies in the intelligence of the home environment. The five most significant advantages of smart home are comfortable, safe and efficient.

This design first describes the development prospects of smart home in human life. Secondly it introduces the concept of the overall framework of this topic and the selection of circuit module components of each module, then it draws the circuit diagram of each module and analyzes the principle[1]. Then it writes the design of software program. Finally it summarizes and predicts the design of this system. This design enables the system to operate efficiently and reliably.

1.2. Function Overview

The design of this project mainly consists of five parts, namely temperature and humidity detection, smoke concentration detection, human infrared sensing, alarm information and display screen information[2].

- The DHT11 temperature and humidity sensor send the detected data to the LCD1602A LCD screen through a single-chip microcomputer. If the detection result is not within the setting range, an audible and visual alarm is issued.

- The function of the smoke sensor can analyze and check the indoor carbon dioxide, methane and smoke concentrations. When the concentration of these gases exceeds a predetermined alarm value, the LED lights up and the buzzer sounds an alarm.
- When the owner leaves the house or rests, he can enter the password on the keyboard of the security system to arm, and the system will enter the arming state. When someone invades the room, the human body infrared sensor can quickly locate the human's range of motion and can send corresponding command signals along with the area where the human body moves.

2. Overall system design

2.1. Design scheme

This design topic is the design of an indoor environment monitoring system based on a single chip microcomputer. The target of the design service is the home of an ordinary family. The reference style chosen for this design is about 120 square meters, with three bedrooms, one living room and one bathroom[3].

2.2. System composition

This system is divided into several modules: the single chip microcomputer (STC89C52) contains a watchdog timer, which supports parallel editing and ISP reline editing[2]. Temperature and humidity sensor (DHT11), this chip has strong anti-interference ability, fast execution of instructions, cost-effective. Liquid crystal display module (LCD1602), MQ-2 smoke sensor, buzzer alarm. The anti-theft module uses a passive infrared sensor. This chip has high cost performance, high security, and strong detection.

3. Hardware circuit design

3.1. Main control circuit

STC89C52 is a low power, high performance CMOS 8-bit microcontroller with 8K system programmable flash memory. The minimum system includes a single-chip microcomputer and its required power, clock, reset, and other components, which can keep the single-chip microcomputer in a normal operating state at all times. Circuits such as power supply and clock are necessary conditions for the single chip microcomputer to operate. The minimum system can be taken as the core part of the application system. By performing memory expansion and A/D expansion, the single chip microcomputer can complete more complex functions.

3.1.1. Reset circuit

The reset operation selected in this design is performed manually. The capacitor in the circuit and the switch used are operated in parallel to meet the predetermined requirements. The resistor R1 plays a role of protecting the circuit. The schematic diagram of the reset circuit is shown in the Fig.1 below.

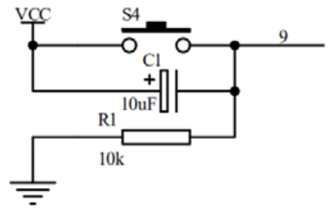


Fig1. Reset circuit schematic

3.1.2. Clock Circuit

This design uses an internal clock circuit, and the selected external crystal is 12MHz. C1 and C2 are load capacitors with a capacitance of 30pF. The function of the two capacitors is to improve the frequency stability and the quickness of oscillation. The selection range of the capacitor is 5-30pF, and the oscillation frequency selection range of the crystal is 1.2-12 MHz.

3.1.3. Interrupt System

Interrupt technology is mainly used for real-time monitoring and control. It is required that the single-chip microcomputer can respond to the service request made by the interrupt request source in time, and make a quick response and timely processing. This is achieved by the on-chip interrupt system. When the interrupt request source issues an interrupt request, if the interrupt request is enabled, the single-chip microcomputer temporarily suspends the main program currently being executed and transfers to the interrupt service handler to process the interrupt service request. After the interrupt service handler has finished processing the interrupt service request .It returns to the original interrupted program (breakpoint) and continues to execute the interrupted main program[4].

3.1.4. Temperature and humidity detection

DHT11 is a temperature and humidity composite sensor with a calibrated digital signal output. It uses special digital module acquisition technology and temperature and humidity sensing technology[3]. The sensor includes a resistive humidity sensing element and an NTC temperature measuring element, and is connected to a high-performance 8-bit microcontroller. Therefore, this product has the advantages of excellent quality, ultra-fast response, strong anti-interference ability, and high cost performance. The actual sensor used this time is shown below. The physical picture of DHT11 sensor, as shown in the Fig.2 below.

Fig.2 DHT11 sensor physical map

3.2. MQ-2 smoke module

The MQ-2 smoke detection module mainly uses the MQ-2 gas sensor to detect hydrogen in natural gas, liquefied petroleum gas, coal gas and other gas components. This chip has the advantages of strong anti-interference ability, high measurement accuracy and fast response. This chip is mainly composed of ceramic tubes, sensitive layers, measuring electrodes, heaters and other sensitive components[4].

3.3. Alarm module

Light-emitting diode is short for LED. It uses a semiconductor device to convert physical signals into electrical signals. In addition, the component uses tungsten wires and phosphors to emit different colors of light as alarm signals. This design scheme is to use LED lights to emit different colors of light as an alarm signal.

3.4. LCD Module Design

LCD1602A uses a digital liquid crystal display, which is easy to operate, small in size, fast in response, high in sensitivity, and clear in picture.

3.5. A/D conversion circuit

In this design, ADC0832 uses a single-channel analog signal input. The connection between ADC0832 and STC89C52 should be 4 data lines: chip select signal CS, clock signal CLK, data output DO, and data output DI. It is an 8-bit resolution operation and uses two channels using A/D conversion elements.

3.6. Button module design

The key design of this project uses a stand-alone keyboard. Because the different output and input pins in this type of key will be stringed together with a certain coded key. The other port is usually grounded. Generally, the single-chip microcomputer operates at a high-potential position in the initial state.

3.7. Design of human detection module

This design uses passive infrared detectors for human detection. Passive infrared sensing technology uses infrared light-sensitive devices to convert trace infrared rays emitted by living organisms into corresponding electrical signals, which are amplified and processed. It can distinguish moving organisms from falling objects[5]. At the same time, it also has the characteristics of large monitoring range, good concealment, strong anti-interference ability and low false alarm rate. This chip is suitable for home use, with low cost and high security.

4. Software design

After the control system powered on, it will reset to the initial state which is the system initialization[6]. This process includes initialization of smoke sensor MQ-2, ADC0809 initialization, temperature and humidity sensor initialization, and so on. If it exceeds the predetermined value, an audible and visual alarm will be issued. If not, it will continue to test the next cycle.

5. System testing and conclusion

5.1. Hardware testing

Before testing the hardware of this design, it makes sure that the welding of the good circuit is correct. The key1 is to set the temperature, humidity and smoke module. Key2 is the plus key, key3 is the minus key, and key4 is the arm disarm key. When key1 is pressed, you can set the upper and lower limits of temperature and humidity and the upper limit of smoke. When the detected humidity falls below the lower limit, the fifth LED turns red and alarms simultaneously. When the indoor smoke concentration is detected to be higher than the set value. The first LED light turns white and an alarm is issued. When the key4 is pressed, the sixth LED turns purple and the LCD screen displays the letter “Z”. After about 20 to 30 seconds, the system enters the armed state and the LCD screen displays the letter "B" The arming status LED is off.

5.2. Software testing

First the switch is turned in the lower left corner of the simulation software and the buzzer will keep alarming. This is because the AT89C52 single-chip computer in the simulation. Software does not have the function of saving data when power is off. When the indoor smoke concentration is lower than this value, the sound and light alarm will stop. For the human infrared sensor module, the LCD screen displays "Z", and the LCD screen displays "B" for about 20 to 30 seconds. The indoor smoke concentration value is shown in the Fig.3 below.

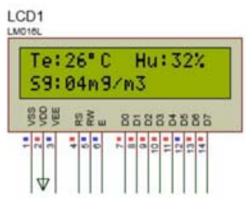


Fig.3 Indoor smoke concentration value

Acknowledgements

The research is partly supported by the Project of Tianjin Enterprise Science and Technology Commissioner to Tianjin Tianke Intelligent and Manufacture Technology Co., Ltd (19JCTPJC53700). It is also supported by the Industry-University Cooperation and Education Project (201802286009) from Ministry of Education, China.

References

1. Terry J., 10 Emerging Technologies that Will Change the World, *Technology Review*, 2003, 106(1): pp.33~49.
2. ISO/IEC14443, Final Committee Draft, 2003, pp.20-35.
3. Texas Instruments, Z-Stack Application Programming Interface, 2013, pp.66-70.
4. Bing Xia, *Design and implementation of data acquisition and identification algorithms in an online logic chip analysis tool*, Henan, PLA Information Engineering University, 2006, pp.22-33.
5. Wei Qin, Based on the AM2301 temperature and humidity meter design, *Hunan Agricultural Machinery*, 2012, 9: pp.59-60.
6. Huailin Zhao, Shunzhou Wang, Yunxiang Liu, Xuyao Hao, Design of an Intelligent Housekeeping Robot Based on IOT, *Journal of Information and Communication Engineering*, 2016, (2): pp.114-118.

Design of a portable instrument for measuring heart rate and blood oxygen

Haokang Wen [1], Fengzhi Dai [1,2,3]*, Jichao Zhao [1], Hongbo Hao [1], Qianqian Zhang [1]

[1] *Tianjin University of Science and Technology, China,*

[2] *Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China,*

[3] *Advanced Structural Integrity International Joint Research Centre, Tianjin University of Science and Technology, Tianjin, 300222, China*

*E-mail: * daifz@tust.edu.cn*

www.tust.edu.cn

Abstract

Aiming at the disadvantage of the inconvenience of the current heart rate blood oxygen detection device, a portable heart rate blood oxygen detector is proposed in this paper. The device can detect the heart rate and blood oxygen value of the human body at any time and place according to the needs of the measured person, and can display the heart rate, blood oxygen and temperature on the liquid crystal display. The instrument gets rid of the cumbersomeness of traditional medical appliances, greatly reducing the volume of the instrument and making it easier to carry.

Keywords: heart rate, blood oxygen, Arduino UNO

1. Introduction

With the advancement of human science and technology, the quality of people's lives is increasing, and the acceleration of the pace of life makes contemporary people pay more and more attention to health issues. Detecting vital signs such as heart rate, arterial oxygen saturation, and temperature is of great significance in the fields of clinical medicine, community medicine, and home health monitoring. However, there are still many problems during the consultation process, such as the small number of hospitals, the long distance, the complicated procedure of the consultation process, the pain caused by invasive medical treatment, the high cost, and the accompanying side effects. All bring a great burden. In response to this situation, a detection system is designed in this article, so that the test does not need to perform a complicated detection process, there is no negative effect in the detection process, and the overall equipment price is not so expensive. Through this system, health signs parameters can be obtained.

2. SpO_2 and pulse wave introduction

2.1. SpO_2 introduction

The human body inhales oxygen from the outside world to the alveoli and expels carbon dioxide through the pulmonary capillaries, while oxygen is sent into the pulmonary capillaries. When oxygen enters the body, it will flow to all parts of the body through the respiratory circulatory system to supply oxygen to the entire body organs. When O_2 is converted to CO_2, O_2 and CO_2 will undergo a second exchange in the capillaries, so that the cells of the whole organ can get sufficient oxygen. The maximum combined amount of hemoglobin and oxygen is usually defined as the oxygen capacity of the blood, which is represented by HbO_2, where Hb is hemoglobin. Blood oxygen saturation is often used to determine blood oxygen content in the human body. Blood oxygen saturation is the percentage of oxygenated hemoglobin in the blood [1], which is

$$S_aO_2 = \frac{HbO_2}{Hb+HbO_2} \qquad (1)$$

Blood oxygen saturation plays an important role in the human respiratory cycle, and can estimate the oxygenation of the lungs and the oxygen-carrying capacity of hemoglobin.

2.2. *Pulse Wave Introduction*

The fluctuation of human arteries is referred to as pulse, The fluctuation of the human arteries is referred to as pulse, which means that when the human left ventricle dilates, the blood is squeezed into the aorta, the blood vessels will block the blood, and the blood will flow into the vein for the first time. At this time, as the blood penetrates into the aorta, the aortic compression will increase. This forces the diameter of the aorta to expand, and the human body can feel this pulsation expansion phenomenon through the ring finger, wrist and other parts. At the same time, the pulse is expanded and contracted by the systolic function of the heart and the elasticity of the wall of the heart. When the ventricle is opened, the pressure of the arteries will increase; when closed, the pressure of the arteries will decrease. This kind of arterial fluctuation will generate a pulse wave.

There are two types of pulse wave acquisition methods: pressure and volume. This article mainly uses the photoelectric volume pulse wave. The volume pulse wave is mainly a waveform obtained by converting the blood vessel volume into a change in light intensity. When blood flows through the microvessels by means of blood circulation, the blood volume exhibits pulsatility under cardiac pulsation. And due to the expansion and contraction of the heart, the volume of blood volume changes from large to small. When the light source enters the human tissue, it will be absorbed by the tissue to produce attenuation, and it will be absorbed by the photodetector, and the vascular rhythm will promote the expansion and contraction of the blood vessel, resulting in the periodic change of the internal blood volume, which will cause the light absorbed by the blood to change. Finally, the pulse wave can be obtained by changing the light intensity.

Electrocardiographic signals (ECG), volume pulse wave signals (PPG), respiratory signals (RESP), and arterial blood pressure signals (ABP) are all important physiological signals for studying human health. Compared with other signals, PPG signals can simply and non-invasively measure the physiological information of the human body, and can continuously measure. Detection with PPG signals has great potential [2].

2.3. *Overall system design*

This system is mainly composed of Arduino microcontroller control module, MAX30100 blood oxygen concentration and heart rate acquisition module, temperature acquisition module, data LCD display module, buzzer alarm module and key input module. The overall system block diagram is shown in Fig. 1.

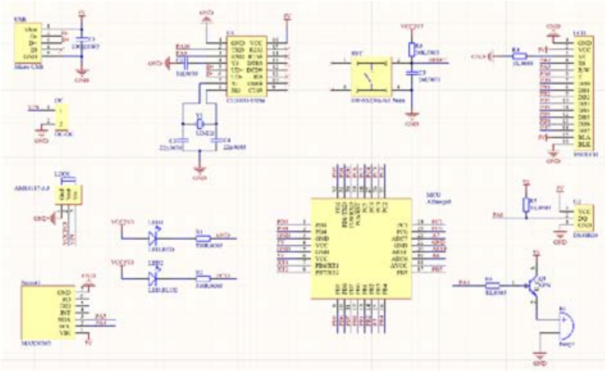


Fig.1 Overall system block diagram

2.4. *MAX30100 module*

The MAX30100 is a module that integrates a pulse oximeter and a heart rate monitoring sensor. The device integrates two LEDs, a photodetector, optimized optics, and a low-noise analog signal processor to detect pulse oximetry and heart rate signals. It can be widely used in fitness auxiliary equipment, medical monitoring equipment and wearable equipment, as shown in Fig.2.

Fig.2 MAX30100

Pin description:

- VIN: LED power input terminal, which is also the I2C bus pull-up level. It is recommended to connect 3.3V or 5V.
- SCL: clock connected to the I2C bus
- SDA: connect to I2C bus data

- INT: Interrupt pin of MAX30100 chip
- IRD: IR LED ground terminal of MAX30100 chip
- RD: RED LED ground terminal of MAX30100 chip
- GND: ground wire

The integrated chip MAX30100 provides a standard I2C compatible communication interface, and external devices communicate through I2C. The working principle of the I2C bus: When SCL is high, SDA transitions generate a start signal from high to low level. First, the highest bit is sent out and the host sends a start signal. When SCL is high, SDA transitions from low to high level to generate a stop signal. The host ends the data communication by sending a stop signal. Because MAX30100's I2C compatible communication interface needs to set SCL and SDA high, a pull-up resistor is added to the SDA and SCL pins in the acquisition circuit.

The working principle of this chip for measuring blood oxygen saturation and pulse is that the internal LED driver sets the timing in advance, alternately irradiates red light and infrared light, and then converts the reflected light signal into an analog electrical signal through a photodetector. The analog electric signal is output, and the signal is converted into digital signal after being amplified and filtered. Finally, the data is placed in the FIFO buffer[3]. The MAX30100 sensor can be transferred to a microcontroller via an I2C digital signal to obtain a pulse wave signal. The working principle of measuring blood oxygen saturation and pulse is shown in the Fig. 3.

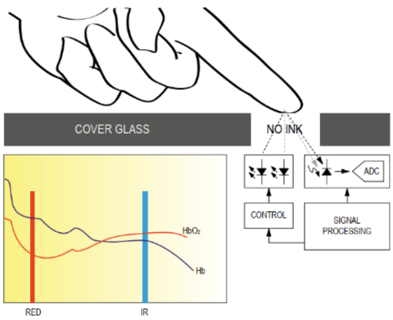


Fig.3 Working principle of measuring blood oxygen saturation and pulse

2.5. MCU main control chip

The Arduino UNO control board is based on the ATmega328 single-chip microcomputer. By improving the control of the Arduino UNO control board, a more convenient and stable development board is obtained. The Arduino uno microcontroller board is flexible and inexpensive, providing a variety of digital and analog inputs, serial interfaces, and digital and PWM outputs[4]. The Arduino UNO control board consists of fourteen digital input and output pins, six analog input ports, USB access ports, power interface, data ICSP interface, reset button, and a 16MHZ quartz crystal oscillator. The physical picture of the Arduino UNO control board is shown in Fig. 4.

Fig. 4 Arduino UNO control board

2.6. DS18B20

The DS18B20 digital thermometer provides 9-bit to 12-bit Celsius temperature measurements and has an alarm function with nonvolatile user-programmable upper and lower trigger points. The DS18B20 communicates over a 1-Wire bus that by definition requires only one data line (and ground) for communication with a central microprocessor. It has an operating temperature range of -55°C to +125°C and is accurate to ±0.5°C over the range of -10°C to +85°C. In addition, the DS18B20 can derive power directly from the data line ("parasite power"), eliminating the need for an external power supply.

1-Wire bus technology has the advantages of saving I/O resources, simple structure, low cost, convenient bus expansion and maintenance [5], so it is widely used in temperature monitoring systems and process monitoring systems inside buildings, equipment or machinery.

The DS18B20 chip relies on a single wire port communication through the Dallas single-bus protocol. When all devices are connected to the bus via a tri-state port or an open-drain port, the control line needs to be connected with a weak pull-up resistor. The external power supply mode is the best working mode of the DS18B20 temperature sensor. It works stably and reliably, has strong anti-interference ability, and the circuit is relatively simple. Therefore, this power supply is used in this design. Fig.5 shows the DS18B20 chip circuit connection method.

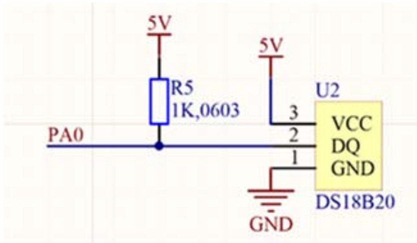


Fig.5 DS18B20 circuit connection method

2.7. 1602 LCD display

LCD1602 character liquid crystal display module is a dot matrix liquid crystal display module specially used for displaying letters, digits, symbols and so on. It is divided into 4-bit and 8-bit data transmission methods. It provides 5×7 dot matrix + cursor display mode, and provides display data buffer DDRAM, character generator CGROM and character generator CGRAM. CGRAM is used to store the font data of up to eight 5×8 dot matrix graphic characters defined by ourselves. LCD1602 has a total of 16 pins, but there are three main pins used for programming, which are: RS (data command selection end), R/W (read and write selection end), E (enable signal). The programming in this design is mainly around these three pins to initialize, write commands, and write data. Fig.6 shows the LCD1602 circuit connection method.

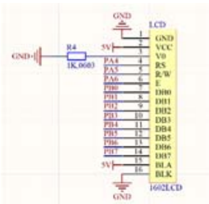


Fig.6 LCD1602

3. Conclusion

This article mainly designs a non-invasive pulse oximetry monitoring system, which is mainly applicable to the family, and is convenient for monitoring the physical condition of the user in the family [6]. In this system, the Arduino uno is used as the main control chip. The heart rate and blood oxygen value are detected by the sensor MAX30100, the temperature is measured by the DS18B20, and the collected data is transmitted to the chip. The heart rate, blood oxygen value and temperature of the human body are displayed on the LCD screen in real time. When the measured data is abnormal, an alarm sound will appear to remind the patient.

Acknowledgements

The research is partly supported by the Project of Tianjin Enterprise Science and Technology Commissioner to Tianjin Tianke Intelligent and Manufacture Technology Co., Ltd (19JCTPJC53700). It is also supported by the Industry-University Cooperation and Education Project (201802286009) from Ministry of Education, China.

References

1. Carmelneiderman N N,Goren I, Wasserstrum Y, et al., Structured, Protocol-Based Pulse-Oximetry Measurement Improves the Evaluation of Hypoxemic Patients at Hospital Admission, *Israel Medical Association Journal*, 2018, 20(3): pp.147-150.
2. Xiaoqin Luo, Fen Miao, Ye Li, PPG and ECG feature comparison for healthy people and hypertensive patients, Proceedings of the 2012 IEEE-EMBS International Conference on Biomedical and Health Informatics, 2012, pp. 785-787.
3. Strogonovs, R., Implementing pulse oximeter using MAX30100, *Morf-Coding and Engineering*, 2017.
4. Al-thobaiti, B. M., Abosolaiman, I. I., Alzahrani, M. H., Almalki, S. H., & Soliman, M. S., *Design and implementation of a reliable wireless real-time home automation system based on Arduino Uno single-board microcontroller*. *International Journal of Control, Automation and Systems*, 2014, 3(3): pp.11-15.
5. LI, G., ZHAO, Y., Principle and Application of 1Wire Bus Digital Thermometer DS18B20. *Modern Electronic Technique*, 2005, 21.
6. Huailin Zhao, Shunzhou Wang, Yunxiang Liu, Xuyao Hao, Design of an Intelligent Housekeeping Robot Based on IOT. *Journal of Information and Communication Engineering*. 2016, 2(2): pp.114-118.

Principal Component Analysis of Wine Based on Three-dim Fluorescent Spectra Characteristic

Di Yin [1,3], Fengzhi Dai [1,2,3]*, Yuxuan Zhu [1], Yasheng Yuan [1]

[1] *Tianjin University of Science and Technology, China;*

[2] *Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China;*
[3] *Advanced Structural Integrity International Joint Research Centre,*
Tianjin University of Science and Technology, Tianjin, 300222, China
*China E-mail: * daifz@tust.edu.cn*
www.tust.edu.cn

Abstract

The traditional sensory analysis method cannot discriminate different kinds of wine objectively and accurately. The three-dimensional fluorescence technology has the advantages of strong selectivity, high resolution and direct access to the fluorescence characteristics of the sample. In this paper, the 3D fluorescence spectroscopy five brands of red wine samples. The five characteristic parameters (the mean value, standard deviation, center of gravity coordinates, first order center moment and correlation coefficient) of the 3D fluorescence spectra of these five dry red wine samples are extracted respectively. The results showed that the three different dimensions of Cabernet Sauvignon the characteristic parameters of the fluorescence spectrum are similar, based on the analysis of five dry red wine samples by principal component analysis, it was found that even the same raw material brewing wine due to different origin of raw materials, soil environment, climate and environment will cause the composition of the wine content, which shows a large difference.

Keywords: Characteristic parameters; Principal component analysis; Wine; Three-dimensional fluorescence spectroscopy

1. Introduction

Studies have shown that wine contains more than 200 nutrients that are beneficial to the human body [1], organic acids, amino acids, vitamins, polyphenols, and inorganic salts, etc. These elements maintain normal development of human function. In China, compared with the industrialized countries in the world, wine started late, has a weak foundation, and has a low level. Although some achievements have been made, there are still deep-rooted problems of healthy and stable development, which have certain effects and limit wine production. The various safety issues discovered have caused serious harm to people's living standards and mental health, and have doubts about food processing safety issues in society. In order to fundamentally improve the safety status of Chinese wine, it is necessary to detect the trace elements of wine in the production process to identify the origin, year and quality of wine [2].

Different brands of wine have different trace components. Existing detection methods in the laboratory can meet the needs of wine detection on a certain basis, but there have some disadvantages of complex pre-treatment processes, time-consuming measurement, high cost, sensitivity, and insufficient accuracy [3].

This article uses different brands of dry red wines as research objects, applies three-dimensional (3D) fluorescence spectroscopy to study 3D fluorescence spectra, and identifies and identifies wines based on the principal component interpretation (PCA) [4].

2. Experiment

2.1. Experimental device

This experiment uses the LS-55 spectrofluorimeter from the PerkinElmer company of America, and the light source is pulsed xenon lamp. The cuvette is a 1cm quartz liquid cell with stopper. This consists of 5 units, namely, light source, monochromator, sample cell, detector and display [5].

2.2. Experimental method

This paper uses 250~350nm ultraviolet light to excite five wines of three brands with an interval of 5nm and measure the emission spectrum of 200~700nm to get fluorescence spectrum data .In order to avoid fluorescence quenching during wine testing, the sample should be diluted 20 times and put into a cuvette for measurement.

2.3. Data processing and analysis

The obtained fluorescence data was normalized using MatlabR2017a software, filtered to obtain a 3D fluorescence spectrum, and characteristic parameters (the number of fluorescence peaks, the peak wavelength of the main fluorescence peak, and the best excitation wavelength) were extracted from the surface fitting.

2.4. Three-dimensional fluorescence spectra of wines of different brands

Wine is rich in phenols, amino acids, anthocyanins, tannins, epicatechins and other flavonoids. These substances contain fluorescent groups that emit fluorescence under the irradiation of light with a certain wavelength. The content of these fluorescent molecules varies according to the variety and maturity of the grape, and it is also affected by the way the wine is made and how long it is stored.

Fig1 to figure 5 are 3D fluorescence spectra of five different wine samples, showing the total fluorescence spectrum information in the samples. Table2 lists the characteristic values of dry red 3D fluorescence spectra of five different wine samples.

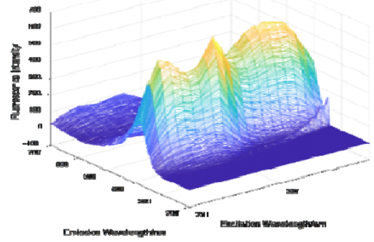


Fig.1 Great Wall Dry Red Wine

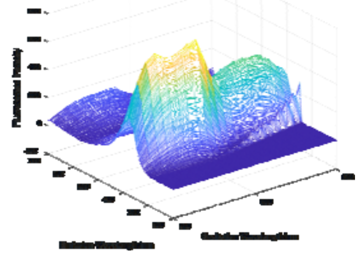


Fig.2 Great Wall Cabernet Dry Red Wine

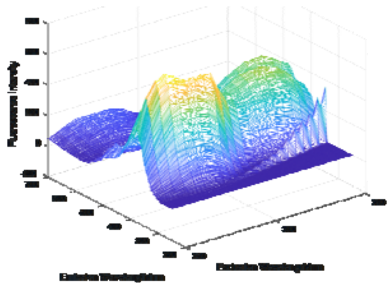


Fig.3 Changyu Dry Red Wine

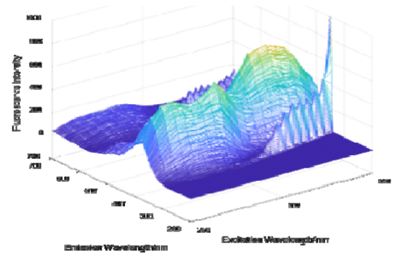


Fig.4 Changyu Dry Red Wine (premium wine)

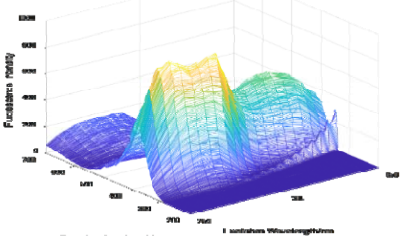


Fig.5 Arabella Dry Red Wine

The 3D fluorescence spectra of the five samples are similar, all of them have three fluorescence peaks, that is located in (290nm / 375nm), (265nm / 375nm), (325nm / 420nm), but the intensities of the three peaks are different. This is because the raw materials of the five samples are the same as Cabernet Sauvignon brewing, the material composition is the same, so the location and number of wave peaks are the same, but because of the different places of origin, the content of each component is different, resulting in the difference of fluorescence peak intensity.

Table1 Characteristic parameters of 3D fluorescence spectra of samples

sample	wave peak number	peak position	fluorescence intensity	optimum excitation wavelength (/nm)
1	3	260/376.5	619.6	290
		290/375	680.3	
		325/418.5	637.2	
2	3	265/372.5	941.3	265
		290/374.5	908.5	
		330/429	567.6	
3	3	265/372.5	710.5	265
		290/373	625.7	
		335/423	483	
4	3	290/374	651.2	325
		325/419.5	783.4	
		265/375	587.2	
5	3	265/376.5	757.3	290
		290/374.5	772.8	
		335/418.5	455.1	

From the visual analysis of the three-dimensional fluorescence spectrum, there are certain differences between different wine samples, which cannot be distinguished directly. Only by analyzing whether the difference is obvious through mathematical statistics can it be more objective and clearer. Therefore, the characteristic parameters of the map need to be extracted. On this basis, the principal component analysis was performed to distinguish the effects of different wine samples.

3. Principal Component Analysis

3.1. The concept of PCA and feature data extraction

PCA is one of the most widely used data reduction algorithms. The main idea of PCA is to map n-dimensional features onto k-dimension. This k-dimension is a brand-new orthogonal feature, which is also called principal component. It is a k-dimension feature reconstructed from the original n-dimension feature. The software used in this experiment is SPSS 19. Analyze the trace components of different brands of wine, extract the characteristic parameters of the original data of three-dimensional fluorescence spectrum of five kinds of wine samples [6]. Table 2 shows the extraction of characteristic parameters of 3D fluorescence spectra.

Table2 Extraction of characteristic parameters of 3D fluorescence spectra

number	mean value	standard deviation	abscissa of center of gravity	vertical coordinates of center of gravity	first order center moment	correlation coefficient
1	118.2692	168.566	302.5984	430.1101	-1.33e+05	-3.12e+05
2	149.9626	204.2795	295.6212	430.4683	-1.27e+05	-3.59e+05
3	156.788	171.6284	300.0153	441.7986	-1.33e+05	-3.52e+05
4	140.4287	177.1754	303.6513	437.7785	-1.33e+05	-3.64e+05
5	118.0955	165.4325	296.6428	428.3699	-1.27e+05	-3.50e+05

3.2. Clustering results

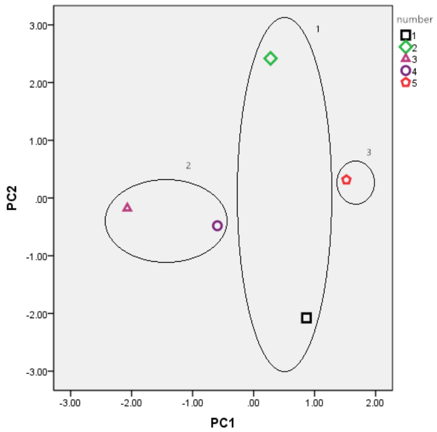
The purpose of cluster analysis is to cluster different brands of wines using extraction parameters. And accurate classification of different brands of wines. Five kinds of wine data were used in this paper. The distance between samples was Euclidean distance and hierarchical clustering method was used for clustering.

Table3 Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	%of Variance	Cumulative%	Total	%of Variance	Cumulative%
1	2.651	44.178	44.178	2.651	44.178	44.178
2	1.943	32.381	76.558	1.943	32.381	76.558
3	0.875	14.575	91.134			
4	0.532	8.866	100.000			
5	2.799E-16	4.665E-15	100.000			
6	8.092E-17	1.349E-15	100.000			

It can be seen from Table 3 that the variance contribution rate of principal component parameters principal component 1 (PC1) and principal component 2 (PC2) is 44.178% and 32.381% respectively, the cumulative variance contribution rate is 76.558%, and the two characteristic values are 2.651 and 1.943 respectively.

Fig6 is the three-dimensional fluorescence spectrum PC1-PC2 principal component score chart of dry red wine from different origins of different brands [7]. Although the wines from different origins of different brands use the same raw material Cabernet Sauvignon, there are significant differences among the five samples, which shows that even the wines from the same raw material have different origins, soil environment and climate environment will cause the formation of wine The results show that there are significant differences.



1: Great Wall Brand 2: Changyu Brand
3: Arabella Brand
Fig.6 PC1-PC2 principal component scores of 3D fluorescence spectra of different brands of wine

4. Conclusion

Based on the characteristics of fluorescence spectrum, this paper studies the 3D fluorescence spectrum of five dry red wine samples from three brands. It is found that the 3D fluorescence spectrum characteristics of the same Cabernet Sauvignon wine are basically the same, which can't be directly distinguished by spectral comparison. Then preprocess the spectral data and extract five characteristic parameters as the characteristic index of wine recognition. On this basis of principal component analysis, the cumulative variance contribution rate is 76.558%, which can be divided into three categories, which can achieve good results for wine brand differentiation. Taken together, even if the wine made from the same raw material is different from the origin of the raw material, the soil environment and climatic environment will affect the content of the wine, and it will show a large difference.

Acknowledgements

The research is partly supported by the Project of Tianjin Enterprise Science and Technology Commissioner to Tianjin Tianke Intelligent and Manufacture Technology Co., Ltd (19JCTPJC53700). It is also supported by the Industry-University Cooperation and Education Project (201802286009) from Ministry of Education, China.

References

1. Ursini, Fulvio, and Alex Sevanian, Wine polyphenols and optimal nutrition, *Annals of the New York Academy of Sciences*, 2002, pp. 200-209.
2. Liu, Fang, and Jamie Murphy, A qualitative study of Chinese wine consumption and purchasing: Implications for Australian wines, *International Journal of Wine Business Research*, 2007, pp. 98-113.
3. Ward, Alastair James, et al., Real time monitoring of a biogas digester with gas chromatography, near-infrared spectroscopy, and membrane-inlet mass spectrometry, *Bioresource technology*, 2011, pp. 4098-4103.
4. Destefanis, Gianluigi, et al., The use of principal component analysis (PCA) to characterize beef, *Meat science*, 2000, pp. 255-259.
5. Yin Di, Dai Fengzhi, et al, Cluster Analysis of Wine based on Three-dim Fluorescent Spectra Characteristic, *Proceedings of the 2019 International Conference on Artificial Life and Robotics*, Oita, Japan, Jan. 2019, pp. 280-283.
6. Kallithraka, S., et al., Instrumental and sensory analysis of Greek wines; implementation of principal component analysis (PCA) for classification according to geographical origin, *Food Chemistry*, 2001, pp. 501-514.
7. Ringnér, Markus, What is principal component analysis? *Nature biotechnology*, 2008, pp. 303.

Control Design of Intelligent Device for Living Environment of Senile Apartment

Yasheng Yuan[1,4], Fengzhi Dai[1,2,4*], Shengbiao Chang[3], Lingran An[1], Di Yin[1], Yuxuan Zhu[1], Yujie Yan[1]

[1] *Tianjin University of Science and Technology, China,*
[2] *Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China,*
[3] *Dawning Information Industry Co., LTD, China,*
[4] *Advanced Structural Integrity International Joint Research Centre, Tianjin University of Science and Technology, Tianjin, 300222, China*
*E-mail:*daifz@tust.edu.cn*
www.tust.edu.cn

Abstract

At present, with the advancement of science and technology, some smart devices in the society are gradually integrated into the daily lives of the elderly. Now wireless network WIFI technology, real-time network communication function, some new Internet of things technology in the society are more and more accepted by the elderly. This paper researches the existing elderly apartments in the society, and uses the WIFI wireless communication method for information transmission in response to the special physiological needs of the elderly. The bottom layer uses various sensor nodes to transmit indoor environmental data to the upper layer in real time, and can reversely control each indoor electrical equipment through the current environmental index.

Keywords: WIFI, senile apartment, the Internet of things, single chip

1. Introduction

Now, the senile apartments in our country have been remodeled from the original nursing homes into small apartments, resulting in the domestic senior citizen apartments being relatively functionally single, thus it is not enough to provide humanized design for the elderly. It can only provide a single life service, and it cannot provide a comfortable place for the elderly [1]. However, with the development of information technology, the quality of the elderly has also improved. In their later years, the requirements of life for the elderly have become higher and higher. People pursuit the stable and comfortable living environment has become the general trend of the development for senior citizen apartments in the future [2].

In conclusion, a set of intelligent research solutions for senior citizen apartments has been designed, which aims to freely monitor indoor environmental parameters and adjust them according to the parameters to satisfy the senior citizens' living needs. This solution uses STM32 for control, uses WIFI wireless communication for information transmission [3], monitors environmental parameters for upper-layer transmission by connecting various sensors, and control electrical appliances according to the current environment.

2. The hardware structure design

This article selects STM32F103ZET6 as the main control core to the environmental control system for the elderly apartment, which have environmental quality sensor detection module and control system in the system. Each part of the system design is modularized as much as possible to reduce the coupling between the module and the system, so that when a certain part of the hardware module of the environment control system in the elderly apartment fails, it will not influence the entire system, and the system can still be operationed. During operation and maintenance, it is only need to replace some of the modules or modify some codes. The simplification of maintenance can reduce the overall system maintenance cost and improve the stability and reliability of the environment control system in the elderly apartment.

2.1. Main control system design

The curtain control system realizes the intelligent operation of the curtain by the MCU, which through the use of the photosensitive sensor which is very sensitive to the sun and the stepping motor of the control system. The window control system will open the windows to ventilation which based on the smoke detected by the sensors. When the temperature and humidity sensor detects that the indoor temperature and humidity value is too high or too low, it will turn on the humidifier or air-conditioning in the elderly apartment to adjust the indoor temperature and humidity. At present, the detection of PM2.5 in the air is more and more important for people health, especially for the elderly. When the PM2.5 content in the air exceeds the standard, the air purifier will be turned on to purify the indoor air. The design is shown in Figure 1.

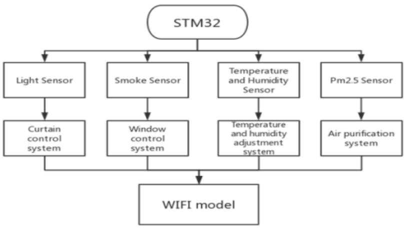


Fig.1.Control system design

2.2. WIFI module

WIFI (Wireless Fidelity) is a certification standard for wireless local area network products which composed of Access Point and wireless network cards.

ESP8266 is a WIFI network solution in the intelligent control system for the living environment of the elderly apartment. It uses 3.3V DC power supply in the intelligent control system for the living environment of the elderly apartment. In the use of the intelligent control system for the living environment of the elderly apartment, it can be set as a client (client) to connect to the sever in the local area network. It can also be used as a slave (router) to run on other hosts (servers). The control system LAN acts as a client and a sever of other clients.

2.3. Power conversion circuit module

This system module uses DC power supply. As shown in Figure 2, the DC-DC chip RT7272 is used in the power module design part of the intelligent control system for the living environment of the elderly apartment. The RT7272 chip is a high-efficiency BUCK IC with an efficiency of up to 98%. It supports an external DC 6 to 24V input voltage in the hardware control system of the elderly apartment living environment. After RT7272, the voltage can be regulated to 5V, then 5V voltage is supplied to the peripheral circuits of the system circuit board [4].

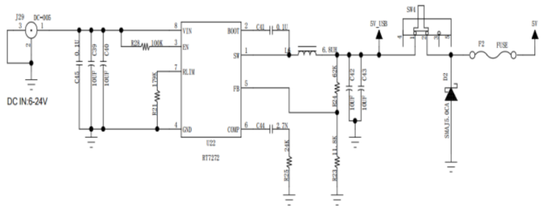


Fig.2. Power supply scheme

2.4. Temperature and humidity module

The temperature and humidity module of this system uses DHT11 digital temperature and humidity sensor. Through the use of specialized digital module and temperature and humidity sensing technology, the temperature and humidity detection of the environment has reliability and excellent long-term stability. The DHT11 module has low cost and strong anti-interference ability. It can quickly measures the relative temperature and humidity of the indoor environment for senior citizen apartments. The circuit schematic is shown in Figure 3. It has humidity measurement elements to measure humidity and NTC temperature measurement elements to measure temperature, and a small, high-performance 8-bit MCU is connected inside the module [5].

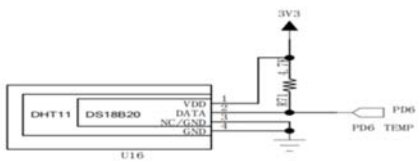


Fig.3. DHT11 circuit design

2.5. Smoke sensor module

The smoke sensor module of this system uses MQ-2 smoke gas sensor. MQ-2 gas sensor has a wide detection radius. It can still maintain high sensitivity and fast response recovery ability in various complex environments [6]. In the harsh environmental conditions can still maintain excellent stable performance. The sensor signal is output through the conditioning circuit to the current indoor smoke concentration value of the elderly apartment and sent to the MCU for processing. Figure 4 is the MQ-2 circuit design.

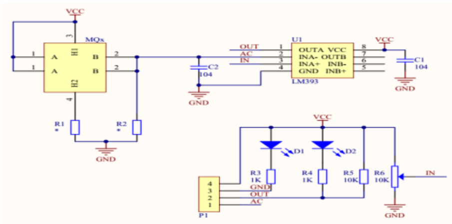


Fig.4. MQ-2 circuit design

2.6. Light sensor module

The ambient light sensor in this system uses a sensitive photoresistor sensor. The working voltage of the living environment monitoring hardware for the elderly apartment is between 3.3V and 5V. The form of output has switching output (0 and 1) and digital analog output (Voltage value). AO analog output is used to collect data in the environmental monitoring system of the elderly apartment. The photosensor circuit design is shown in Figure 5.

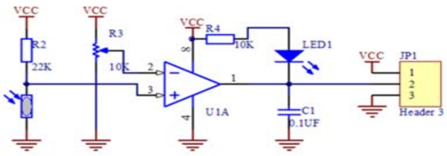


Fig.5. Photosensor circuit design

2.7. PM2.5 detector module

The GP2Y1050 sensor is used to detect PM2.5 in this system. The sensor is not sensitive to some small particles in the elderly apartment, it can only detect low-level air pollution in the elderly apartment. The effect is particularly effective for very small particles [7,8]. For the relatively dense gas such as cigarettes, the level output of GP2Y1050 is a series of high voltages continuously output.

3. Design of System Intelligent Control Module

3.1. Temperature and humidity detection system

After the temperature and humidity information is detected by the DHT11, the system will transmit the temperature and humidity information to the main chip STM32 which can judgment and processing this information. When the temperature and humidity information is not good for the health of the elderly, multiple relay drive circuits with optocoupler isolation are used to control the electrical equipment.

3.2. Intelligent curtain control system

The STM32 compares the indoor lighting information and outdoor lighting information by the photoresistors. When the outdoor light and the indoor light form a certain difference, the STM32 will send a start command to the motor to control the curtain opening and closing. The system have two control modes, the first is the automatic control mode, that is, the MCU automatically controls the curtains based on the measured lighting information, and the second is the manual control, that is, the user can directly send curtain opening and closing commands to the MCU, when user turn on the manual control mode, the automatic mode will turn off.

3.3. Intelligent window control system

When the indoor MQ-2 detects the presence of harmful gases in the room, the STM32 will start the window control motor to open the window for ventilation. When the outdoor MQ-2 detects weather that is harmful to the elderly, such as sandstorms, the STM32 will close the windows to avoid harmful weather affecting the room in time. This system is similar to the intelligent curtain control system, and it also has two control methods, manual and automatic.

3.4. Air purification system

This system collects the indoor air quality in senior citizen apartments by PM2.5 sensors, and acquires the collected information through AD conversion and sends them to the residential monitoring system MCU of the senior citizen apartments. After that, the MCU controls the operation of the air purifier by internal logic.

4. Testing and conclusion

By testing the intelligent control device for the living environment in senior citizen apartments, a preliminary evaluation of the entire system can be performed. By using the temperature and humidity module, MQ-2 module, ambient light sensor module and PM2.5 module, the environmental parameters such as temperature, humidity, light and various harmful gas concentrations can be transmitted to the host in real time through the WIFI module. As shown in Figure 6.

Fig.6.test data

During the test, the hardware side transmits data to the server every 4s. The data contains sensor information such as temperature and humidity, PM2.5, smoke, and light intensity. At the same time, the paper also tests the packet loss rate and communication distance of the WIFI module. In the case of having obstacles, 11,729 data transmission tests were tested within a range of 20m without any data loss, and 10249 data transmission tests were performed within a range of 30m. It was found that amount of data was lost within 40m. After 15015 data transmission tests, it was found that some data was lost during the transmission.

Without obstacles, 10183 data transmission tests were performed within a range of 30m for the first time, and 10028 data transmission tests were performed for a range of 48m within the second time. There were no data during the two data transmission lost, and 10611 data transmissions were performed at a distance of 90m. As a result, a small amount of data was lost at this time. The maximum hop count of the mesh network set during the test is 4 hops. The results are shown in Tables 1 and 2.

Tab.1. Communication distance – setting obstacle

	20 meters	30 meters	40 meters
Times	11729	10249	15015
Success rate	100%	99.84%	93.05%

Tab.2. Communication distance - without obstacle

	30 meters	48 meters	90 meters
Times	10183	10028	10611
Success rate	100%	100%	99.29%

Acknowledgements

The research is partly supported by the Project of Tianjin Enterprise Science and Technology Commissioner to Tianjin Tianke Intelligent and Manufacture Technology Co., Ltd (19JCTPJC53700). It is also supported by the Industry-University Cooperation and Education Project (201802286009) from Ministry of Education, China.

References

1. Yunliang Lu. *Research on the problems and countermeasures of smart pension*, Jiangsu Normal University, Jiangsu, China, 2018.
2. Lili Liu, Eleni Stroulia, Ioanis Nikolaidis, Antonio Miguel-Cruz, Adriana Rios Rincon, Smart homes and home health monitoring technologies for older adults, *International Journal of Medical Informatics*, 2016, pp.91.
3. Biljana L.Risteska Stojkoska, Kire V.Trivodaliev. A review of Internet of Things for smart home. *Journal of Cleaner Production*, 2017, pp.140.
4. Haotian Shi, Zhe Wang. Design of an intelligent temperature and humidity control system indoor. *Communication World*, 2018 (09): pp.292.
5. Baoquan Zhang, Bo Zhang, Ruiya Zhang, Liujun Ma, Dengfeng Zheng, Tingmao Cao, Qiong Liu, Feiyue Wu. Application of temperature and humidity real-time monitoring system based on Internet of Things in tobacco curing. *Tianjin Agricultural Sciences*, 2018 (10): pp.61-63+72.
6. Qingliang Song, Yi Ruan. Analysis of the status and progress of independent smoke alarms. *Application of Integrated Circuits*, 2018, 35 (09): pp.1-6.
7. Lei Lü, Rui Xing, Jincheng Li, Yitao Zhang. Multifunctional intelligent home control system. *Electrical Engineering Technology*, 2018 (18): pp.1-2+11.
8. Huailin Zhao, Shunzhou Wang, Yunxiang Liu, Xuyao Hao, Design of an Intelligent Housekeeping Robot Based on IOT. *Journal of Information and Communication Engineering*, 2016, 2(2), pp. 114-118.

Multiple model adaptive control based on switching/weighting intelligent fusion algorithm

Weicun Zhang*, **Yongnan Jia, Qing Li**
School of Automation and Electrical Engineering, University of Science and Technology Beijing,
30 Xueyuan Road Haidian District, Beijing, China
E-mail: weicunzhang@ustb.edu.cn
www.ustb.edu.cn

Abstract

In this paper, a multiple model adaptive control scheme based on switch/weighting intelligent fusion algorithm is presented with the following contributions: 1) switching/weighting intelligent fusion algorithm against disturbances and noises of the system; 2) stability and convergence analysis of the corresponding closed-loop control system.

Keywords: Multiple model adaptive control, switching/weighting, intelligent fusion.

1. Introduction

Multiple model adaptive control (MMAC) is an important control strategy to deal with uncertainties and nonlinearities in control theory and engineering.

There are mainly two types of multiple model adaptive control (MMAC), i.e., weighted MMAC[1-12] and switching MMAC[13-23]. The advantage of switching MMAC is the fastness in response to the system parameter change, but its drawback is the robustness to the disturbances and noises; and the advantage of weighted MMAC is the smooth of adaptive process under noise or disturbance, but its drawback is the slow response to the system parameter change. Then we wanted to make use of advantages of these two schemes. For this goal, this paper proposes an MMAC scheme based on switch/weighting intelligent fusion algorithm. The contributions of this paper are as follows: 1) switching/weighting intelligent fusion algorithm against disturbances and noises of the system; 2) stability and convergence analysis of the corresponding closed-loop control system.

The basic idea is to adopt a switch/weighting intelligent fusion algorithm joint with multiple local models and corresponding multiple local controllers to achieve global multiple model adaptive control. To be specific, this paper will adopt weighting algorithm to unify the switching function and weighting function to construct a switching/weighting intelligent fusion algorithm and a corresponding multiple model adaptive control system.

2. The Details of the Control System

Consider a plant with known structure information but with unknown or jumping parameters

$$A(q^{-1})y(t) = q^{-d}Bu(t) + C(q^{-1})e(t) \tag{1}$$

where

$$\begin{cases} A(q^{-1}) = 1 + a_1 q^{-1} + \cdots + a_n q^{-n} \\ B(q^{-1}) = b_0 + b_1 q^{-1} + \cdots + b_m q^{-m} \\ C(q^{-1}) = 1 + c_1 q^{-1} + \cdots + c_l q^{-l} \\ d \geq 1, b_0 \neq 0 \end{cases} \tag{2}$$

A switching/weighting intelligent fusion algorithm based MMAC system consists of three parts, i.e., a model set to cover the uncertainty of the plant to be controlled,

*Corresponding author：W. Zhang, weicunzhang@ustb.edu.cn

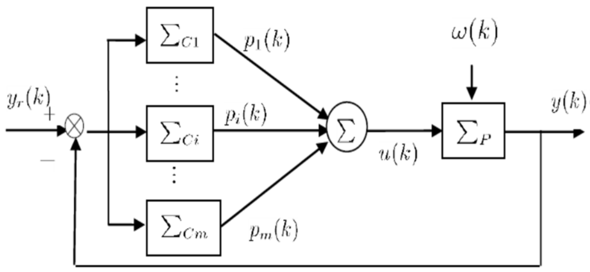


Fig. 1. The block diagram of the proposed system.

a controller set corresponding to the model set, and a switching/weighting intelligent fusion algorithm.

Model set consists of n fixed models to cover the uncertainties of the parameters in equation 1), i.e.,

$$\Omega = \{M_i | i = 1, \cdots, n\} \tag{3}$$

Controller set

$$C = \{C_i | i = 1, \cdots, n\} \tag{4}$$

Controller C_i is designed according to any existing control strategy which stabilizes M_i .and formulates a tracking closed-loop system. In this paper, C_i is designed according to pole-placement strategy. Denote the output of the 'local' controller C_i as u_i, then the global control is calculated by

$$u(k) = \sum_{i=1}^{n+1} p_i(k) u_i(k)$$

Switching/weighting intelligent fusion algorithm

It is designed to select the most suitable model and switch in the corresponding controller when the local model output errors $e_{mi}(k)$ are significantly different; or make weighted sum of all local control law when the local model output errors $e_{mi}(k)$ are not significantly different.

To be specific, suppose each local model M_i generates output error $e_{mi}(k)$, based on which the switching/weighting intelligent fusion algorithm is realized by the following equations (adapted from the weighting algorithm in Ref. 11).

$$l_i(0) = \frac{1}{n}, p_i(0) = l_i(0) \tag{5}$$

$$l_i(k) = 1 + \frac{1}{k} \sum_{p=1}^{k} e_{mi}^2(p) \tag{6}$$

$$l'_{\min}(k) = \min_i \{l'_i(k)\} \tag{7}$$

$$l_i(k) = \begin{cases} l_i(k-1) \frac{l'_{\min}(k)}{l'_i(k)}, \text{if sig} \left(\frac{l'_{\min}(k)}{l'_i(k)}\right) \le \alpha \\ \text{int}\left(\frac{l'_{\min}(k)}{l'_i(k)}\right), \text{if sig}\left(\frac{l'_{\min}(k)}{l'_i(k)}\right) \ge \alpha \end{cases} \tag{8}$$

$$p_i(k) = \frac{l_i(k)}{\sum_{i=1}^{n} l_i(k)} \tag{9}$$

Where sig$(\cdot)$ is a significance test function; int$(\cdot)$ is a integer-valued function; α is a confidence level which can be learned (or estimated) on line.

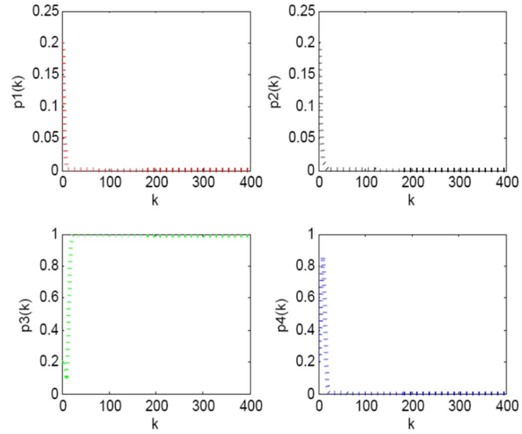


Fig. 2. The weighting signals of the proposed intelligent fusion algorithm.

3. Simulation Results

Consider the following plant:

$$y(t) + a_1 y(t-1) + a_2 y(t-2) = b_0 u(t-1) + b_1 u(t-2) + w(t) \tag{10}$$

where $w(t)$ is white noise with a zero mean and a variance of 0.01.

The deterministic part of Eq. (10) is obtained by converting the following continuous-time LTI model to a discrete-time model with sample time $t_s = 0.5$s and the zero-order hold.

$$\frac{k}{s^2 - 3s + 2} \tag{11}$$

The uncertainties of Eq. (10) originate from k in Eq. (10). For the real plant, we suppose $k = 1$.

For simplicity, the model set includes four models, corresponding to $k = 0.97, k = 0.98, k = 1.02$ and $k = 1.03$, respectively. Four 'local' controllers were designed by pole assignment strategy. Each controller stabilizes a possible model by formulating an expected closed-loop characteristic polynomial, say $A_m(q^{-1})$ and track the reference signal $y_r(k)$.

$$A_m(q^{-1}) = 1 - 1.3205q^{-1} + 0.4966q^{-2}$$

which corresponds to the characteristic polynomial of the following continuous time second-order system

$$A_m(q^{-1}) = \frac{\omega_n^2}{s^2 + 2\varsigma\omega_n s + \omega_n^2}$$

with $\varsigma = 0.707, \omega_n = 1$, and sampling time $t_s = 0.5$s.

The simulation results, i.e., the four weighting signals, the closed-loop output $y(k)$ against reference signal $y_r(k)$, and the control signal $u(k)$ are shown in Fig.2 and Fig.3, respectively.

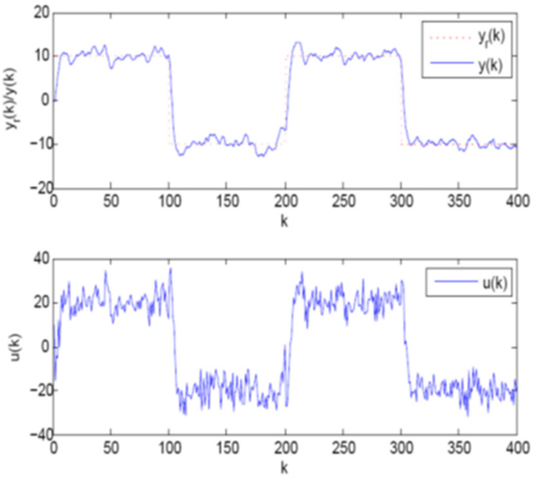


Fig. 3. The output, reference input, and control signals of the proposed MMAC system.

4. Conclusions

In this paper, a new MMAC scheme has been proposed with simulation verification. The main contribution is the proposed switching/weighting intelligent fusion algorithm which inherits the advantages of both switching and weighting.

Acknowledgements

This work was partly supported by National Natural Science Foundation of China (No. 61520106010; 61741302).

References

1. Fekri, M. Athans and A. Pascoal, Issues, progress and new results in robust adaptive control, *Int. J. Adapt. Control*, 2006, **20**(10): 519-579.
2. D. T. Magill, Optimal adaptive estimation of sampled stochastic processes, *IEEE Trans. Automat. Contr.*, **10**, 1965: 434-439.
3. D. G. Lainiotis, Partitioning: A unifying framework for adaptive systems-I: Estimation; II: Control, Proc. IEEE, 1976; 64:1126-1143 and 1182-1197.
4. M. Athans et al., The stochastic control of the F-8C aircraft using a multiple model adaptive control (MMAC) method-Part I: Equilibrium flight, *IEEE Trans. Automat. Contr.*, 1977, **22**: 768-780
5. A. Badr, Z. Binder, D. Rey, Weighted multi-model control, *Int. J. Syst. Sci.*, 1992; **23**(1): 145-149.
6. M. Kuipers, P. Ioannou, Practical Robust Adaptive Control: Benchmark Example, *Proc. of American Control Conference*, Seattle, Washington, USA, 2008:5168-5173
7. Kuipers and P. Ioannou, Multiple model adaptive control with mixing, *IEEE Trans. Automat. Contr.*, 2010; **55**(8), 1822-1836
8. S. Baldi, P. A. Ioannou, and E. Mosca, Multiple Model Adaptive Mixing Control: the Discrete-time case, *IEEE Trans. Automat. Contr.,* 2012; **57**(4):1040-1045
9. N. Sadati, G. A. Dumont, and H. R. Feyz Mahdavian, Robust Multiple Model Adaptive Control Using Fuzzy Fusion, 42nd South Eastern Symposium on System Theory, Tyler, TX, USA, March 7-9, 2010
10. Z. Han, and K. S. Narendra, New Concepts in Adaptive Control Using Multiple Models, *IEEE Trans. Automat. Contr.*, 2012; **57**(1): 78-89
11. W. Zhang, Stable Weighted Multiple Model Adaptive Control: Discrete-Time Stochastic Plant, *Int. J. Adapt. Control*, 2013; **27**(7): 562-581
12. W. Zhang, Further Results on Stable Weighted Multiple Model Adaptive Control: Discrete-Time Stochastic Plant, *Int. J. Adapt. Control*, 2015; **29**(12): 1497-1514
13. A. S. Morse, D. Q. Mayne, and G. C. Goodwin (1992), Applications of hysteresis switching in parameter adaptive control, *IEEE Trans. Automat. Contr.*, **37**, pp. 1343–1354.
14. A.S. Morse (1993), Supervisory control of families of linear set point controllers, *Proc. IEEE Conf. Decision Contr.*, San Antonio, TX.
15. B. Martensson (1986), *Adaptive stabilization*, Ph.D. dissertation, Lund Institute of Technology, Lund, Sweden.
16. K.S. Narendra, and Osvaldo Driollet (2001), Stochastic Adaptive Control Using Multiple Estimation Models. *Int. J. Adapt. Control and Signal Process*, Vol.15, pp.287-317.
17. K. S. Narendra and C. Xiang (2000), Adaptive control of discrete-time systems using multiple models, *IEEE. Trans. Automat.Contr.*, Vol.45, No.9, pp.1669-1685.
18. K. S. Narendra and J. Balakrishnan (1997), Adaptive control using multiple models, *IEEE. Trans. Automat. Contr*, Vol.42, pp.171-187.
19. K. S. Narendra and S. Mukhopadhyay (1994), Intelligent control using neural networks, *Intelligent Control*, M. M. Gupta and N. K. Sinha, Eds.
20. K. S. Narendra and J. Balakrishnan (1994), Improving transient response of adaptive control systems using multiple models and switching, *IEEE. Trans. Automat.Contr*, Vol **40**, pp.1861-1866.
21. Lingji Chen, K.S. Narendra (2001), Nonlinear adaptive control using neural networks and multiple models. *Automatica,* Vol.**37**, No. 8, pp.1245-1255.
22. M. Fu and B. R. Barmish (1986), Adaptive stabilization of linear systems via switching control, *IEEE Trans. Automat. Contr.*, Vol. **31**, pp. 1097–1103.
23. S. R. Weller and G. C. Goodwin (1994), Hysteresis switching adaptive control of linear multivariable systems, *IEEE Trans. Automat. Contr.*, Vol. **39**, pp. 1360–1375.
24. W. Zhang, On the Stability and Convergence of Self-Tuning Control-Virtual equivalent system approach, *Int. J. Control*, 2010; **83**(5): 879-896.

Detection Algorithm of Porosity and Crack Defects on Surface of Micro-precision Glass Insulated Terminals

Qunpo Liu*, **Gaowei Wang, Fuzhong Wang, Manli Wang**
School of Electrical Engineering and Automation, Henan Polytechnic University, 2001 Century Avenue Jiaozuo, (454003), Henan, P.R.China,

Hanajima Naohiko
College of Information and Systems, Muroran Institute of Technology, 27-1 Mizumoto-cho,Hokkaido Muroran-shi, Hokkaido (050-8585), Japan
E-mail: lqpny@hpu.edu.cn, 211807020075@home.hpu.edu.cn, wanghq2013@yeah.net, 311508070330@home.hpu.edu.cn, hana@mondo.mech.muroran-it.ac.jp
www.hpu.edu.cn, www.muroran-it.ac.jp

Abstract

A miniature precision glass-insulated terminal introduced by glass powder and metal wires through a special complicated process. Aiming at the two common defects of pores and cracks on the surface, a defect detection algorithm propose based on threshold segmentation and feature extraction. Pre-operation, global threshold segmentation processing and feature extraction (based on area, circularity aspect ratio, compactness, and contour length) are preformed to detect the defects. Experimental results show that the algorithm can accurately identify pores and seam defects.

Keywords: Micro-precision glass-insulated terminal; Manual inspection; Surface defect inspection; Feature extraction

1. Introduction

The micro-precision glass-insulated terminal (referred to as the glass terminal) is mainly used to connect a module, fix its wires and insulate it from the "tube shell", as shown in Fig. 1. The glass terminal structure is composed of three parts: the terminal, the high-temperature glass layer, and the low-temperature glass layer in order from the inside. Due to the low density of the low-temperature glass layer on the outside of the glass terminal, which is fluffy, and the material contains lead, the operating temperature during the sealing process is abnormal (such as too high or too low), and the accuracy of the equipment is not high or malfunctions. As a result, part of the surface structure of the low-temperature glass layer is not compact enough, thereby causing porosity and crack defects. Therefore, the detection of porosity and crack defects on the surface of glass terminals is mainly the detection of defects at the low temperature glass welded to the outside.

a) Glass terminal installation b) Glass terminal

Fig.1 Glass terminal and its installation

[*] Corresponding author

The glass terminal has a small volume (about 2.85mm in diameter) and requires high accuracy (less than 0.01 mm). At present, the detection of defects such as pores and cracks on the surface of the glass terminal achieve manually by combining with a magnifying glass. However, people who work for a long time are prone to eye fatigue and labor intensity[1]. There are too many subjective factors during manual inspection. Over-fatigue caused by high-intensity work and fast inspection speed may lead to missed inspection or false inspection. As a result, the defective terminal assembly applied, resulting in poor or even unusable product quality and need to rework, causing huge economic losses. In addition, glass terminals are in great demand in some respects and show expansive growth. However, the automatic detection of porosity and crack defects on the surface of glass terminals at home and abroad is still blank. Therefore, the automatic detection of pores and crack defects on the surface of glass terminals has become a major problem that needs to solve urgently.

For the research of surface defect detection, the Refs. 2-7 and so on use deep learning or support vector machine to detect the surface defect of the object, but the calculation time is long and the sample demand is large. The Refs.8-11 conducts defect detection by extracting features such as geometric features of defect areas on the surface of objects, and then identifies and classifies the defects. The method is simple, fast and practical. Based on the characteristics of pores and crack defects on the surface of glass terminals, a defect detection of surface algorithm for micro-precision glass insulated terminals based on threshold segmentation and feature extraction was proposed.

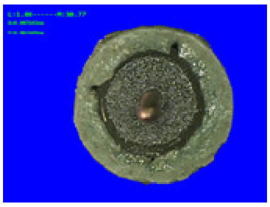
2. Detection algorithm basic structure

The algorithm uses Halcon HDevelop (18.11) to perform pre-processing operations such as background removal, and G channel acquisition of the image with defects. Then, these two defects are detected using global threshold segmentation processing and feature extraction methods based on area, circularity, aspect ratio, compactness, and contour length, and the characteristic parameters of the defect area are calculated and analyzed. Finally, the classification is based on the difference between the pore and crack defect characteristics, and the specific process is image pre-processing, defect detection, feature parameter analysis, and final classification decision.

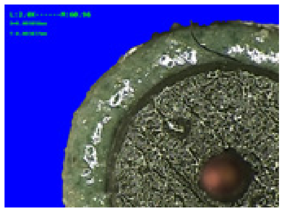
3. Image pre-processing

3.1. *Background removal*

The pictures acquire with an image-measuring instrument and the colored image is 640 × 480. Among them, the magnification of the collected picture and the length of each pixel obtained are: stomatal defect map--30.77 times, 0.0076245mm; crack -- 60.96 times, 0.0038475mm.

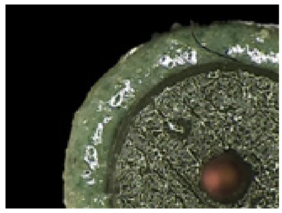


(1) Stomatal defect sample



(2) Crack defect sample

(3) Background removal of stomatal samples



(4) Background removal of cracked samples

Fig.2 The removal result of background

In order to facilitate defect detection, it is necessary to separate the end surface area of the glass terminal from the background to reduce the interference. Since blue does not occur naturally in glass terminals and is one of the three channels in the RGB color space, it is easier to segment the foreground area from a blue background than other colors[12]. In order to verify the feasibility of this method from the aspect of practical application, and use a large number of samples with different colors for experimental verification and theoretical analysis. By manually changing the background to ten common colors such as black, gray, white, red, orange, yellow, green, cyan, blue, and purple, the experimental verification and comparative analysis of glass terminal defect samples were performed. In order to meet the requirements of practical applications, dope the background color appropriately, and it finally find that when the background is biased to pure blue, it can completely meet the requirements of removing the main background. In addition, because the color of the upper-left character of the collected picture is green, which is different from the main background blue color. It is necessary to fill the character area (that is, the area from the image origin to one-sixth long and one-third wide) to the purpose of process of background which include color image to HSV space, blue threshold MASK and bitwise negation to generate the target area MASK, make the original image and itself bitwise operate in the target area MASK, character area color fill. The removal result show in Fig.2. Where, the range of values of H, S, and V when generating MASK in Fig.2 is: H: [100,124], S: [43,255], V: [46,255].

3.2. *Obtaining the glass terminal G channel*

During the implementation of detection technology of machine vision, due to the requirements of detection speed and algorithm implementation, it is often necessary to convert color images into grayscale images for processing[13]. For the convenience of detection, the three-channel color image is converted into six single-channel grayscale images of R, G, B, H, S, and I. After testing and observation, it is found that the defect part of the glass channel of the G channel is more obvious, so the image of the G channel is selected as the detection map, and the G channel of the glass terminal is shown in Fig.3.

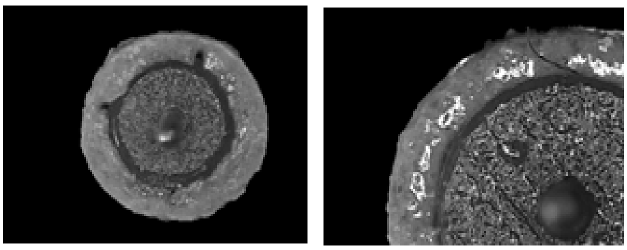


Fig.3 G-channel for extracting defective samples

4. Defect Detection

The porosity and crack defects on the surface of the glass terminal are more clearly identified by color, and gray value is lower than other surrounding areas, and they are approximately black. Therefore, global threshold segmentation and feature-based detection and recognition methods use for the detection of porosity and crack defects.

4.1. *Threshold Segmentation*

Global threshold segmentation perform on the pre-processed glass terminal image. The threshold segmentation show in Eq.(1).

$$G_{min} \le g(x,y) \le G_{max} \tag{1}$$

where, $g(x,y)$ is the gray-value of the image, G_{min} and G_{max} is the minimum pixel value (15 here) and the maximum pixel value (35 here). That is, the area formed by the gray-value of the image between the minimum pixel value and select the maximum pixel value in the image as the defect area to select.

After the above color feature processing, the detection range of glass terminal defects can reduce, and the distribution of suspicious defects can be roughly determined. Then, the defect area can be detected by the geometric features of the defects. The segmentation result is shown in Fig.4.

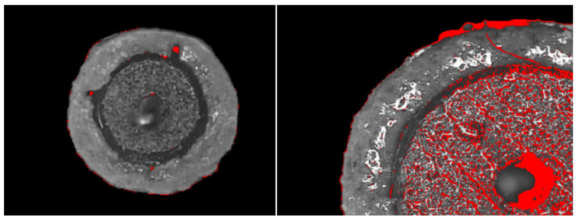


Fig.4 Threshold segmentation map (red area is a suspected defect area)

4.2. *Detection and identification of defect geometric features*

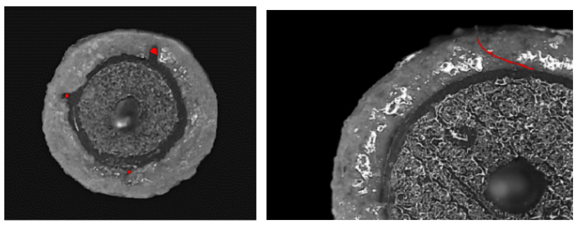


Fig.5 Defect detection results (red areas are defects)

According to the distribution of suspicious defects of the glass terminal, finally determining the pore and crack defects also requires multiple features for detection and identification. Common defect features in defect detection are area, circularity, perimeter, length, width, ratio of major axis to minor axis, defect center of gravity, circular or stripe defects, average grayscale and contrast, etc.[14]. This article mainly uses area, circularity, aspect ratio and compactness for feature selection. The glass terminal defect detection results are shown in Fig.5.

4.2.1 Defect area

The area of a defect can be simply defined as the number of pixels contained in the boundary of the defect, which is only related to the size of the target area, and has nothing to do with the gray value of the pixels at each point[8]. The area is determined as shown in Eq. (2).

$$A_s = \sum_{(x,y)\in s} f(x,y) \qquad (2)$$

where, S is the target area, $f(x,y)$ is pixel value, A_s is object area.

4.2.2 Roundness of defects

It can be defined as the ratio of the area to the maximum area, as shown in Eq. (3).

$$C = A_s/\pi d_{max}^2 \qquad (3)$$

where, d_{max} is the maximum distance from the defect center to the area boundary.

4.2.3 Aspect ratio of defects

The aspect ratio refers to the ratio of the major axis to the minor axis of the circumscribed ellipse of the region, as shown in Eq. (4).

$$\mathrm{R}_t = R_a/R_b \qquad (4)$$

where, R_a and R_b are the major and minor axes of the circumscribed ellipse, respectively.

4.2.4 Tightness of defects

Tightness is a form factor that represents the tightness of a region. The tightness of the circle is 1, if the area is long, the shape factor is greater than 1, and the longer the shape factor, the larger the shape factor, as shown in Eq. (5).

$$C_t = L^2/4\pi A_s \qquad (5)$$

where, L is the outline length of the area.

4.2.5 Defect contour length

The contour length refers to the total length of the contour line at the boundary of the specified area, that is, the sum of the distances of the pixels on the edge of the contour. Among them, the distance between two adjacent contour points parallel to the coordinate axis is 1, and the distance on the diagonal is $\sqrt{2}$.

5. Data analysis and classification decision

5.1. *Data analysis*

The characteristic values of feature detection and identification of stomatal and crack defects are shown in *Table 1.*

Table 1. Characteristic values of stomatal and crack defects

Defect type	pore	crack
As (mm^2)	[0.0018,0.01]	[0.0045,0.008]
C	[0.38,1]	[0.01,0.1]
Rt	[1.25,1.70]	[8,10]
C_t	[1.0,1.6]	[15,20]
Contour length (mm)	[0.07,0.68]	[0.76,1.46]

In the detection of stomatal and crack defects, image segmentation based on color features and various geometric features are mainly used to detect defects. According to the two defect detection algorithms, 10 glass terminals with stomatal and crack defects are randomly selected for experiments, and the range of various characteristic values of these two defects is found in *Table 2*. by experiments.

Table 2. Ranges of various characteristic values of defects

Defect type	pore1	pore 2	pore 3	crack
As (mm^2)	0.0093	0.0051	0.0022	0.0065
C	0.42	0.66	0.58	0.03
Rt	1.33	1.39	1.51	9.06
Ct	1.42	1.21	1.01	18.47
Contour length (mm)	0.41	0.28	0.17	1.23

As shown in *Table 2.*, as a whole, among the defect characteristics of glass terminal pores and cracks, the roundness characteristic value of the pore defect is much larger than that of the defects. The eigenvalues of aspect ratio, compactness and contour length of missing defects are much larger than those of stomatal defects.

5.2. *Classification decision*

According to the defect data in Table 2, it can be known that the roundness, aspect ratio, tightness, and contour length of the glass terminal can be used as the basis for defect classification decisions. Here, roundness C is used as the basis for classification decision. The decision algorithm is as follows:

```
if (C>0.25)
  disp_message（'pore'）
else
  disp_message（'crack'）
endif
```

The results after classification are shown in Figure 6.

Fig.6 Classification decision results

6. Conclusion

Based on the detection of pores and missing defects on the surface of glass terminals, a surface defect detection algorithm for glass terminals based on threshold segmentation and feature extraction is proposed based on the characteristics of the abnormal texture of the end surfaces of glass terminals. The background is removed by image pre-processing, and the defect can be obtained by threshold segmentation. Then, through the characteristics of area, roundness, aspect ratio, tightness, and contour length, the detection of pores and cracks on the surface of the glass terminal is realized. Laid the foundation for automatic quality inspection and assembly.

Acknowledgements

This work is partially supported by National Key Research and Development Project (2016YFC0600906), Innovation Scientists and Technicians Troop Construction Projects of Henan Province (CXTD2016054) and also supported by Innovative Scientists and Technicians Team of Henan Provincial High Education (20IRTSTHN019) .

References

[1] Shen Zhenteng, Chen Jianping, Du Houdong, et al. Research and application of machine vision-based insulator RTV coating quality inspection method [M]. 2018: 233-240.

[2] WANG Z, ZHU D. An accurate detection method for surface defects of complex components based on support vector machine and spreading algorithm[J]. Measurement, 2019,147:106886.

[3] HE D, XU K, WANG D. Design of multi-scale receptive field convolutional neural network for surface inspection of hot rolled steels[J]. Image and Vision Computing, 2019,89:12-20.

[4] OZTURK S, AKDEMIR B. Novel BiasFeed cellular neural network model for glass defect inspection: IEEE, 2016.

[5] Zhu Suya, Du Jianchao, Li Yunsong, et al. Bridge crack detection method using U-Net convolutional network [J]. Journal of Xidian University, 2019: 1-8.

[6] Li Yiting, Xie Qingsheng, Huang Haisong, et al. Surface defect detection based on fast region calibration based on convolutional neural network [J]. Computer Integrated Manufacturing System, 2019,25 (08): 1897-1907.

[7] Hu Jiacheng, Wang Xiangyang, Liu Ye. Deep defect detection of continuous casting slabs based on deep learning [J]. Journal of Shanghai University (Natural Science Edition), 2019 (04): 445-452.

[8] Chen Yanyan, Wang Yun, Ye Hanxiao, et al. Research on defect detection of machine vision toughened glass insulator bubbles [J]. Automation Instrumentation, 2010,31 (07): 51-53.

[9] Wang Fei, Cui Fengkui, Liu Jianting, et al. Research on glass surface defect detection system based on machine vision [J]. Glass and Enamel, 2009.

[10] Wang Shihao, Cai Yanguang, Cai Hao. Research on glass surface defect detection algorithm [J]. Journal of Dongguan University of Technology, 2018,25 (01): 45-49.

[11] Liu Chuanze, Wang Xiao, Chen Longxian, et al. Surface defect recognition of fiberboard based on random forest algorithm [J]. Forestry Science, 2018,54 (11): 121-126.

[12] ZHANG D, LEE D, TIPPETTS B J, et al. Date maturity and quality evaluation using color distribution analysis and back projection[J]. Journal of Food Engineering, 2014,131:161-169.

[13] Hou Yujia, Yang Wenliang, Wang Qiong, et al. Analysis of fracture causes and defect detection techniques of high voltage pillar porcelain insulators [J]. Inner Mongolia Electric Power Technology, 2019: 1-4

[14] Zhou Pengfei, Wang Fei, Xiao Hui, et al. Automatic detection and recognition of stomatal defects in DR images [J]. Nondestructive Testing, 2017,39 (10): 37-41.

Concept Drift Adaptation by Multi-stream Data Knowledge Transfer

Hongpeng Yin*, **Chenglin Liao, Yi Chai**
College of Automation, Chongqing University, Shazheng Street No.174
Chongqing, 400000, China
E-mail: yinhongpeng@gmail.com, liaochenglin1127@gmail.com, chaiyi@cqu.edu.cn
www.cqu.edu.cn

Abstract

The classifier of new concept cannot be adapted after concept drift unless massive new concept instances gathered. To handle this issue, this paper proposed a multi-stream knowledge transfer approach when just few new concept data are collected. Multi-stream data are represented by some labeled source data streams and one unlabeled target data stream. Several sub-classifiers learnt from source data streams individually are composed to an ensemble to predict the target. Empirical studies indicated the effectiveness comparing with other state-of-the-art methods.

Keywords: Concept drift, knowledge transfer, non-stationary environment, ensemble learning

1. Introduction

Machine learning tasks which data producing in a stream fashion have drawn growing attentions because of their wide existence in many real-world applications. Representative examples include industrial internet-of-things data analysis, financial data management, recommendation system and so on[1]. Usually, it is implicitly assumed that the process generating such continuously sequential data is stationary[2], which means the data are drawn from an unknow but fixed probability distribution. However, the assumption above is not true most of the time, in fact, the underlying distribution of incoming data may change overtime under a real-world scenario. This phenomenon, referred as concept drift[3], is one of the key challenges needed to be dealt with for deploying machine learning in real life applications.

Formally, concept drift between time point t_i and $t_{i+\Delta}$ is defined as:

$$\exists \mathcal{X} : p_{t_i}(\boldsymbol{x}, y) \neq p_{t_{i+\Delta}}(\boldsymbol{x}, y) \tag{1}$$

where the p denotes the joint distribution between input variables $\boldsymbol{x}$ and the target variable y [4]. Change of joint distribution makes the existing predicting model suffers performance drop according to Bayesian theory:

$$p(y \mid \boldsymbol{x}) = \frac{p(\boldsymbol{x}, y)}{p(\boldsymbol{x})} \tag{2}$$

As a result, the change of posterior probabilities $p(y \mid \boldsymbol{x})$ affects the prediction performance after concept drift (Fig. 1).

It is normal to train a new classifier from scratch after a concept drift was detected by a sliding window- based technique to adapt the new probability distribution[5].

However, amounts of labeled data of the new concept are needed in this kind of methods, which is unavailable

*Corresponding author.

in many real-world applications. What's worse, the accuracy of such approaches are tended to be poor when there have been not enough data of new probability for training, and the real-time demand that is commonly existed cannot be met.

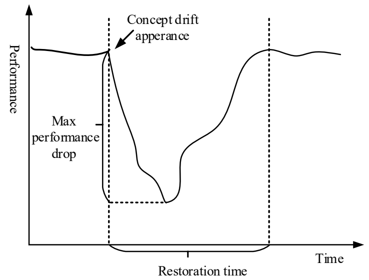


Fig. 1. Performance of classifier experiencing concept drift

To address the problems above, this paper proposed a multi-stream data knowledge transfer approach when just few new concept data are collected[6]. Naturally, ensemble methods offer a feasible approach combining knowledge from various classifiers to reduce the requirements of labeled data of new concept and restoration time by knowledge transfer[5]. The multi-stream data consist of multiple independent source data streams with labeled instances and one target data stream with unlabeled instances from the same domain. Knowledge learnt from the source streams are used to speed up updating the new concept[7]. Further, minimum the max performance drop and shorten the restoration time. For example, in remaining useful life prediction, data of other machine under the same working condition, the source data, can be helpful to improve prediction of the target machine. Even though the incoming data drawn from a new probability, i.e., concept drift occurred, the proposed method can transfer relevant knowledge from the source to overcome this issue. Empirical studies indicate the effectiveness comparing with other state-of-the-art methods, especially when just few target data are available.

2. The Proposed Method

An overview of our framework is shown in Fig. 2. In a non-stationary environment, there are N source stream with labeled instances and only one target data to predict. This framework works on an online paradigm, assuming that the target data are available sequentially in a fixed chunk, i.e., $B_1, B_2, ..., B_t, ...$. An ensemble classifier learnt on source data is used to improve the online classification model by re-weighting sub-classifiers of their performance on the target stream. As new instances arriving from the target, a drift detection model detects any kind of drift. If there is no drift beyond the threshold is detected, then the model is simply optimized by sub-classifiers re-weighting, otherwise, a sub-classifier is trained with concept new

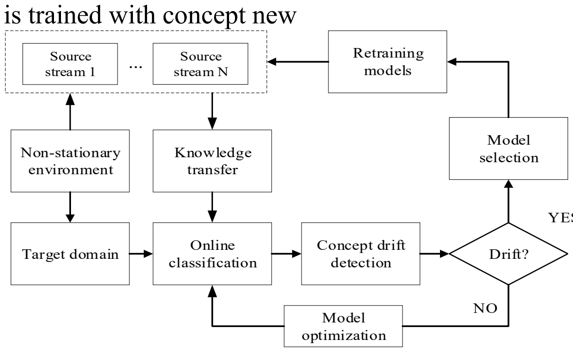


Fig. 2. Illustration of the proposed method

data. Some details of the proposed approaches will be discussed in the following sections.

2.1. *Drift Detection Method*

In the proposed approach, a Drift Detection Method (DDM)[8] is simply applied to discover concept drift. For each point j in a data chunk, the error rate is the misclassifying probability p_j with standard deviation given by $s_j = \sqrt{p_j\left(1-p_j\right)/j}$. Then, the instances will be stored in a buffer for retraining until $p_i + s_i$ reaches the minimum value, i.e., $p_{\min}$ and $s_{\min}$ if there comes a warning value:

$$p_i + s_i \geq p_{\min} + 2 \cdot s_{\min} \tag{3}$$

The concept drift is supposed to be true for the drift level:

$$p_i + s_i \geq p_{\min} + 3 \cdot s_{\min} \tag{4}$$

A new sub-classifier will be trained using the instances stored since the warning level reached.

2.2. *Knowledge Transfer Method*

The proposed method considers that the target concept is composed by several related concepts deduced from the source. When a new probability distribution comes, i.e., in the beginning of learning or after a concept drift, the proposed method can figure out which existing sub-classifiers share the most similarities with the new probability distribution, then transfer the knowledge learnt so far to build the new classifier much more efficiently.

The decision tree is employed as the base learner in the proposed method. First of all, N sub-classifiers will be trained with the source data streams, denoted as $\{f_1^t, ... f_N^t\}$. When a new data chunk B_t arrives, the ensemble classifier will firstly predict the labels of new arriving instances, then, making adaptation with different situation. A new sub-classifier f^t will be trained with newly collected data, and the initial weight of the new sub-classifier will be set by equation (5), if a concept drift is detected. Otherwise, a reweighting method is simply applied with equation (5) and (6). The weight for a sub-classifier is calculated according to the following equations:

$$w_t = \frac{1}{\text{MSE}_r^t + \epsilon} \tag{5}$$

$$w_i^t = \frac{1}{\text{MSE}_r^t + \text{MSE}_i^t + \epsilon} \tag{6}$$

where the MSE_i^t estimates the prediction error of the sub-classifier f_i^t, on the data chunk B_t, MSE_r^t represent the mean square error of a random classifier, ϵ is a very small positive number to avoid the denominator becoming 0. The details of the knowledge transfer method are shown in Table 1.

Table 1. The details of knowledge transfer method

Algorithm 1

Input: $B_1, B_2, ..., B_t, ...$: the data stream;
$\{f_1^t, ... f_N^t\}$: the sub-classifiers trained with source data stream;
Output: E_t :the new ensemble classifier

```
While data chunk B_t is available do:
    if Drift-detection == WARNING then
        Collect the new arriving data
    End if
    if Drift-detection == TURE then
        Train a new sub-classifier with newly collect data
    Else:
        Keep the current sub classifiers
    End if
    Reweight f^t with equation (5)
    Reweight {f_1^t,...f_N^t} with equation (6)
```

$$E_t = \left(\sum_i w_i^t f_i^t + w_t f_t\right) / \left(\sum_i w_i^t + w_t\right)$$

3. Experiment Setup

This paper focuses on adapt data steam distribution change after a concept drift with just few new instances of new concept collected. To prove the effectiveness of the proposed method, real-world datasets are used for verification.

3.1. *Datasets*

Two widely used real-world datasets SEA Concepts[9] and Usenet[10] are used. There are 60,000 examples, 3 attributes and 3 classes in SEA Concept. Attributes are numeric between 0 and 10, only two are relevant. And there are four concepts, 15,000 examples each. Usenet is a text dataset with 659 attributes and 5,931 examples simulated news filtering with a concept drift related to the change of interest of a user over time. Over time the virtual user decides to unsubscribe from those groups that he was not interested in and subscribe for two new ones that he becomes interested in. The previously interesting topics become out of his main interest. Attribute values are binary indicating the presence or absence of the respective word.

To generate multi-stream data, some instances from the real-world dataset were extracted as source dataset. In our experiment, 60% of the dataset are extracted as the source.

3.2. *Setup*

The following algorithms are compared in our experiments: SEA[9], DDD[11], Online Bagging[12] and an offline decision tree method. To make the comparisons fair, decision tree was employed as the base learner in all of these approaches. Every data chunk size is set to 300 instances in SEA Concept and 100 instances in Usenet to simulate the online environment. And all of the ensemble method set ensemble size to 25 according to Ref. 9.

3.3. *Performance*

The average classification accuracy and F1 score on Usenet dataset of every data chunk are shown in table 2. All the methods are also repeatedly performed ten times to report their average classification performance.

Table 2. Average Performance on Usenet

	DDD	Online Bagging	SEA	Offline Decision Tree	Proposed Method
Accuracy	0.509	0.515	0.493	0.495	**0.603**
F1 Score	0.401	0.323	0.171	0.448	**0.561**

It is obviously that the proposed approach achieves the best results compared to the competing method, both accuracy and F1 score. And the accuracy over time is shown in Fig. 3, which proves that even if there is a concept drift happened, the proposed method has the most stable performance over these five approaches. This shows that the proposed method benefits a lot from the knowledge transfer mechanism to recover perform from a concept drift, especially for small datasets. Table 3 shows the Average performance of these five approaches on SEA Concept. Based on accuracy and F1 scores, we can see that the proposed method presented the best results over SEA Concept.

Table 3. Average Performance on SEA Concept

	DDD	Online Bagging	SEA	Offline Decision Tree	Proposed Method

Accuracy	0.823	0.862	0.758	0.767	**0.877**
F1 Score	0.890	0.892	0.808	0.817	**0.903**

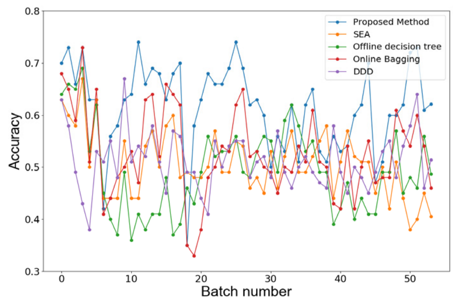


Fig. 3. Accuracy over time on Usenet

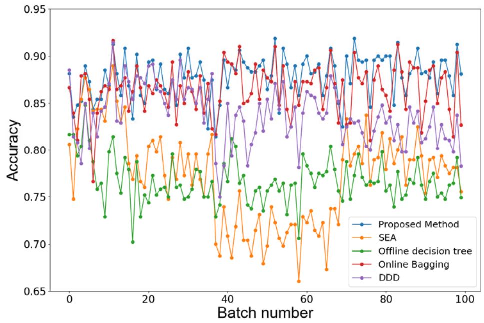


Fig. 4. Accuracy over time on SEA Concept

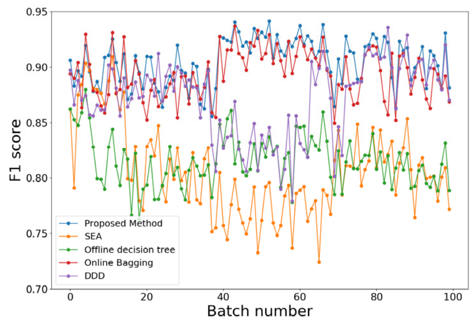


Fig. 5. F1 score over time on SEA Concept

4. Conclusion

In this paper, a multi-stream data knowledge transfer approach is proposed to overcome probability distribution changing in non-stationary environment. To minimum the max performance drop and shorten the restoration time after concept drift, multi-stream knowledge is utilized to train an ensemble classifier. Accuracy-based metric is used to measure the dynamic weight of sub-classifiers. Experiment results showed the effectiveness of the proposed approach.

However, a target concept may be composed by different sub-concept, just like the formula of total probability. Therefore, diversity-based sub-classifier selection approaches probably get better results in knowledge transfer.

Acknowledgements

This project was supported by Fundamental Research Funds for the Central Universities (2019CDYGZD001) and graduate research and innovation foundation of Chongqing, China (Grant No. CYS19072)

References

1. Krawczyk, B., Minku, L. L., Gama, J., Stefanowski, J., & Woźniak, M. (2017). Ensemble learning for data stream analysis: A survey. *Information Fusion*, *37*, 132-156.
2. Ditzler, G., Roveri, M., Alippi, C., & Polikar, R. (2015). Learning in nonstationary environments: A survey. *IEEE Computational Intelligence Magazine*, *10*(4), 12-25.
3. Gomes, H. M., Barddal, J. P., Enembreck, F., & Bifet, A. (2017). A survey on ensemble learning for data stream classification. ACM Computing Surveys (CSUR), 50(2), 23.
4. Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., & Bouchachia, A. (2014). A survey on concept drift adaptation. *ACM computing surveys (CSUR)*, *46*(4), 44.
5. Sun, Y., Tang, K., Zhu, Z., & Yao, X. (2018). Concept drift adaptation by exploiting historical knowledge. *IEEE transactions on neural networks and learning systems*, *29*(10), 4822-4832.
6. Haque, A., Tao, H., Chandra, S., Liu, J., & Khan, L. (2018, April). A framework for multistream regression with direct density ratio estimation. In *Thirty-Second AAAI Conference on Artificial Intelligence*.
7. Chandra, S., Haque, A., Tao, H., Liu, J., Khan, L., & Aggarwal, C. (2018, April). Ensemble direct density ratio estimation for multistream classification. In *2018 IEEE 34th International Conference on Data Engineering (ICDE)* (pp. 1364-1367). IEEE.
8. Baena-Garcıa, M., del Campo-Ávila, J., Fidalgo, R., Bifet, A., Gavalda, R., & Morales-Bueno, R. (2006, September). Early drift detection method. In *Fourth international workshop on knowledge discovery from data streams* (Vol. 6, pp. 77-86).
9. Street, W. N., & Kim, Y. (2001, August). A streaming ensemble algorithm (SEA) for large-scale classification. In *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 377-382). ACM.
10. Katakis, I., Tsoumakas, G., & Vlahavas, I. (2010). Tracking recurring contexts using ensemble classifiers: an application to email filtering. *Knowledge and Information Systems*, *22*(3), 371-391.
11. Minku, L. L., & Yao, X. (2011). DDD: A new ensemble approach for dealing with concept drift. *IEEE transactions on knowledge and data engineering*, *24*(4), 619-633.
12. Oza, N. C. (2005, October). Online bagging and boosting. In *2005 IEEE international conference on systems, man and cybernetics* (Vol. 3, pp. 2340-2345). Ieee.

Neuro-Adaptive Control of High-Speed Trains under Uncertain Wheel-Rail Relationship

Zhechen Wang*
The Seventh Research Division and the Center for Information and Control,
School of Automation Science and Electrical Engineering,
Beihang University (BUAA), Beijing 100191, P.~R.~China

Yingmin Jia
The Seventh Research Division and the Center for Information and Control,
School of Automation Science and Electrical Engineering,
Beihang University (BUAA), Beijing 100191, P.~R.~China
E-mail: yipin_yipin@163.com, ymjia@buaa.edu.cn

Abstract

Traditional automatic controller designing in train systems is almost based on urban rail transit where the influence of changing wheel-rail relationship caused by the variation of speed and environment is ignored. However, high-speed railway operates in more open environment and higher speed, which leading to a more complex variation of wheel-rail relationship occurring. In this paper, we design an automatic train controller in high-speed railway which can realize the automatic velocity tracking even if the uncertain and nonlinear variation of complex wheel-rail relationship happens. First of all, the train dynamic model is established where the wheel-rail relationship is expressed as an uncertain unknown function and the train operation system is expressed as a third-order nonlinear system. Then, a neural network adaptive controller is designed by using the backstepping method and barrier Lyapunov function. Based on this controller, position and velocity tracking errors are semi-globally uniformly ultimate boundedness. Finally, the effectiveness of the algorithm is verified by simulation experiments.

Keywords: Train velocity tracking, adaptive control, wheel-rail relationship, barrier Lyapunov function.

1. Introduction

Unmanned operation has been attempted successfully in urban mass transit. However, it fails in high-speed railway due to higher speed, wider range and more complex environmental changes. Precisely position and velocity tracking is one of the crucial task for high-speed railway. Thus, considerable research focuses on automatic operation or cruise control.

In the last decade, research in the field of cruise control could be categorized into three groups, PID control, robust adaptive control and intelligent control.[1] Typical PID control was the most widely used methods in train control system in early years.[2] Nevertheless, frequent switches of control law and appropriate parameter chosen are two issues that are not able to overcome. Intelligent approaches such as expert systems and reinforcement learning help improve the performance of safety, energy efficiency, comfort and punctuality as shown in Refs. 3-4. Robust adaptive approaches are generally model-based. As the problems of fault tolerant in Ref. 5, actuation notches in Ref. 6, antiskid constraints, input saturation in Refs.7-8, communication delays in Ref. 9 were considered about, train model got closer to the reality. In terms of the control methods, H∞ control in Ref. 10, terminal sliding mode control in Ref. 11 were used in research.

Although lots of research results have been achieved in the field of cruise control, the influence of complex relationship between wheels and rail on train control have not been considered. On the one hand, the wheel-rail model is difficult to describe by an accurate dynamic model. On the other hand, the adhesion coefficient changes drastically in different weather condition. These characteristics cause the increasing difficulty of controller designing. Some research studied the optimal adhesion control problem which means to find a controller to keep the wheel-real relationship in the best state which can be found in Refs 12-15. Different from optimal adhesion control, we are supposed to achieve the cruise control by taking advantage of the idea of active adhesion control. The train dynamics equation combined with the wheel dynamic equation formulate the system dynamics, which can express not only the relationship between rail and wheels but also the effect of the environmental changes on train dynamic system.

2. Problem Formulation

The multi-point motion model and wheel dynamic model of a high-speed train are as follows.

$$\dot{p}_i = v_i, \qquad (1)$$

$$m_i \dot{v}_i = \mu_i m_i g + f_{i+1,i} - f_{i,i-1} - f_{ri}, \qquad (2)$$

$$j_i \dot{\omega}_i = T_i - \mu_i m_i g r_i, \qquad (3)$$

where p, v, ω, T are train position, train velocity, angular velocity of wheels and the traction torque respectively. m, j, r and g are mass, wheel inertia, wheel radius and gravity constant respectively. The subscript i represents the ith carriage of the train. $f_{i+1,i}$ represents the coupler force between the (i+1)th and the ith carriage, which can be expressed as

$$f_{i+1,i} = k_{i+1,i} \cdot \left(p_{i+1} - p_i - l_{i+1,i}\right) + \delta\left(p_{i+1} - p_i - l_{i+1,i}\right), i = 2, \ldots, n-1.$$

$$f_{n+1,n} = f_{1,0} = 0, \qquad (4)$$

where $k_{i+1,i}$ is the stiffness and f_{ri} describes the resistance force of a carriage by Davis formula,

$$f_{ri} = a_i + b_i v_i + c_i v_i^2, \qquad (5)$$

where a, b and c are unknown positive constants with known upper bound. Adhesion coefficient μ which depends on plenty of factors, such as temperature, humidity, surface roughness, surface cleanliness, etc,[3] illustrates the complex wheel-rail system. According to previous research[13], μ can be described by a uncertain function related to the slip ratio given as

$$\lambda_i = \frac{\omega_i v_i - v_i}{v_i}. \qquad (6)$$

From fig.2, it is obvious that μ is an unimodal function in the first and the third quadrant respectively. Additionally, $d\mu/d\lambda > 0$ is satisfied in the stable region $\Omega := \left\{(\lambda, \mu) \middle| -\lambda_{\max} < \lambda < \lambda_{\max}\right\}$.

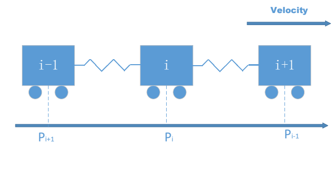


Fig.1. Train model

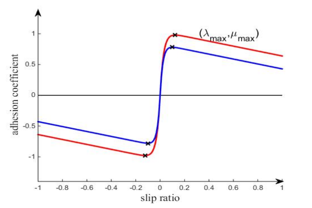


Fig. 2. The wheel-rail relationship

Based on the train model and wheel model, the object of this paper is to find a controller which is able to realize the following goals.

1. To force the velocity and position track a desired trajectory, i.e. $x_i(t) \to x_{id}(t)$, $v_i(t) \to v_{id}(t)$.
2. To ensure the coupler force in a safe range, i.e. $\left|f_{i+1,i}\right| \le f_{c,\max}, \forall t > 0, i = 2, \ldots, n-1$.
3. To ensure train anti-slip operation i.e. $|\mu| \le \mu_{\max}, \forall t > 0$.

It is easy to find that the problem we are talking about is a tracking problem for nonlinear and uncertain system with constrained output.

To this end, the following definitions, assumptions and lemma are imposed on system (1)-(3).

Define 1

$$x_i = p_i - \sum_{k=1}^{i} l_{k+1,k} \qquad (7)$$

Assumption 1

Slip ratio is in stable area when train moves, i.e.

$$\theta = \frac{d\mu}{d\lambda} \ge \underline{\theta} > 0, \forall t > 0.$$

Assumption 2

The position, velocity and adhesion coefficient of a train is measureable.

Lemma 1[16] *If there exists a C^1 continuous and positive definite Lyapunov function $V(x)$ satisfying $\kappa_1(\|x\|) \le V(x) \le \kappa_2(\|x\|)$ with bounded initial*

conditions, such that $\dot{V}(x) \le -\rho V(x) + c$,
where $\kappa_1, \kappa_2 : \mathbb{R}^n \to \mathbb{R}$ *are class* $\boldsymbol{K}$ *functions, and* ρ, c *are two positive constants, then the solution* $x(t)$ *is uniformly bounded.*

Take (7) into (1), we obtain that

$$\dot{x}_i = \dot{p}_i \tag{8}$$

In order to separate the uncertain parts from the system, (2) can be rewritten as

$$\dot{v} = \mu_i g + f_{i1} + \Delta f_{i1}, \tag{9}$$

where

$$f_{i1} = \frac{1}{m_i}\left[\hat{k}_{i+1,i}\left(x_{i+1} - x_i\right) - \hat{k}_{i,i-1}\left(x_i - x_{i-1}\right) - \left(\hat{a}_i + \hat{b}_i v_i + \hat{c}_i v_i^2\right)\right]$$

presents the certain part, $\Delta f_{i1} = \Delta f_{i1}\left(x_{i+1} - x_i, x_i - x_{i-1}, v_i\right)$ is the uncertain part, $\hat{k}_{i+1,i}, \hat{k}_{i,i-1}, \hat{a}_i, \hat{b}_i, \hat{c}_i$ are the estimated parameters obtained by experience and is closely near the real value. Similarly, through (2),(6) and (7), (3) is transformed as

$$\begin{aligned}\dot{\mu}_i = \theta_i \Bigg[& \frac{r_i}{j_i v_i} T_i - \left(\frac{m_i r_i^2}{j_i v_i} + \frac{\lambda_i + 1}{v_i}\right) g \mu_i \\ & - \frac{\left(\lambda_i + 1\right)\left(f_{i+1,i} - f_{i,i-1} - f_{ri}\right)}{m_i v_i}\Bigg],\end{aligned} \tag{10}$$

where $\theta_i = \dfrac{d\mu_i}{d\lambda_i}$, $f_{i2} = -\left(\dfrac{m_i r_i^2}{j_i v_i} + \dfrac{\lambda_i + 1}{v_i}\right) g \mu_i - \dfrac{\left(\lambda_i + 1\right) f_{i1}}{v_i}$ is the certain part, $\Delta f_{i2} = \Delta f_{i2}\left(\lambda, x_{i+1} - x_i, x_i - x_{i-1}, v_i\right)$ is the uncertain part.

For the sake of estimating the uncertain terms, neural networks is proved as an appropriate method. From Ref. 17, we know that for any given continuous function $f(X): R^m \to R$ and an arbitrary $\delta > 0$, there exists a neural network $W^{*T} h(X)$, such that $\left|f(X) - W^{*T} h(X)\right| \le \delta$, where

$$W^* := \arg\min_{W}\left\{\sup_{X}\left|f(X) - W^T h(X)\right|\right\} \tag{11}$$

is the optimal weight vector,

$$h(X) = \left[h_1(X), h_2(X), \ldots, h_l(X)\right]^T \tag{12}$$

is the Gaussian basis function vector which can be expressed as

$$h_i(X) = \exp\left(-\frac{\left\|X - c_i\right\|^2}{\sigma_i^2}\right), i = 1, 2, \ldots, l \tag{13}$$

In general, W^* and are bounded by the constants .
According to the NN estimation, system(8)(9)(10) can be detailed as

$$\begin{aligned}\dot{x}_i &= v_i \\ \dot{v}_i &= g\mu_i + f_{i1} + W_{i1}^{*T} h_{i1}(\cdot) + \varepsilon_{i1}(\cdot) \\ \dot{\mu}_i &= \theta_i\left(\frac{r_i}{j_i v_i} T + f_{i2}\right) + W_{i2}^{*T} h_{i2}(\cdot) + \varepsilon_{i2}(\cdot)\end{aligned} \tag{14}$$

Then the control problem becomes to a stability problem for system (14) with state constraints.

3. Controller Design and Stability Analysis

In this section, the backstepping procedure is employed to design a neuro-adaptive controller.
Denote that
$z_{i1} = x_i - x_{id}$, $z_{i2} = v_i - v_{id} - \alpha_{i1}$, $z_{i3} = \mu_i - \alpha_{i2}$.
Give the Lyapunov and barrier Lyapunov function candidate,

$$V_{i1} = \frac{1}{2}\ln\frac{z_{i1m}^2}{z_{i1m}^2 - z_{i1}^2}, \tag{15}$$

$$V_{i2} = \frac{1}{2} z_{i2}^2 + \frac{1}{2}\tilde{\varphi}_{i1}^2, \tag{16}$$

$$V_{i3} = \int_0^{z_{i3}} \frac{\tau \mu_{im}^2}{\mu_{im}^2 - \left(\tau + \alpha_{i2}\right)^2} d\tau + \frac{1}{2}\tilde{\varphi}_{i2}^2, \tag{17}$$

$$V_i = V_{i1} + V_{i2} + V_{i3} \tag{18}$$

where $\tilde{\varphi}_{ij} = \varphi_{ij} - \hat{\varphi}_{ij}, j = 1, 2,$ $\hat{\varphi}_{ij}$ is the estimation of φ_{ij} which is an adaptive parameter of NN weight vector to reduce the online computer burden.

Step 1.
Take derivative of (15), we have

$$\dot{V}_{i1} = \frac{z_{i1} z_{i2}}{z_{i1m}^2 - z_{i1}^2} + \frac{z_{i1}\alpha_{i1}}{z_{i1m}^2 - z_{i1}^2}. \tag{19}$$

Design the virtual control law as

$$\alpha_{i1} = -k_{i1} z_{i1}, \tag{20}$$

then

$$\dot{V}_{i1} = -\frac{k_{i1} z_{i1}^2}{z_{i1m}^2 - z_{i1}^2} + \frac{z_{i1} z_{i2}}{z_{i1m}^2 - z_{i1}^2}. \tag{21}$$

Step 2.
The derivative of (16) is

$$\begin{aligned}\dot{V}_{i2} &= z_{i2}\dot{z}_{i2} + \tilde{\varphi}_{i1}\dot{\tilde{\varphi}}_{i1} \\ &= g z_{i2} z_{i3} + z_{i2}\left(W_{i1}^{*T} h_{i1} + \varepsilon_{i1}\right) - \frac{z_{i1} z_{i2}}{z_{i1m}^2 - z_{i1}^2} \\ &\quad + z_{i2}\left(g\alpha_{i2} + f_{i1} - \dot{v}_{id} - \dot{\alpha}_{i1} + \frac{z_{i1}}{z_{i1m}^2 - z_{i1}^2}\right).\end{aligned} \tag{22}$$

From Young's inequality, we get

$$z_{i2} W_{i1}^{*T} h_{i1} \le \frac{z_{i2}^2 \varphi_{i1}^2}{2 a_{i1}^2} h_{i1}^T h_{i1} + \frac{a_{i1}^2}{2}, \tag{23}$$

$$z_{i2}\varepsilon_{i1} \le \frac{z_{i2}^2}{2b_{i1}^2} + \frac{b_{i1}^2\varepsilon_{i1m}^2}{2}. \tag{24}$$

Design the virtual control law as

$$\begin{aligned}\alpha_{i2} = &\frac{1}{g}\Big(-k_{i2}z_{i2} - f_{i1} + \dot{v}_{id} + \dot{\alpha}_{i1} \\ &-\frac{z_{i2}}{z_{i1m}^2 - z_{i1}^2} - \frac{z_{i2}\hat{\varphi}_{i1}}{2a_{i1}^2}h_{i1}^T h_{i1} - \frac{z_{i2}}{2b_{i1}^2}\Big),\end{aligned} \tag{25}$$

With the adaptive law

$$\dot{\hat{\varphi}}_{i1} = \frac{z_{i2}^2}{2a_{i1}^2}h_{i1}^T h_{i1} - \sigma_{i1}\hat{\varphi}_{i1}. \tag{26}$$

From (22)-(26), we have

$$\begin{aligned}\dot{V}_{i2} \le\ & g z_{i2}z_{i3} - \frac{z_{i1}z_{i2}}{z_{i1m}^2 - z_{i1}^2} - k_{i2}z_{i2}^2 + \frac{a_{i1}^2}{2} + \frac{b_{i1}^2\varepsilon_{i1m}^2}{2} \\ &+\tilde{\varphi}_{i1}\frac{z_{i1}^2}{2a_{i1}^2}h_{i1}^T h_{i1} - \sigma_{i1}\tilde{\varphi}_{i1}\hat{\varphi}_{i1}\end{aligned} \tag{27}$$

Step3.

The derivative of (17) is

$$\begin{aligned}\dot{V}_{i3} &= \frac{\partial V_{i3}}{\partial z_{i3}}\dot{z}_{i3} + \frac{\partial V_{i3}}{\partial \alpha_{i2}}\dot{\alpha}_{i2} + \tilde{\varphi}_{i2}\dot{\tilde{\varphi}}_{i2} \\ &= \frac{\mu_{im}^2}{\mu_{im}^2 - \mu_i^2}z_{i3}\theta_i\left(\frac{r_i}{j_i v_i}T_i + f_{i2}\right) \\ &\quad+\frac{\mu_{im}^2}{\mu_{im}^2 - \mu_i^2}z_{i3}\left(W_{i2}^{*T}h_{i2} + \varepsilon_{i2}\right) + \tilde{\varphi}_{i2}\dot{\tilde{\varphi}}_{i2},\end{aligned} \tag{28}$$

Design the control law as

$$\begin{aligned}T_i = &\frac{j_i v_i}{r_i}\Big[-k_{i3}z_{i3} - f_{i2} \\ &-\frac{z_{i3}}{\underline{\theta}_i}\frac{\mu_{im}^2}{\mu_{im}^2 - \mu_i^2}\left(\frac{\hat{\varphi}_{i2}}{2a_{i2}^2}h_{i2}^T h_{i2} + \frac{1}{2b_{i2}^2}\right)\Big],\end{aligned} \tag{29}$$

with the adaptive law

$$\dot{\hat{\varphi}}_{i2} = \frac{z_{i3}^2}{2a_{i2}^2}\left(\frac{\mu_{im}^2}{\mu_{im}^2 - \mu_i^2}\right)^2 h_{i2}^T h_{i2} - \sigma_{i2}\hat{\varphi}_{i2}, \tag{30}$$

Then we have

$$\begin{aligned}\dot{V}_{i3} \le\ & -k_{i3}z_{i3}^2 - g z_{i2}z_{i3} + \sigma_{i2}\tilde{\varphi}_{i2}\hat{\varphi}_{i2} \\ &+\frac{a_{i2}^2}{2} + \frac{b_{i2}^2\varepsilon_{i2m}^2}{2}.\end{aligned} \tag{31}$$

Thorem1. *Consider the train system (14) with the assumptions. If the control law (29) and the virtual control law (20)(25) with the adaptive law (26)(30) are implemented, then*

1) *The uniformly ultimate bounded position and velocity tracking is ensured;*
2) *The constraints of coupler force and anti-slip condition are not violated if the initial conditions are satisfied.*

Proof. From (15)-(18),(21),(27) and (31) we obtain

$$\begin{aligned}\dot{V}_i &= \dot{V}_{i1} + \dot{V}_{i2} + \dot{V}_{i3} \\ &\le -\frac{k_{i1}z_{i1}^2}{z_{i1m}^2 - z_{i1}^2} - k_{i2}z_{i2}^2 - k_{i3}z_{i3}^2 + \frac{a_{i1}^2}{2} + \frac{a_{i2}^2}{2} \\ &\quad+\frac{b_{i1}^2\varepsilon_{i1m}^2}{2} + \frac{b_{i2}^2\varepsilon_{i2m}^2}{2} + \sigma_{i1}\tilde{\varphi}_{i1}\hat{\varphi}_{i1} + \sigma_{i2}\tilde{\varphi}_{i2}\hat{\varphi}_{i2} \\ &\le -\rho_i V_i + c_i,\end{aligned} \tag{32}$$

where

$\rho_i = \min\{2k_{i1}, 2k_{i2}, 2k_{i3}, 2\sigma_{i1}, 2\sigma_{i2}\}$,

$c_i = \frac{a_{i1}^2}{2} + \frac{a_{i2}^2}{2} + \frac{b_{i1}^2\varepsilon_{i1m}^2}{2} + \frac{b_{i2}^2\varepsilon_{i2m}^2}{2} + \frac{1}{2}\sigma_{i1}\varphi_{i1}^2 + \frac{1}{2}\sigma_{i2}\varphi_{i2}^2$

By Lemma 1, we can obtain the results of uniformly ultimate bounded position and velocity tracking.

It is not difficult to prove by the barrier Lyapunov function that position error $z_{i1} \le z_{i1m}$ and adhesion coefficient $\mu_i \le \mu_{im}$. □

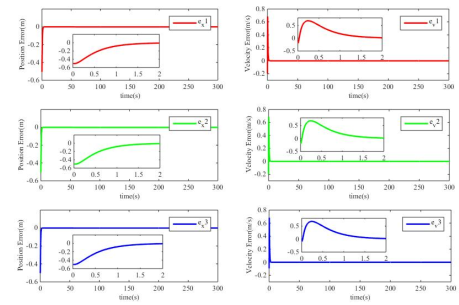


Fig.3. Position errors and velocity errors

4. Simulation Results

In this section, controller (29) will be verified for the train system (14) whose parameters are given as follows:

number of carriages $n = 3$, radius of wheels $(r_1, r_2, r_3) = (0.31, 0.31, 0.33)(m)$, inertia of wheels $(j_1, j_2, j_3) = (36, 36, 39)(kg \cdot m^2)$, mass of carriage $(m_1, m_2, m_3) = (10100, 10020, 10080)$ (kg) . The running resistance is expressed as $f_r = 1809 + 26v + 0.26v^2 + 0.1447\sin(0.02vt)$, where the last term represents the oscillation excitation force from the rail. The coupler force is calculated by $f_c = 800\xi(1 + 0.5\xi^2)$, where ξ is relative displacement between two adjacent carriages. The wheel-rail model use Burckhardt model given as $\mu = sign(\lambda)\cdot\kappa_1(1 - \exp(-\kappa_2|\lambda|)) - \kappa_3\lambda$.

The reference acceleration of the train and the parameters of Burckhardt model are as follows: $Acc = 0.6, 0 < t \le 100;\ Acc = -0.8, 200 < t \le 250;\ Acc = 0, otherwise.$
The initial state of the train is $v(0) = 11.5(m/s)$, $\omega(0) = 11.5(rad/s)$.
The performance of the controller is shown in Fig.3-4. Fig.3 indicates that position errors and velocity errors are convergent quickly. It is obvious to find that the tracking goals and coupler force constraints are satisfied. As shown in Fig.4, adhesion coefficient satisfies the requirement of antiskid control, $|\mu| \le \mu_{\max}$. shows the position error and velocity error of each carriage of a train.

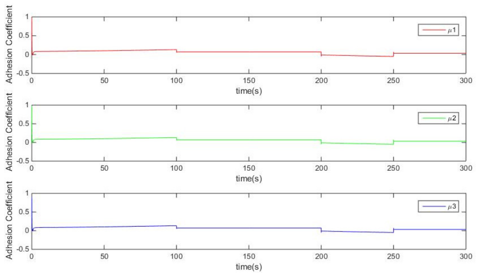


Fig.4. Adhesion coefficient of each carriage

5. Conclusions

In this paper, a neuro-adaptive controller is designed for train dynamic system with considering the uncertain nonlinear wheel-rail relationship. By using the barrier Lyapunov methods and backstepping designing methods, the tracking errors are convergent in a small region. The effect of the proposed controller is validated through the numerical simulation.

6. References

1. Jiateng Yin, Tao Tang, Lixing Yang, Jing Xun, Yeran Huang, Ziyou Gao, Research and development of automatic train operation for railway transportation systems: A survey, *Transportation Research Part C: Emerging Technologies*, Volume 85,2017
2. Oshima H, Automatic train operation system based on predictive fuzzy control *International Workshop on Artificial Intelligence for Industrial Applications*, 1988
3. Dong Hairong, Gao Shigen Extended fuzzy logic controller for high speed train, *Neural Computation & Application*, 2013.
4. J. Yin, D. Chen, L. Li, Intelligent train operation algorithms for subway by expert system and reinforcement learning, *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 6, pp. 2561-2571, Dec. 2014.
5. Z. Mao, G. Tao, B. Jiang, X. G. Yan, Adaptive actuator compensation of position tracking for high-speed trains with disturbances, *IEEE Trans. Veh. Technol*, vol. 67, no. 7, pp. 5706-5717, Jul. 2018.
6. D. Li, P. Li, W. Cai, X. Ma, B. Liu and H. Dong, Neural Adaptive Fault Tolerant Control for High Speed Trains Considering Actuation Notches and Antiskid Constraints, *IEEE Transactions on Intelligent Transportation Systems*, 2018.
7. Lin X, Dong H, Yao X, et al. PID-type sliding mode fault-tolerant control for high-speed trains using neural networks. In: *The 56th annual Conference on Decision and Control (CDC)*, Melbourne, Australia, 12–17 December 2017, pp.6364–6369.
8. Li, S., Yang, L., Gao, Z, Adaptive coordinated control of multiple high-speed trains with input saturation, *Nonlinear Dyn*, 2016, 83, (4), pp. 2157–2169.
9. K. Gao, Z. W. Huang, J. Wang, J. Peng, W. L. Liu, Decentralized control of heavy-haul trains with input constraints and communication delays, *Control Engineering Practice*, Vol.21, 420-427, 2013.
10. C. Yang, Y. Sun, Mixed H_2/H_∞ cruise controller design for high speed train, *Int. J. Control*, vol. 74, no. 9, pp. 905-920, Jan. 2001.
11. X. Yao, J. H. Park, H. Dong, L. Guo, X. Lin, Robust Adaptive Nonsingular Terminal Sliding Mode Control for Automatic Train Operation, *IEEE Transactions on Systems Man and Cybernetics: Systems*, pp. 1-10, 2018.
12. B. Picasso, D. Caporale, P. Colaneri, "Braking control in railway vehicles: A distributed preview approach", *IEEE Transactions on Automatic Control*, vol. 63, no. 1, pp. 189-195, 2018.
13. Chen Y, Dong H, Lu J, et al. A super-twisting-like algorithm and its application to train operation control with optimal utilization of adhesion force, *IEEE Trans Intell Transp Syst*, 2016, 17: 3035–3044.
14. W.-C. Cai, D.-Y. Li, B. Liu, Y.-D. Song, Uniform rolling-wear-based robust adaptive control of high-speed trains in the presence of actuator differences, *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 12, pp. 3591-3601, Dec. 2016.
15. W C Cai, D Y Li, Y D Song, A Novel Approach for Active Adhesion Control of High-Speed Trains Under Antiskid Constraints, *IEEE Transactions on Intelligent Transportation Systems*, pp. 1-10, 2015.
16. Wei He, Yuhao Chen, Zhao Yin. Adaptive neural network control of an uncertain robot with full-state constrains. *IEEE Transactions on Cybernetics*, 2016, 46(3):620-629.
17. Yanjun Liu, Jing Li, Shaocheng Tong, C. L. Philip Chen, Neural Network Control-Based Adaptive Learning Design for Nonlinear Systems With Full-State Constraints, *IEEE Transactions on Neural Networks & Learning Systems*, 2016, 27(7):1562.

Acknowledgements

This work was supported by the NSFC (61327807,61521091, 61520106010, 61134005) and the

National Basic Research Program of China (973 Program: 2012CB821200, 2012CB821201) .

Route planning of teleoperation mobile robot Based on the Virtual Reality Technology

Jiwu Wang
School of Mechanical and Electronic Engineering, Beijing jiaotong University
Beijing, Haidian District, China

Xuechun Yuan
School of Mechanical and Electronic Engineering, Beijing jiaotong University
Beijing, Haidian District, China
E-mail: jwwang@bjtu.edu.cn, 18121262@bjtu.edu.cn
www.bjtu.edu.cn

Abstract

Mobile teleoperation robot is one of the effective methods to help operators to work in complex environments. However, the time delay by distance is a key factor that restricts its application. To solve this problem, the motion trajectory of the robot is simulated using virtual reality technology and the obtained optimization data are applied to control teleoperation robot.

Keywords: teleoperation robot ; virtual reality ;route planning

1. Introduction

Teleoperation robots are one of the effective methods to help operators to work in complex environments (such as nuclear power plants, underwater, space, etc). However, one of the main problems in teleoperation is the time delay. [1]Specifically, in actual work, due to the distance between the operator′s side and the slave side of the robot is far away, the communication will inevitably occur time delay, resulting in the fact that the operator′s side and the slave side of the robot cannot achieve synchronous action.

In order to solve this problem, the motion trajectory of the robot is simulated by virtual reality technology and the obtained optimization data are applied to control teleoperation robot. In order to improve the efficiency of simulation, a 3D simulation environment of the real world is built in the unity 3D platform, and the physical model and kinematic model of the robot are established. At the same time, the environmental conditions of various adjustable parameters such as light and material are simulated in the platform. The physics engine for the robot is added in the simulation platform and the physical properties of the robot are configured.

Since there is no delay between the operator and the robot model, the robot model can immediately respond to the operator′s input, compensate for the influence of the delay, and assist the operator to complete the operation task smoothly and reliably.

2. The Construction of the Virtual Environment

The vraisemblance of virtual environment construction is an important guarantee to realize this project. Only by ensuring that the virtual world and the real world are as consistent as possible, can we ensure the accuracy of our entire operation. The construction of the virtual environment is divided into the construction of the robot model and the construction of the virtual simulation environment.

2.1. Construction of the robot model

2.1.1. *The construction of motion model*

Wheeled mobile robot is an important branch of autonomous mobile robot, which has the advantages of simple structure and easy control. It is often used as a platform to study robot positioning, navigation, path planning and other technologies.

The mobile robot with two wheels differential drive is composed of vehicle body, power supply, control system and positioning system. By controlling the speed of the two driving wheels, the robot can track different trajectories: when the speed and direction of the two wheels are the same, the robot moves forward in a straight line; when the speed of the two wheels is different, the robot turns. [2]The rear wheels are two independent driving wheels, while the front wheels are a universal wheel that can roll in any direction. The front wheel only plays a supporting role to keep the body stable. Its kinematic model is shown in figure 1.

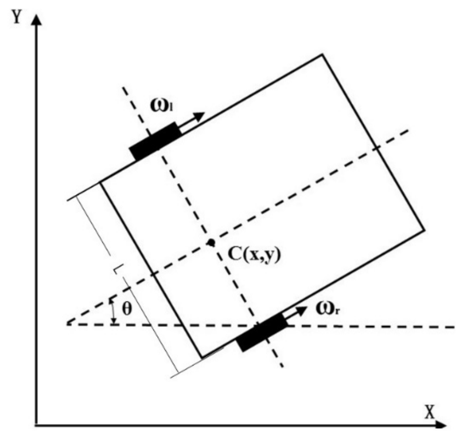


Fig. 1. Motion model of a mobile robot

Take the midpoint of the center line of the left and right wheels as the reference point. w_L and w_r are respectively the angular velocity of the left and right driving wheels. v and w are respectively the linear velocity and angular velocity of rotation of the robot. r and L are respectively the radius of the wheel and the distance between the two wheels. In the derivation process, $w_r > w_L$ is assumed, and its kinematic equation is shown as follows:

$$\begin{cases} x = v\cos\theta \\ y = v\sin\theta \\ \theta = w \\ v = r(w_L + w_r)/2 \\ w = r(w_L - w_r)/L \end{cases} \tag{1}$$

2.1.2. *The construction of physics model*

The robot is modeled by SolidWorks 2010. First, models of the parts of the robot such as head, arm and joint are built, then these parts are assembled. The assembly drawing is saved into the format of .stl and imported into 3d max, as shown in figure 2. After the model is output as .FBX format, it is imported into unity 3D for simulation, as shown in figure 3. After importing, set the physical properties for the robot in unity 3D, add the rigidbody component to the robot in unity 3D, and add the box collider component to each part.

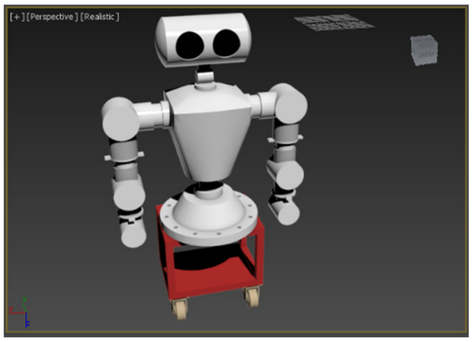


Fig. 2. The robot model in 3d max imported from Solidworks

Fig. 3. The components of the robot in unity 3D imported from 3d max

2.2. Construction of the virtual simulation environment

Virtual environment modeling is realized by synthesizing two-dimensional environment images into three-dimensional models.

Firstly, multiple photos of the surrounding environment are collected from the vision system of the end-robot, and the photos are transmitted to the ground control end. Then a three-dimensional model is generated through Agisoft Photoscan software. Finally, the generated model is imported into unity 3D, and the light source is set for the whole scene, material, and texture are set for the objects in the scene. The Sunken Square of Beijing Jiaotong University is modeled as an example, and the model after importing unity 3D was shown in the figure 4.

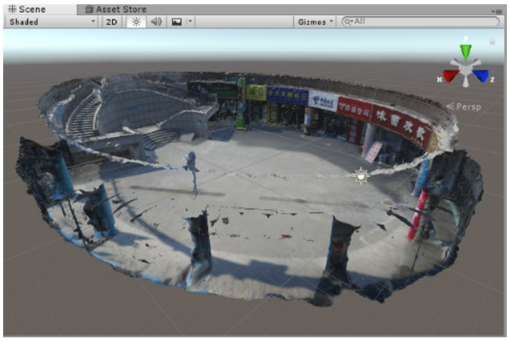


Fig. 4. The model in unity 3D imported from 3d max

3. Path Planning of Robots in Virtual Platform

Path planning is mainly implemented in unity 3D platform. First, several obstacles are set up in the scene. The four orange cuboids in the figure simulate obstacles. What we want to achieve is that by clicking on any position on the ground with the mouse, the robot can avoid obstacles and successfully reach the mouse click point. In this process, the walking track of the robot is displayed, and the three-dimensional coordinates of the walking track point of the robot are output to a. TXT file.

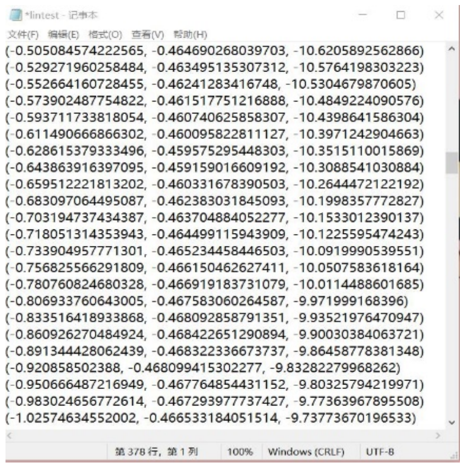
To realize the obstacle avoidance function of the robot, the pathfinding plug-in Navigation that comes with unity is used to realize automatic path finding through a series of calculations.

First, we baked the walking area of the robot, as shown in the blue area in figure 6, and added the Nav Mesh Agent component for the robot to set the speed and acceleration of the robot and other parameters.

Fig. 5. The walkable area of the robot

Write C# script and attach it to the robot to realize the function of mouse clicking on the ground robot to avoid obstacles and reach the target point, and add Line Render component to automatically draw the robot's walking track when the robot moves through the script. Use the StreamWriter class in C# to output real-time 3d coordinates of the robot as it moves to a.txt file, as shown in figure 6.

After the scene runs, the effect is shown in the figure 7 below. We can also zoom in and out the mobile scene through the middle mouse button.



*lintest - 记事本
文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)
(-0.505084574222565, -0.464690268039703, -10.6205892562866)
(-0.529271960258484, -0.463495135307312, -10.5764198303223)
(-0.552664160728455, -0.46241283416748, -10.5304679870605)
(-0.573902487754822, -0.461517751216888, -10.4849224090576)
(-0.593711733818054, -0.460740625858307, -10.4398641586304)
(-0.611490666866302, -0.460095822811127, -10.3971242904663)
(-0.628615379333496, -0.459575295448303, -10.3515110015869)
(-0.643863916397095, -0.459159016609192, -10.3088541030884)
(-0.659512221813202, -0.460331678390503, -10.2644472122192)
(-0.683097064495087, -0.462383031845093, -10.1998357772827)
(-0.703194737434387, -0.463704884052277, -10.1533012390137)
(-0.718051314353943, -0.464499115943909, -10.1225595474243)
(-0.733904957771301, -0.465234458446503, -10.0919990539551)
(-0.756825566291809, -0.466150462627411, -10.0507583618164)
(-0.780760824680328, -0.466919183731079, -10.0114488601685)
(-0.806933760643005, -0.467583060264587, -9.971999168396)
(-0.833516418933868, -0.468092858791351, -9.93521976470947)
(-0.860926270484924, -0.468422651290894, -9.90030384063721)
(-0.891344428062439, -0.468322336673737, -9.86458778381348)
(-0.920858502388, -0.468099415302277, -9.83282279968262)
(-0.950666487216949, -0.467764854431152, -9.80325794219971)
(-0.983024656772614, -0.467293977737427, -9.77363967895508)
(-1.02574634552002, -0.466533184051514, -9.73773670196533)
第 378 行，第 1 列 100% Windows (CRLF) UTF-8

Fig. 6. The three-dimensional coordinates of the robot in motion

Fig. 7. the operation result after improvement the scene runs

Draw the exported data points on the xz plane with Matlab, as shown in the figure 8.

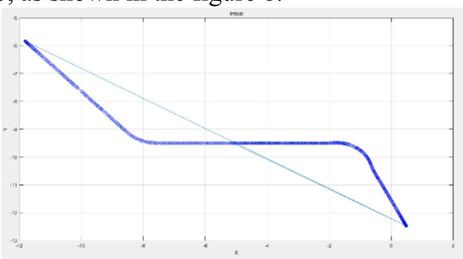


Fig. 8. The trajectory of the robot in the xc plane

4. Conclusion

In this paper, the robot motion simulation platform is built in the virtual reality environment. With this platform, the robot's motion trajectory was simulated for any specified target points. At the same time, the obstacle avoidance strategy was also tested in that platform. The satisfactory results were obtained. After the simulation, the true moving trajectory in the true environments for the robot is gotten. Then the true robot can be controlled to move based on the above data. In the coming future, the robot will be verified with the simulation data.

References

1. Bejczy A K,Kim W S,Venema S C，The phantom robot：predictive display for teleoperation with time delay (IEEE International Conference on Robotics and Automation，1990)
2. Farzad P, Mattias P K, Adaptive control of dynamic mobile robots with nonholonomic constrains. (Computers and Electrical Engineering , 2002(28):241-253)

Geometric Measurement Based on The Single Image with a Rectangular Structure

Jiwu Wang
School of Mechanical and Electronic Engineering, Beijing jiaotong University
Beijing, Haidian District, China

Bo Dai
School of Mechanical and Electronic Engineering, Beijing jiaotong University
Beijing, Haidian District, China
E-mail: jwwang@bjtu.edu.cn, 18121251@bjtu.edu.cn
www.bjtu.edu.cn

Abstract

Image-based geometric measurement is getting more attention in machine vision field due to its contact-less and low-cost characteristics. Here the measurement of single image with the rectangular structure is studied, in which only one side length of the rectangle is known, and the coordinates of points in the measured plane are obtained. The intrinsic parameters of the camera are calibrated firstly by using the three vanishing points of mutually orthogonal directions. Then the homography between the image plane and the model plane which the rectangle structure belongs to is derived, and the plane measurement method of the latter is described in details. Furthermore, the measurement experiment is done with the single image taken by the camera, and the validity of the proposed method is verified by comparing with real data.

Keywords: measurement; single image with rectangular structure; vanishing points; homography.

1. Introduction

As characterized by its non-contact and universality, real-time and repeatability, Image measuring gradually becomes one of the important methods for geometric measurement. In binocular measurement systems, precise matching between two images and camera calibration have always been classic problems in photogrammetry. Geometric measurement based on a single image is getting more attention in photogrammetry owing to the avoidance of such problems.

Considering what camera model is appropriate, Ref.1 made blade angle measurements of variable stator vanes in jet engines based on a single image.Ref.2 achieved plane metric rectification based a single view of multiple coplanar circles, but the camera's intrinsic parameters are known. The research was evaluated to verify the utility of a single image approach for scene of crime dimensional analysis in Ref.3.

Compared to theses above methods, we propose a new measurement method based a single image. The method avoids camera calibration by specialized equipment and only requires that the image scene contains parallel lines in three mutually orthogonal direction, where the length of one side of arbitrary spatial rectangle is known. The measurement task is completed in accordance with steps in section 2 and section 3.

2. Using vanishing points for camera calibration

The vanishing point is useful information for the measurement method based on a single image. It can be used to estimate intrinsic parameters and extrinsic parameters of the camera[4,5], and assist in measuring the height of target on the reference plane[6]. The accuracy of vanishing point detection directly affects the accuracy of measurement result. Currently, there are various methods for automatic vanishing point detection. Owing to inevitable interferences in some special scenes, these methods tend to be misjudged. In order to avoid such possible measurement errors, the method in this paper determines the parallel lines corresponding to three mutually orthogonal directions manually and estimates the vanishing point by using Least Squares.

As shown in figure 1, ***Px***, ***Py***, ***Pz*** are vanishing points corresponding to three mutually orthogonal directions. The parallel lines that determine the vanishing points are respectively parallel to ***OX***, ***OY*** and ***OZ*** coordinate axis in the spatial coordinate system ***O−XYZ***. ***C*** is the orthocenter of ∆***PxPyPz***, and ***D*** is the perpendicular foot on the side ***PxPy***.

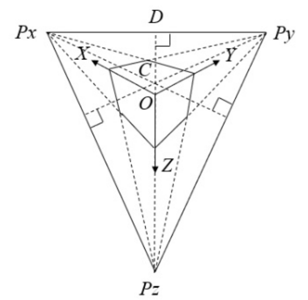


Figure 1: vanishing points in three mutually orthogonal directions

According to Ref.4 and Ref.6, there are such basic properties which are very useful for the purpose of camera calibration.

PROPERTY 1. Point ***C*** is the intersection of the optical axis and the image plane.

PROPERTY 2. The focal length of the camera satisfies the following formula.

$$f = \sqrt{|P_x D||P_y D| - |OD|^2}$$

Afterwards the camera intrinsic matrix ***K*** is determined.

$$K = \begin{pmatrix} f & 0 & u \\ 0 & f & v \\ 0 & 0 & 1 \end{pmatrix}$$

Let us consider a typical indoor scene with plenty of information about line segments as shown in figure 2. Coordinates of vanishing points corresponding to three mutually orthogonal directions on the image plane are estimated based on the above properties.

Figure 2: typical indoor scene with plenty of information about line segments

Px *= (1303.41, −2.06)*, ***Py*** *= (593.14,2014.20)*, ***Pz*** *= (−294.08,101.98)*. It is assumed that the distortion coefficient equals 0 and aspect ratio equals 0 and then the camera intrinsic matrix is calculated.

$$K = \begin{pmatrix} 730.37 & 0 & 486.51 \\ 0 & 730.37 & 376.96 \\ 0 & 0 & 1 \end{pmatrix}$$

3. Estimating the homography based vertexes of the rectangle

It is necessary to specify a spatial rectangle mapping with arbitrary quadrilateral in the image as a reference and the known length of one side of the rectangle serves as the input of the measurement system. We choose the spatial rectangle $A_1A_2A_3A_4$ corresponding to quadrilateral $a_1a_2a_3a_4$ on the image plane as the reference. Spatial points are represented by uppercase letters and subscripted numbers, and points on the image plane are represented by lowercase letters and subscripted numbers.

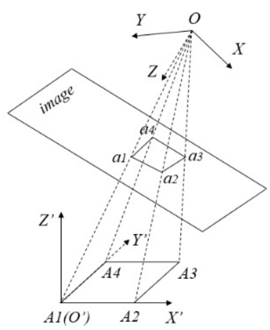


Figure 3: perspective projection schematic of spatial rectangle $A_1A_2A_3A_4$

In figure 3, the coordinate system $\boldsymbol{O-XYZ}$ is the camera coordinate system and the coordinate system $\boldsymbol{O'-X'Y'Z'}$ is the world coordinate system. According to the theory in Ref.7, if camera intrinsic parameters are known, the ratio of one side to the other adjacent side in the spatial rectangle is a function of camera intrinsic parameters and the coordinates of vertexes of quadrilateral $a_1a_2a_3a_4$ on the image plane.

$$\frac{|A_1A_2|}{|A_2A_3|} = \sqrt{\frac{(C_{12}x_2 - x_1)(C_{12}x_2 - x_1)}{(C_{13}x_3 - C_{12}x_2)(C_{13}x_3 - C_{12}x_2)}}$$

$$C_{12} = \frac{\det(x_3, x_4, x_1)}{\det(x_2, x_3, x_4)}, C_{13} = \frac{\det(x_4, x_1, x_2)}{\det(x_2, x_3, x_4)}$$

Notation: x_i represents the vector $\overrightarrow{Oa_i}$ in the camera coordinate system.

The length of $\boldsymbol{A_1A_2}$ and $\boldsymbol{A_2A_3}$ are respectively $\boldsymbol{a}$ mm and $\boldsymbol{b}$ mm based the above algorithm. In the world coordinate system $\boldsymbol{O'-X'Y'Z'}$, the coordinates of $\boldsymbol{A_1,A_2,A_3,A_4}$ are *(0,0,0), (a,0,0), (a,b,0), (0,b,0)*.The coordinates of $\boldsymbol{a_1,a_2,a_3,a_4}$ correspond to (x_1,y_1), (x_2,y_2), (x_3,y_3), (x_4,y_4) on the image plane.

Only the homogeneous solution for estimating the homography matrix $\boldsymbol{H}$ is described herein Ref.8. Each image-to-world point correspondence provides two equations which are linear in the elements of the matrix $\boldsymbol{H}$. They are:

$$xh_1 + yh_2 + h_3 - xXh_7 - yXh_8 - Xh_9 = 0$$
$$xh_4 + yh_5 + h_6 - xYh_7 - yYh_8 - Yh_9 = 0$$

The homography matrix $\boldsymbol{H}$ has 8 degrees of freedom[9]. For 4 correspondences, we obtain a system of 8 equations in eight unknows and then an exact solution about $\boldsymbol{H}$ is obtained. For the reference rectangle $\boldsymbol{A_1A_2A_3A_4}$, 8 equations from its vertexes can be rewritten as a matrix equation $\boldsymbol{Ah=0}$.

$$A = \begin{pmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1X_1 & -y_1X_1 & -X_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1Y_1 & -y_1Y_1 & -Y_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x_4X_4 & -y_4X_4 & -X_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -x_4Y_4 & -y_4Y_4 & -Y_4 \end{pmatrix}$$

$\boldsymbol{A}$ performs singular value decomposition and then $\boldsymbol{A=UDV^T}$. $\boldsymbol{h}$ equals the last column of the matrix $\boldsymbol{V^T}$. Afterwards h can be rewritten as a 3 by 3 matrix and that is $\boldsymbol{H}$.

$$H = \begin{pmatrix} 0.0002 & 0.0000 & -0.7296 \\ -0.0001 & -0.0001 & 0.6839 \\ 0.0000 & 0.0000 & 0.0008 \end{pmatrix}$$

4. Performing Experimental Measurement

As shown in figure 4, we select 3 line segments ($\boldsymbol{L_1}$, $\boldsymbol{L_2}$, $\boldsymbol{L_3}$) which all belong to the spatial plane corresponding to quadrilateral $\boldsymbol{a_1a_2a_3a_4}$, and measure the lengths based on the single image. The experimental results are compared with the results using handheld measurement tool.

Figure 4: estimation of length in typical indoor scene

Table 1: experimental results (mm) and results (mm) using handheld measurement tool

	L_1	L_1'	L_2	L_2'	L_3	L_3'
1	247.7	253.0	341.3	364.0	246.4	264.1
2	254.5	252.3	350.9	367.1	242.3	263.0
3	245.9	251.8	359.0	365.2	242.5	263.4

Three replicates of the measurement experiment are carried out and table 1 is sorted out after the experimental results are recorded accurately. $\boldsymbol{L_i}$ (i=1,2,3) represents the result based on the single image and $\boldsymbol{L_i'}$(i=1,2,3) represents the result using handheld measurement tool. From table 1, the relative measurement error does not exceed 8%. Taking the measurement principle into account, the error seems to be caused by the following two aspects:

- Vanishing point detection involves two steps of manually selecting parallel lines and solving overdetermined linear equations by using Least Squares when using vanishing points for camera calibration.
- There is no doubt that extracting the coordinates of pixels causes the error. In case 1, the error of estimating the homography matrix occurs if the mouse chooses the adjacent pixels rather than the pixels corresponding to vertexes of the reference rectangle. In other case, selecting the coordinates of pixel wrongly is certain to cause the error when trying to measure based on the single image.

5. Conclusions

The paper proposes a new method of geometric measurement based on a single image which makes full use of line segments in the image and requires quite few conditions. In general, the method takes an obvious advantage than other methods owing to its flexibility and simplicity in these cases where measurement accuracy is not required to be so precise. Currently, ways for image acquisition present a diversified tendency without getting camera parameters. The significance of this method is becoming increasingly apparent. One key issue of future research is to solve possible problems when the method is applied in specific fields.

References

1. Mendonca P R S, Kaucic R. Single View Metrology: A Practical Example (IEEE Workshop on Applications of Computer Vision, 2008).
2. Lourakis M I A. Plane metric rectification from a single view of multiple coplanar circles (Image Processing (ICIP), 2009 16th IEEE International Conference on. IEEE, 2009).
3. Gonzalez-Aguilera D, Gomez-Lahoz J. Dimensional Analysis of a Crime Scene from a Single Image Image and Signal Processing, 2008. CISP '08. Congress on. IEEE, 2008).
4. Caprile B, Torre V. Using Vanishing Points for Camera Calibration (International Journal of Computer Vision, 1990, 4(2):127-139).
5. Criminisi A. Single-View Metrology: Algorithms and Applications (Joint Pattern Recognition Symposium. Springer Berlin Heidelberg, 2002).
6. CAO X, COROOSH H. Metrology from vertical objects (British Machine Vision Conference,2004).
7. VANDENHEUVEL F A. Exterior orientation using coplanar parallel lines (Proceedings of the 10th Scandinavian Conference on Image Analysis, 1997:71-78).
8. Criminisi A, Reid I D, Zisserman A. A plane measuring device (Image & Vision Computing, 1997, 17(8):625-634).
9. Hartley R, Zisserman A. Multiple View Geometry in Computer Vision, 2nd Edition (Cambridge University Press,2003. 87~129).

Heavy-duty Spherical Mobile Robot Driven by Five Omni Wheels

Wei Liu, Junyang Sun, Runjiao Wang, Guanwang Geng, Xiaolong Han

Mechanical Engineering Department, School of Mechanical, Electronic and Control Engineering,

Beijing Jiaotong University, Beijing, 100044, P.R.China

E-mail: liuw@bjtu.edu.cn

www.bjtu.edu.cn

Abstract

This paper presents a heavy-duty spherical mobile robot with 5 omni wheels (omni-directional wheels). The chassis is supported by four drive omni wheels, and they touch the internal spherical shell. One omni wheel arranged at top can prevent the chassis from overturning. Four omni wheels are driven by two motors so that both forward and steering motions can be achieved simultaneously. The four omni wheels can support heavy load comparing to one omni wheel. The mechanism of the robot is described in detail, and its motion and dynamics are built and analyzed. Preliminary multi-body dynamics simulation shows that it has good function and dynamic characteristics etc.

Keywords: Spherical mobile robot, Omnidirectional wheel, Omni wheel, Eccentric gravitational moment, Heavy-duty spherical mobile robot.

1. Introduction

The sphere is an excellent rolling body and it can easily change the rolling directions. The spherical structure offers the extraordinary motion properties that the other mobile robot could not provide. The characteristics of sphere inspire the enthusiasm of scholars to create the

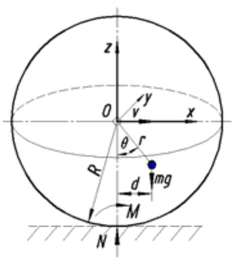


Fig. 1. Eccentric gravitational moment drive

Fig. 2. Omni wheel

spherical robots[1]. Research on spherical robots has been going on for nearly three decades. During the period, scholars have successively developed many different kinds of spherical robots. The existing spherical robots with omnidirectional motion could be divided into two categories according to the physical driving principles: 1) **Inertia moment driving type** relies on the rotating inertia moment to drive the omnidirectional rolling of a spherical robot.[2] 2) **Eccentric torque drive type** relies on the motion of eccentric mass centroid to produce the eccentric torque, and that drives the spherical robot rolling, as shown in Fig.1, simply the torque $M=mg \cdot r \cdot \sin(\theta)$ under the condition without considering the influence of rolling resistance. Only $\boldsymbol{M}$ is driving or steering the robot rolling. θ is the attitude angle. $\boldsymbol{r}$ is the distance from the sphere center to the mass centroid $\boldsymbol{m}$ directly determines the quantity of the driving moment $\boldsymbol{M}$. There are lots of spherical robots developed by using the eccentric moment mechanism to drive.[1] One of the earliest and most concise spherical robots was invented by Halme etc. (1996).[3] Its driving mechanism is made up of a driving wheel, a steering shaft and a balance wheel. Recent years, the several spherical mobile robots driven by omnidirectional or Mecanum wheel(s) are invented. The omnidirectional or Mecanum wheel(s) (shown in Fig. 2) simplifies the driving structures of spherical mobile robots.[4,5]

This paper proposes a novel spherical mobile robot mechanism, which uses 5 omni wheels to realize the driving principle of eccentric torque. As 5 omni wheels support the frame, they can bear a large load, just like the heavy truck supported by many wheels. The robot behavior such as detail mechanism, kinetics, dynamics and simulation of the robot are analyzed as following.

2. Mechanic Structure

The heavy-duty spherical mobile robot are supported by 5 omni wheels which touch the internal of spherical shell, as shown in Fig. 3. One omni wheel is on the top of frame to prevent the chassis from overturning. Other 4 ones are arranged under chassis. Their layout is shown in Fig 3

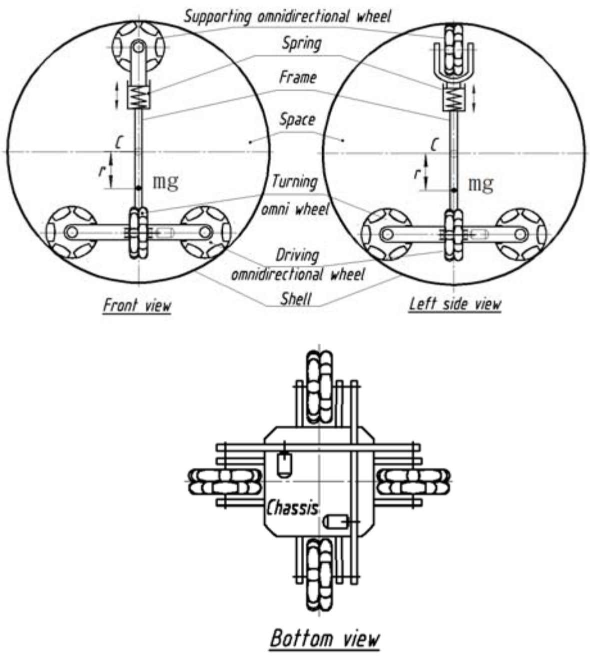


Fig. 3. The mechanism of spherical mobile robot with 5 omni wheels

bottom view. The omni wheels are arranged in vertical. Each pair is by one motor. One motor is in charge of moving forward and backward, and another motor for steering motions, both of them work simultaneously. As the weight of internal parts and spring provide pressure, it produces friction at the inner surface of the spherical shell when omni wheels driven by motors. That causes the eccentric mass ***m*** deviates from the vertical center line of the sphere (distance ***d***), so the eccentric torque ***M*** appears. ***M*** drives the sphere shell rolling and steering.

3. Mechanics Analysis

Before carrying out the mechanism mechanics analysis, we assume that: 1) the rolling contact of the whole system is pure rolling (i.e. temporarily excluding the influence of sliding); 2) all the components are rigid bodies, including omni wheels. In this way, the analysis can focus on key issues.

3.1. *Kinematics analysis*

G-XYZ is the coordinate system established on the ground, and *O-xyz* is the sphere inertial coordinate system established on the spherical center. Since the spherical shell always rotates with respect to the frame, the *O-xyz* rotates with respect to the frame, as shown in Fig.4. To facilitate the description of the frame position-orientation, the $O\text{-}x^f y^f z^f$ coordinate system is built on the mass centroid of the frame, as shown in Fig.5. In order to facilitate the kinematic analysis, the parameters are set as following: ***r*** is the distance between this mass centroid and the spherical center, $\boldsymbol{m_s}$ is the mass of the spherical shell, ***t*** is the wall thickness of the spherical shell, ***v*** is the velocity of the sphere center, $\boldsymbol{\omega_f}$ is the angular velocity of the sphere rolling, $\boldsymbol{r_0}$ is the radius of omni wheel. $\boldsymbol{\omega_0}$ is the angular velocity of the driving omni wheel, and $\boldsymbol{v_A}$ is the instantaneous velocity at point *A* outside the shell.

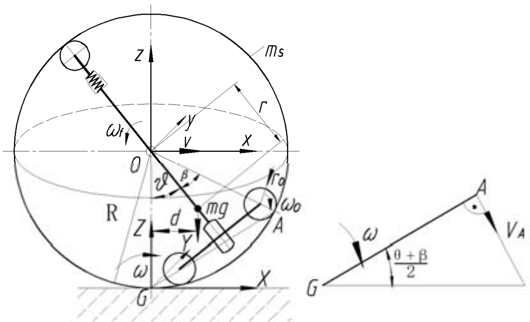


Fig. 4. The sphere inertial coordinate system and moving position

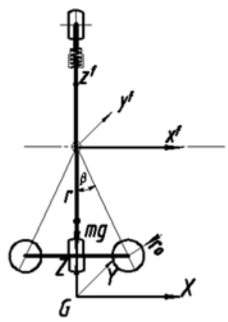


Fig. 5. The coordinate system of the frame

To demonstrate the theoretical and mechanical characteristics of the robot, its motion can be decomposed into several phases: start-up, uniform motion, obstacle crossing, taking off and ground landing etc.

At the phases of start-up and uniform motion, the velocity and acceleration at point *A* can be analyzed. As shown in Fig. 4, the absolute velocity of point *A* can be obtained from *G-XYZ*:

$$v_A = 2v \cdot \sin\frac{\theta+\beta}{2} \tag{1}$$

In addition, the absolute velocity at point *A* can be obtained by setting up a convected coordinate based on the sphere center:

$$\overrightarrow{v_A} = \vec{v} + \vec{R} \times \overrightarrow{\omega_f} - \overline{(r_o+t)} \times \overrightarrow{\omega_o} \tag{2}$$

where ω_f is the convected angular velocity of the frame, $\vec{v} + \vec{R} \times \overrightarrow{\omega_f}$ is the instantaneous velocity at point A, $\overline{(r_o+t)} \times \overline{\omega_o}$ is the relative velocity at point *A*. When $\omega_f = 0$, the robot is in a state of uniform movement.

Put the Eq. (2) into Eq. (1), the relationship between the angular velocity of the omni wheel $\boldsymbol{\omega_0}$ and the velocity of the spherical center $\boldsymbol{v}$ can be acquired.

The acceleration of point A is as follows:

$$\overrightarrow{a_A} = \vec{a} + \vec{R} \times \overrightarrow{\alpha_f} + \overrightarrow{\omega_f} \times \overrightarrow{\omega_f R} + \overrightarrow{(r_o + t)} \times \overrightarrow{\alpha_w} \quad (3)$$

where $\vec{a} + \vec{R} \times \overrightarrow{\alpha_f}$ is the instantaneous convected acceleration of the frame at point A, α_f is the relative acceleration of the frame; $\overrightarrow{\omega_f} \times \overrightarrow{\omega_f R}$ is the centripetal acceleration of the frame; $\overline{(r_o + t)} \times \overline{\alpha_w}$ is the relative acceleration at point A, α_w is the angular acceleration of the omni wheel, and the Coriolis acceleration in Eq. (3) is 0. In addition, the acceleration at point A can be obtained from G-XYZ：

$$a_A = a_t + a_n \quad (4)$$

Where $a_t = 2a \sin \frac{\theta + \beta}{2}$， $a_n = 2R\omega^2 \sin \frac{\theta + \beta}{2}$

The relationship between the angular acceleration at the spherical center and its acceleration is:

$$\alpha_O = a / R \quad (5)$$

3.2. *Dynamics analysis*

The parameters are set to facilitate kinetic analysis as follows, as shown in Fig. 6 and Fig. 7: $\boldsymbol{M_f}$ is the resistance couple of spherical shell rolling friction, $\boldsymbol{F_{ar}}$ is air resistance, $\boldsymbol{M_{wr}}$ is the rolling friction resistance couple of the walking omni wheel, $\boldsymbol{M_{ws}}$ is the rolling friction resistance couple of the steering omni wheel, $\boldsymbol{F_s}$ is the spring force on the frame, $\boldsymbol{F_{wr}}$ is the friction force of the walking omni wheel, $\boldsymbol{F_{ws}}$ is the friction force of the steering wheel, $\boldsymbol{F_p}$ is the frame pressure on omni wheels, $\boldsymbol{N_A}$,$\boldsymbol{N_B}$ are the pressure to the omni wheel inside the spherical shell, $\boldsymbol{M_d}$ is the driving torque of the omni wheel, and $\boldsymbol{M_e}$ is the effective torque.

The rolling of the spherical shell is due to the principle moment of the mass eccentricity. Here I_{shell} is the momentum moment of the spherical shell to the G-Y axis. According to the angular momentum moment, there is:

$$I_{(\theta)} = I_{shell} + I_{frame} = mg \cdot r \cdot \sin\theta - M_f \quad (6)$$

where, the moment of inertia of the spherical shell is

$$I_{shell} = \frac{2m_s}{5} \left(\frac{R^s - (R-t)^5}{R^3 - (R-t)^3} \right) + m_s R^2 \quad (6\text{-}1)$$

The moment of inertia of the frame is

$$I_{frame} = mr^2 + m\left((R - r\cos\theta)^2 + (r\sin\theta)^2 \right) \quad (6\text{-}2)$$

The resistance couple of spherical shell rolling friction is

$$M_f = \delta(m + m_s)g \quad (6\text{-}3)$$

The Eq. (6) can determine the angular acceleration of the spherical center as

$$\alpha = \frac{d\omega}{dt} \quad (6\text{-}4)$$

The dynamical motion equation is

$$I_{(\theta)}\alpha + \delta(m + m_s)g = mg \cdot r \cdot \sin\theta \quad (7)$$

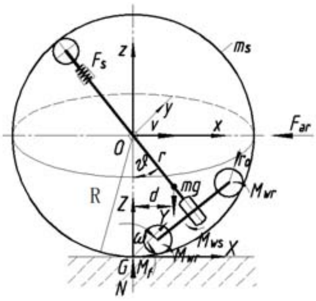


Fig. 6. The eccentric moment of couple

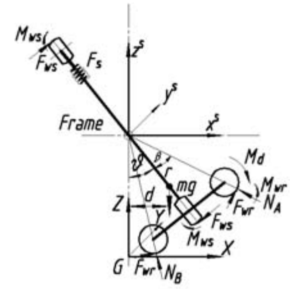


Fig. 7. The forces of the frame system

When considering the effect of the rolling friction resistance couple between the omni wheel and the spherical shell, the rolling resistance is shown in Fig. 8.

$$M_e = M_d - M_{wr} = F_{wr} \cdot r_0 - \delta N_A \quad (8)$$

Here, $\boldsymbol{F_c}$ is the force from chassis, the frame pressure on omni wheels is $F_P = N_A$, the rolling friction resistance couple of the walking omni wheel is $M_{wr} = \delta N_A$.

In the spherical shell coordinate system O^s-$x^s y^s z^s$, when the frame is in the attitude angle, the moment of couple reaches balance:

$$M_d - M_{wr} - M_{ws} - mg \cdot r \cdot \sin\theta = 0 \quad (9)$$

The equation can deduce the position angle θ.

The start-up phase can be divided into the following sub-phases: 1) When the driving torque of the omni wheel $M_d > M_{wr} + M_{ws}$,the omni wheel starts rolling inside the shell, the attitude angle $\theta > 0$; 2) When the eccentric moment accumulates to $mg \cdot r \cdot \sin\theta > M_f$, the robot shell starts rolling on the ground.

4. Simulation

The simulation purpose is to show the robot behaviors and verify the analysis above. The 3D solid model of the omni wheel spherical robot is built，as shown in Fig. 9. then some preliminary simulations of the kinematics and dynamics are tested by using ADAMS (Automatic

Fig. 9. 3D model of the spherical robot

Dynamic Analysis of Mechanical System). When the rotation speed of the omni wheel is constant, the angular velocity of the spherical robot changes periodically with time, as shown in Fig. 10. And the curve of the position angle θ is shown in Figure 11. The periodic phenomena

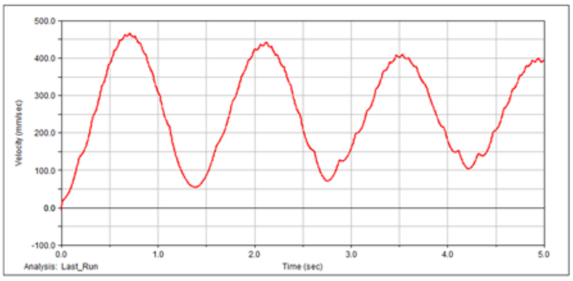


Fig. 10. Velocity of sphere under constant omni wheels speed

rise the interest waiting to explore them deeply in further study.

5. Conclusion

The spherical mobile robot with 5 omni wheels makes use of the characteristics of the omni wheel to simplify the driving mechanics and realizes the omnidirectional motion freely. The bottom 4 omni wheels can increase the load capacities. As the touch area of the sphere with the ground is limited and the sphere velocity is periodic, it is necessary to explore adaptive control method to improve the motion stability. As the robot is true four-

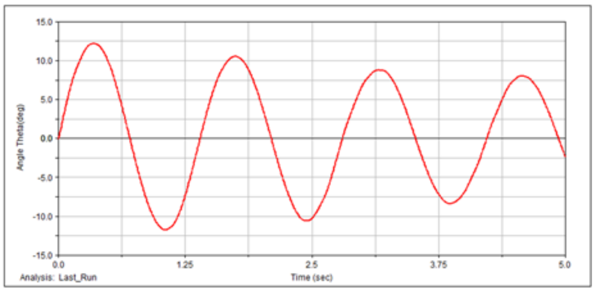


Fig. 11. The curve of the position angle θ by ADAMS

wheel drive and all the parts are not rigid, it is possible that one motor can cause the interference between the omni wheel pairs. So, it is necessary to explore the effective mechanism to avoid the interference. Furthermore, the behaviors of robot are waiting to explore further.

Acknowledgements

This research was supported by China National High Technology Research and Development Plan (863 Plan) 2013AA041302.

References

1. Rhodri H A, Julian F v., Rolling in nature and robotics: a Review, *Journal of Bionic Engineering*, **3**(4) (2006) p195-208.
2. Otani T et a1, Position and attitude control of a spherical rolling robot equipped with a gyro[A], International Workshop on Advanced Motion Control[c], Istanbul, Turkey：IEEE, 2006, p416—421
3. Halme A, Schonberg T, Wang Y., Motion control of a spherical mobile robot, *4th IEEE International Workshop on Advanced Motion Control*, (Mie, Japan, Jan,1996) p259–264.
4. Liu Wei, Ma Shuanglong, Duan Lunqin, Yu Jiangtao, Spherical Mobile Robot Driven by Biorthogonal Omniwheels, Proceedings of the 2018 International Conference on Artificial Life and Robotics, (Beppu, Oita, Japan, Feb. 2018) p174-177.
5. Liu Wei, Lian Luo, Jiwu Wang, Spherical Mobile Robot Designed with Single Omnidirectionsl Wheels, *Proceedings of the 2019 International Conference on Artificial Life and Robotics*, (Beppu, Oita, Japan, Jan. 2019) p 320-323.

Robustness Analysis of SLAM Algorithm in Simulation Environment of Gazebo

Jiwu Wang
School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University
Beijing, China

Yafan Liu
School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University
Beijing, China
E-mail: jwwang@bjtu.edu.cn, 17121246@bjtu.edu.cn
www.bjtu.edu.cn

Abstract

The SLAM algorithm applied to mobile robots is mainly used in complex and variable unknown environments. Therefore, in order to reduce the experimental cost and improve the robustness of the SLAM algorithm, it is necessary to test the SLAM algorithm in the virtual simulation environment before the robot enters the actual working conditions. The Gazebo used in this paper is a three-dimensional multi-robot dynamics simulation system, which can provide high-fidelity physical simulation environment and a complete set of robot sensor models for SLAM testing. In this paper, our method perform simulation tests on the Gmapping SLAM algorithm applied on the Turtlebot3 robot, and obtain experimental data and compare them. Then, according to the experimental results, we propose corresponding improvement suggestions for Gmapping algorithm.

Keywords: Gazebo, Gmapping SLAM, Kinematic simulation of Turtlebot3

1. Introduction

With the rapid development of some advanced technologies in the field of intelligent mobile robots, the applications of mobile robots have expanded to many aspects such as service, rescue, military, and space exploration, and the application environment has also expanded from structured to unstructured environments and even completely unknown environments.[1]

SLAM (Simultaneous Localization and Mapping), [2] it means that the robot acquires effective information of the surrounding environment via the sensors carried by the robot in the process of continuous movement, and completing map construction and real-time pose estimation for unknown environments. Currently, SLAM is classified by sensor type into laser SLAM (e.g. Gmapping)[3] and visual SLAM (e.g. ORB-SLAM2).[4] However, all kinds of sensors are very expensive, especially laser sensors. Therefore, in order to reduce the cost and avoid the accidental collision of physical robots in the unknown environment, it is essential to construct a virtual simulation environment for SLAM algorithm verification and simulation. In this paper, Gmapping, the most popular laser SLAM algorithm, is selected for simulation analysis. At the same time, Suggestions for improvement of the corresponding algorithm can be proposed based on the simulation results.

2. The construction of simulation environment

In this paper, gazebo simulation platform is selected to build a mobile robot simulation environment to test the robustness of the Gmapping algorithms. Gazebo is a 3D multi-robot dynamics simulation system. The system is a distributed architecture with independent libraries that provide high fidelity physical simulation and a complete set of sensor models. With the help of ROS（Robot Operating System）, the robot simulation models in Gazebo can subscribe to topics published by SLAM algorithms, and all simulation data of the robot can be displayed in Rviz(a 3D visualization tool), which is convenient for developers to debug the algorithm.

2.1. *Construction of the robot model*

The robot model invoked in the simulation environment is Turtlebot3 (size:*176mm*187.35mm*137.55mm*, differential control system, speed adjustable: *0.2m/s, 0.4m/s, 0.8m/s*) produced by ROBOTIS enterprise of South Korea, which can be equipped with Intel® Realsense™ R200 camera and 360 Laser Distance Sensor lds-01 radar, so as to meet the simulation requirements of the Sensor. The Turtlebot3 model is shown in figure 1.

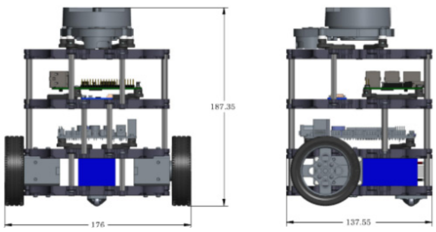


Fig. 1. Turtlebot3 model

2.2. *Construction of the virtual simulation environment*

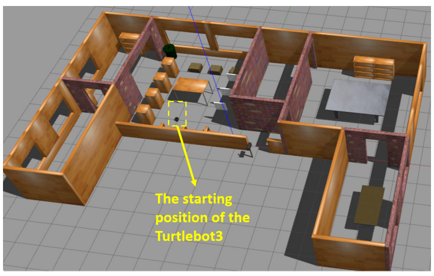


Fig. 2. 3D simulation environment from Gazebo

In our experiment, a simple indoor scene (model size: *15m*9m*1.5m*) was constructed by the gazebo simulation platform. In order to improve the fidelity of the model and increase the richness of the scene, all the walls were pasted Wooden texture; We set multiple fixed-size landmarks in the scene in advance, in order to facilitate the subsequent measurement of the map size constructed by SLAM; Meanwhile, to reduce the system calculation burden, the physical parameters of all models have selected default values. This model is shown in Figure 2.

2.3. *Gmapping algorithm simulation overview*

Gmapping algorithm needs to input the information of laser and IMU sensors to construct the map and estimate the position. The core idea of Gmapping algorithm is to use the Rao-Blackwellized Particle Filter (RBPF) to separate the locating and mapping process, and then construct the map after completing the locating. Therefore, Gmapping has a small amount of computation and high accuracy in constructing small scene maps.

In our method, the communication mode between gazebo, Turtlebot3, keyboard control port and Gmapping is shown in Figure 3.

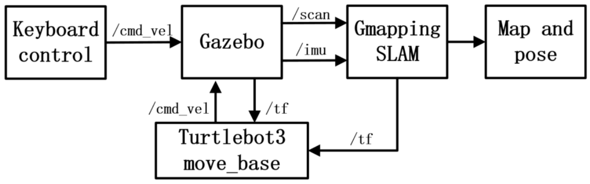


Fig. 3. Communication between functional nodes in ROS

The keyboard control port controls the movement of the Turtlebot3 robot through a differential controller. Turtlebot3 obtains its position based on the calculation of the IMU information. In the process of locating, the laser information obtained by Turtlebot3 is continuously transmitted to the Gmapping algorithm to construct the environmental map. At the same time, accurate map information can be used to adjust the robot position. Finally we output two-dimensional raster map and robot trajectory in Rviz. The display interface during simulation is shown in Figure 4.

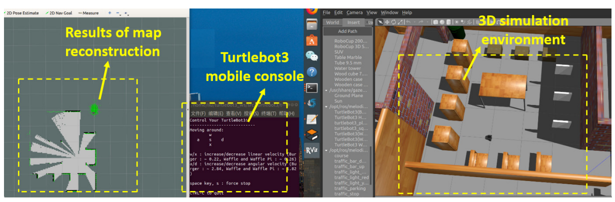


Fig. 4. Maps, consoles, and simulation environments in ROS

3. The Analysis of algorithm simulation results

In order to facilitate experimental comparison and data analysis, we set 10 different sizes and shapes of landmarks (known size) and a simple desk in the simulated experimental environment. The Turtlebot3 robot moves a complete circle around the desk, and completes the environment map construction and self-locating through the environmental information obtained by the sensors carried by robot. As shown in Figure 5, the left is the result of creating a two-dimensional map, and the right is a three-dimensional simulation environment.

Fig. 5. 2D map and corresponding 3D environment

After the mapping is completed, the map and the landmarks in the map are measured by the measuring rule in Rviz. The measurement results and the true values of the objects in the map are recorded in Table 1.

Tab. 1. Dimensional accuracy and analysis of markers in map

Object category	Measured values in map	Map size actual value	Accuracy
a.Overall map size	5.38m*4.96m	5.4m*5m	99.42%
b.Cabinet	0.536m	0.54m	99.36%
c.wooden case	0.122m	0.54m	23.60%
d.Cardboar box	0.459m	0.46m	99.82%
e.White board	Mapping failed	0.12m	Failed

According to Table 1, We can get the results that a complete map is successfully constructed and the size accuracy rate is as high as 99.42%; for b and d objects with rich texture and regular size, the size accuracy also reaches 99.36% and 99.82%; However, for c, which has a complex shape and is not within the view of the laser sensor, the construction of the object size is not accurate enough, only 23.6%; for a whiteboard or a reflective surface, the laser sensor does not obtain any information, and the map construction fails.

During the movement of the simulated Turtlebot3 robot, we calculated the robot's coordinate information, obtained the robot's motion trajectory, and plotted it as shown in Figure 6. It can be seen from the trajectory map that the experimental test data completely matches the real trajectory, which fully verifies the robustness of the Gmapping algorithm.

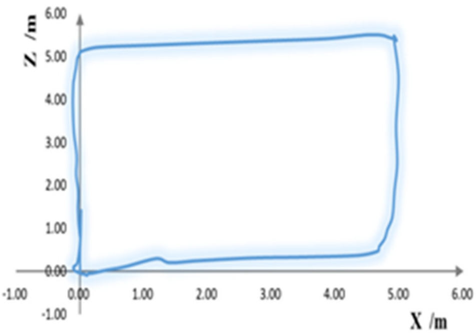


Fig. 6. Turtlebot3's motion trajectory in the simulation environment

4. Conclusions

Our method completes the construction of a high-precision map in Gazebo with the help of the Gmapping algorithm provided by Turtlebot3. According to the results of the Turtlebot3 kinematics simulation experiment, we learned that compared to white walls or reflective surfaces, the Gmapping SLAM algorithm is more accurate in locating and mapping in texture-rich scenes. Meanwhile, due to the use of two-dimensional laser sensors, the robot cannot obtain all the effective information in the environment, resulting in many holes and shadows in the constructed map. This problem may be improved in the future by constructing a more realistic simulation environment or improving our SLAM algorithm.

References

1. Petillot Y, Salvi J, et al. The SLAM problem: a survey [M]. AMSTERDAM: IOS PRESS, 2008:363.
2. Cadena C, Carlone L, Carrillo H, et al. Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age[J]. IEEE TRANSACTIONS ON ROBOTICS, 2016,32(6):1309-1332.

3. Grisetti, G., C. Stachniss and W. Burgard, Improved Techniques for Grid Mapping With Rao-Blackwellized Particle Filters. IEEE Transactions on Robotics, 2007. 23(1): p. 34-46.
4. Mur-Artal R, Montiel J M M, Tardos J D. ORB-SLAM: A Versatile and Accurate Monocular SLAM System[J]. IEEE TRANSACTIONS ON ROBOTICS, 2015,31(5): 1147-1163.

Hidden Surface Removal for Interactions between User's Bare Hands and Virtual Objects in Augmented Reality

Takahiro Ishizu, Makoto Sakamoto*, Kenji Sakoma, Takahiro Shinoda and Amane Takei
Faculty of Engineering, University of Miyazaki,
Miyazaki-City, Miyazaki, Japan

Takao Ito
Graduate School of Engineering, Hiroshima University,
Higashi-Hiroshima, Hiroshima, Japan

E-mail: hm14002@student.miyazaki-u.ac.jp, fruits2000jp@yahoo.co.jp
**Corresponding Author*

Abstract

Augmented reality (AR) technology is a technique of superimposing information generated by a computer on perceptual information that we receive from real space. Recently, much attention has been focused on interaction techniques between users and virtual objects, such as the user directly manipulating virtual objects with his/her bare hands. On the other hand, in AR technology, since the 3-dimensional (3D) model is superimposed on the image of the real space afterwards, it is always displayed on the front side than the hand. Thus, it becomes an unnatural scene in some cases (occlusion problem). In this study, this system considers the object-context relations between the user's hand and the virtual object by acquiring depth information of the user's finger. In the evaluation experiment, it is confirmed that the hidden surface removal in this study not only makes it possible to consider the object-context relations but also can distinguish between finger boundaries and to clarify and process finger contours.

Keywords: Augmented Reality (AR), Occlusion Problem, Hidden Surface Removal, Hand Detection.

1. Introduction

In recent years, augmented reality (AR) has become widespread [1]-[3]. Further, interaction with the virtual objects is required in those papers.

Therefore, we think that it requires a more accurate technology, because it is expected that interaction technology using AR technology will be more and more developed in the future.

On the other hand, since hands are our main means of interaction with objects in real life, it is necessary for AR interfaces to be able to manipulate virtual objects with the user's bare hand.

However, since the 3-dimensional (3D) model displayed by the AR is superimposed on the image of the real space afterwards, the 3D model is always displayed on the front side and user's hand is hidden by virtual objects. Thus, the scene may become an unnatural scene, and the user cannot see the object-context relations of the virtual object and his/her hand, and feels that it is difficult to manipulate the virtual objects.

In the existing study [4], the system used a transparent 3D model and the 3D model followed each fingers of the user. In this way, they performed hidden surface removal based on the depth information of user and 3D model. However, since the 3D model to be followed is larger than the finger, a wider range than the actual finger was displayed on the front (see Fig. 1).

In this study, we realize hidden surface processing along the fingertips of the user. Thus, we will be able to apply this system to more advanced interaction operations.

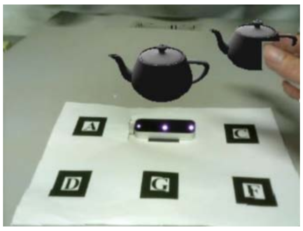


Fig. 1. Hidden surface removal in existing study.

2. Proposal

In order to realize hidden surface removal along the contour of the user's hand, this system detects hand area of the user by extracting depth information, color information and edge detection in the user's hand.

2.1. *System Component*

This system consists of a Web camera, the Leap Motion Controller, PC and AR marker.
In this study, in order to get accurate position information of user's fingers, we use the Leap Motion Controller.

This system controls the position and orientation of the virtual objects by recognizing the AR marker with Web camera.

However, when the user manipulates virtual objects, the AR marker may be covered by his/her own hand.

Therefore, this system adopts a method of treating Marker A to Marker F as one marker (see Fig. 2). Thus, even when a part of the marker is covered, the virtual objects can be displayed properly. We place the Leap Motion Controller on the position of Marker B.

This system displays the virtual objects on the position of Marker C and Marker G (see Fig. 2).

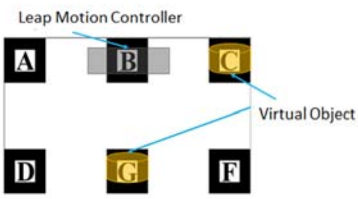


Fig. 2 AR marker.

2.2. *Determination of "skin color"*

This system extracts the area of skin color for detection of the user's hand area. Therefore, it is necessary to define the "skin color" color.

TaFirst, this system plots blue points on the finger which are the closest to the camera from the position information obtained from Leap Motion Controller (see Fig. 3).

Fig. 3 Blue Points.

This system converts RGB images obtained from a web camera into HSV images.

Next, the HSV image pixel corresponding to the pixel position where the blue point group is located is partially acquired.

In this study, in order to determine the range of "skin color", this system separates H, S and V for each pixel of the image, and creates the histograms. In the histogram, pixels with a small number are rejected. After that, a certain range is determined for the average value, and a color falling within the range is recognized as "skin color".

In this study, we assume that each of obtained H, S, and V data follow a normal distribution.
3 σ rule [5] is known as a simple method for outlier detection.

In this method, an observation is considered as an outlier if its least squares residual exceeds three times its standard deviation.

Thus, in this study, a pixel that falls within the range of 3σ from the average value of each of the HSV values of the colors acquired from the red dot group is taken as "skin color" using 3σ rule.

2.3. *Generation of binary image*

This system generates binary image to detect the user's hand area using "skin color".
The generated image contains noise, so remove it. (see Fig. 4)

Fig. 4 Binary Image.

2.4. *Generate image considering depth information*

This system acquires 3D coordinates of the distal bone, middle phalanx, basal bone and metacarpal bone of the user's hand with Leap Motion Controller.

Based on the acquired position information of the fingers, this system generates images plotting blue point group (following the position of the thumb) and green point group (following the position of a finger other than the thumb).

Blue point group displays at the positions of each joint of the thumb when the user's thumb is front of the back object (see Fig. 5).

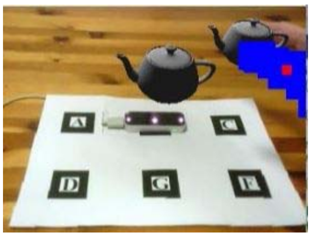


Fig. 5 Blue point group following the thumb.

In addition, green point group displays at the positions of each joint of the fingers except thumb when the user's fingers except thumb are front of the back object (see Fig. 6).

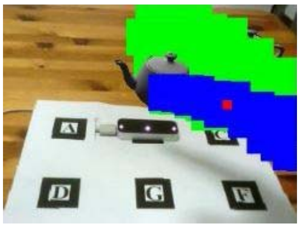


Fig. 6 Blue point group and green point group following the whole hand.

Furthermore, this system makes it possible to handle multiple virtual objects of different depths. The object-context relations between the back virtual object and the hand of the user is determined depending on whether or not blue point group and green point group are displayed. On the other hand, the object-context relations between the virtual object on the near side and the user's hand is determined by using the Z-buffer method.

Using the Z-buffer method, when a finger is positioned behind the virtual object, blue point group and green point group following the position of the finger are hidden by the virtual object.

In this way, this system can determine the object-context relations between the user's hand and the virtual objects even if there are multiple virtual objects.

2.5. *Generate images which always display the hand in front of virtual objects*

This system synthesizes the original RGB image (see Fig. 7-a) only in the white area in Fig. 4 to the image in Fig. 7-b. Thus, this system generates images that the user's hand is always displayed on the front of the virtual object (see Fig. 7-c).

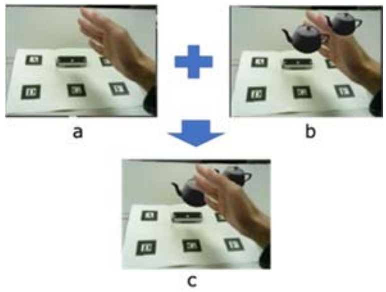


Fig. 7 Generation of images which hand is displayed front.

2.6. *Generate result image*

The area in Fig. 7-b corresponding the area of the blue point group, the green point group and the red point group (in the case where the position of the proximal phalanges of thumb is in front of the virtual object) of Fig. 8-a are replaced with the image shown in Fig. 8-b (Fig. 7-c).

Thus, based on the depth information, this system generates images that the hand is displayed in front of the virtual object (see Fig.8-c).

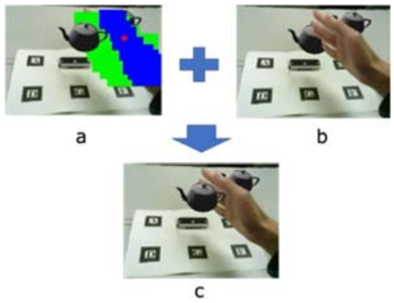


Fig. 8. Generation of result images.

3. EXPERIMENTAL RESULTS

We executed this system to confirmed whether hidden surface removal along the fingers is possible or not.

First, we checked whether this system can cope with multiple virtual objects.

The object-context relations between the virtual object on the near side and the user's hand is determined by using the Z-buffer method.

In order to confirm whether this is properly processed, we placed the virtual object at position G in Fig. 2 and performed manipulation that a user grabs the virtual object.

3.1. *In the case of an object behind*

Since only the thumb comes to front side than Marker C, the blue point group follows only the thumb (see Fig. 5) and draw the thumb properly on the front of the virtual object (see Fig. 9).

Fig. 9 Result at the position of C in the AR marker.

3.2. *In the case of an object in front*

This system conducted the hidden surface removal considering multiple virtual objects by the Z-buffer method. The blue point group follows only the thumb and the green point group follows the other fingers (see Fig. 6).

Besides, only the point group of the thumb was drawn on the front of the virtual object by the Z-buffer method. As a result, it was possible to draw only the thumb on the front (see Fig. 10).

The blue point group and green point group in Fig. 6 are based on the depth information of the user's fingers.

Also, the anteroposterior relationship between the blue point group and the green point group is correctly displayed by the Z-buffer method.

In this way, when the thumb is positioned behind the other fingers, the green point group is displayed on the front side of the blue point group.

Fig. 10 Result at the position of G in the AR marker.

4. Consideration

We confirmed that this system can perform correctly hidden surface removal even when a user manipulates multiple virtual objects by Z-buffer method from the experiment result.

The distinction between the virtual object and the user's hand was not clear when using the Z-buffer method when considering the depth between the virtual object and the user's hand.

Therefore, if there is only one virtual object or if the depth coordinates of the virtual objects are all at the same position, it is better to have blue point group follow the thumb of the user only than to have blue point group and green point group follow the palm of the user's hand.
On the other hand, the processing may be delayed, so we need to solve this.

5. Conclusion

In this study, we paid attention to detection the contour of the hand to realize hidden surface removal along the user's fingers. In this system, processing based on hand depth information, color information and edge detection is conducted to detect hand contours.

By these processes, hidden surface processing along the fingertip was realized.
In the evaluation experiment, it was shown that it can cope with multiple virtual objects.

References

1. Takahiro Ishizu, Makoto Sakamoto, Takaaaki Toyota, Masamichi Hori, Satoshi Ikeda and Amane Takei, *Proposal for Interaction Techniques for Intuitive Virtual Objects Manipulation in Augmented Reality*, Proceedings of the 2019 International Conference on Artificial Life and Robotics, pp.237-240, 2019.
2. Kangdon Lee, *Augmented Reality in Education and Training*, TechTrends, vol.56, no.2, pp.13-21, 2012.
3. Carmigniani J, Furht B, Anisetti M, Ceravolo P, Damiani E, Ivkovic M, *Augmented reality technologies, systems and applications*, Multimedia Tools and Applications 51:341-77, 2011.
4. R. Katahira, M. Soga. *Development and Evaluation of a System for AR enabling Realistic Display of Gripping Motions using Leap Motion Controller*, Procedia Computer Science 60:1595–1603, 2015.
5. Hekimoglu, S., and Koch, K. R. *How can reliability of the test for outliers be measured?,* Allgemeine Vermessungsnachrichten, 7, 247–253, 2000.

Proposal of Interactive Projection Mapping using Human Detection by Machine Learning

Takahiro Shinoda, Makoto Sakamoto*, Takahiro Ishizu, Kenji Sakamoto and Amane Takei
Faculty of Engineering, University of Miyazaki,
Miyazaki-City, Miyazaki, Japan

Takao Ito
Graduate School of Engineering, Hiroshima University
Higashi-Hiroshima, Hiroshima, Japan
E-mail: hm14021@student.miyazaki-u.ac.jp, fruits2000jp@yahoo.co.jp
**Corresponding Author*
E-mail: fruits2000jp@yahoo.co.jp

Abstract

In recent years, "Entertainment Computing" (EC) has attracted attention and has become one of the major industries in Japan. "Projection Mapping" is well known in this EC. Projection mapping is a video technology that creates a new space by synthesizing space and video using a projector. Among them, many people are fascinated by works that create a fantastic world by combining dancer performance and projection mapping. However, these works require the performer to accurately align with the coordinates of the image objects in the projection mapping, which is not easy for everyone. In this study, we aim to entertain not only the people who see the projection mapping but also performers. Therefore, we have prototyped an interactive projection mapping that changes according to user movement. This time, we focused on sports and projected the ball to the user to experience baseball pitching and soccer lifting. Furthermore, we conducted a questionnaire survey to evaluate the sense of use of this system, and the results showed that many people can enjoy by this projection mapping.

Keywords: Entertainment Computing, EC, Projection Mapping, Kinect.

1. Introduction

In recent years, exciting projection mapping has received more and more attention. Many people have seen projection mapping to buildings such as theme parks.

For example, there is a closing ceremony for the "Rio Olympics" held in August 2016.

Currently, projections are also being made on familiar items such as clothes, faces, and notebooks. For example, during a performance at the world's largest advertising festival in Cannes, the artist's white costume became a screen, and colorful graphics were projected one after another.

Projection mapping will delight us in many ways. However, in the conventional projection mapping, the viewer mainly enjoys watching and enjoying the projected image, so that the viewer's feeling of immersion in the content is considered insufficient [1].

In this study, we propose a participatory projection mapping that changes according to the movement of participants by projecting to participants.

2. Study method

2.1. *Equipment Used*

2.1.1. *Kinect for Windows*

Kinect for Windows is available from Microsoft. It's a peripheral device that enables operations by body movement, gestures, sound, etc. without using a controller.

Kinect has an infrared sensor, an RGB camera that acquires 8-bit 3-channel (RGB) image data, a depth image sensor that acquires distance (depth) image data from Kinect, and a sound microphone that estimates the position of a sound source.

Also, Kinect's feature is "posture recognition technology", recognizes the whole human body and performs operations by its movement. Thereby, it is possible to estimate where the parts of the body are by using the depth image.

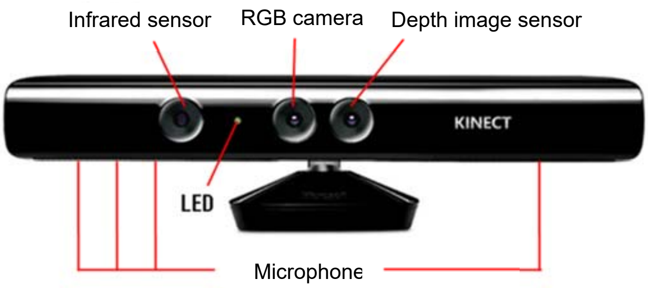


Fig. 1. Kinect for Windows v1 [2]

2.1.2. *Projector*

The projector is used for projection.

2.2. *Development Environment*

- OS: Windows10,
- Development Environment: Visual Studio 2017,
- Development Language: C++,
- OpenNI2,
  OpenNI is an interface that allows you to configure Kinect, acquire images, measure distance, and use a microphone.
  It is also used in this study because it can be linked with OpenCV, an image processing library.
- NiTE,
  NiTE is used to get data (eg gesture recognition, motion tracking etc) from Kinect [2].
- OpenCV,
  Create an object detector using the OpenCV library.
- OpenGL.

3. Outline of the system

It was programmed to display seven screens: Ball, Color, Depth, User, Skeleton, Combination (for projection), and Combination PC (for PC).

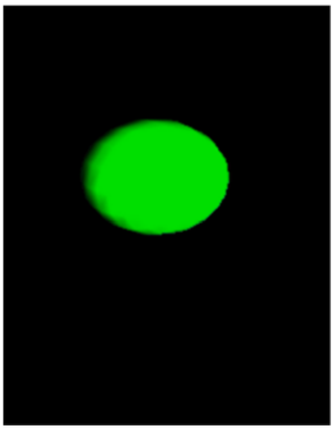


Fig. 2. "Ball" screen display

Fig. 3. "Color" screen display

Fig. 4. "Depth" screen display

Fig. 5. "User" screen display

Fig. 6. "Skeleton" screen display

Fig. 7. "Combination" " Combination PC" screen display

3.1. *User Detection*

3.1.1. *Boosting*

Boosting is a learning algorithm that sequentially generates weak classifiers and combines them to create a strong classifier [3]. In this study, we decided to use a method called AdaBoost among various Boosting methods. AdaBoost is a technique for creating a classifier with high accuracy by learning by adaptively weighting the recognition rate of the classifier during the learning process [4].

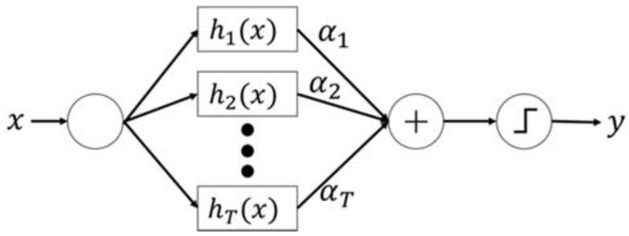


Fig. 8. AdaBoost

3.2. *Extract image features*

The classifier learns by extracting the feature amount of the image at the time of creation. There are the following three feature extraction methods.

- Haar-Like,
- Local Binary Pattern (LBP),
- Histogram of Oriented Gradients (HOG).

In this study, we conducted experiments using LBP.

3.2.1. *LBP*

LBP is one of feature quantities that can be used for image recognition and classification [5]. It is usually calculated in a 3x3 pixel area and extracts local features. It is particularly resistant to lighting changes and has the advantage of being able to calculate at high speed. In the calculation method, first, the luminance value of the center is compared with the luminance values of pixels in the vicinity of the surrounding eight. Of the 8 neighborhoods, 1 is set when the luminance value is equal to or greater than the central luminance value, and 0 is set otherwise. This is multiplied by a mask, the sum is obtained, and this value is replaced with the luminance value of the central pixel. A mask is a weight assigned 2^n clockwise from the top left. This operation is performed for all pixels, and the resulting image is called an LBP image. The object is recognized using the LBP feature value thus obtained.

3.3. *Soccer mode*

The first screen is in the black initial state. When the user raises his knee, a soccer ball and a soccer stadium are projected. In addition, a sound is now played when kicking the ball. If the ball coordinate is lower than the y coordinate threshold, the ball disappears and is initialized. In addition, physics was used for the lifting operation.

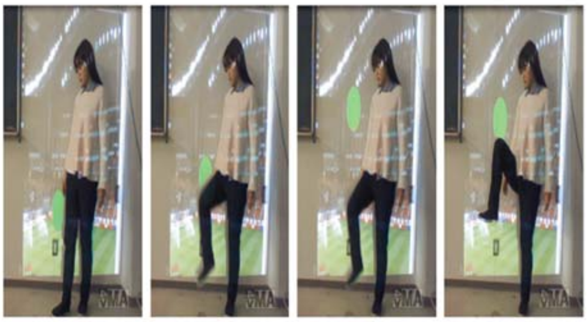


Fig. 9. Soccer mode execution result

3.4. *Baseball mode*

The first screen is in the black initial state. When the user raises his hand, a baseball ball and a baseball stadium are projected. In addition, a sound is now played when throwing the ball. In order to make the ball appear to be thrown in the back, the ball gradually gets smaller as the flight distance increases. When the ball position is greater than the threshold value of the x coordinate (the left edge of the screen as seen by the user), the ball disappears and is initialized.

Fig. 10. Baseball mode execution result

4. Experimental result

The experiment using the threshold of skeleton coordinates and the method of this study were tested and compared [6]. In this study, we performed grayscale on the “Color” screen display obtained from Kinect and detected the user using object detector. Figure 11 shows human body recognition in soccer mode.

Fig. 11. User detection using human detector in soccer mode

In previous study, when the subject disappeared from the Kinect field of view and entered the field of view again, there was a problem that the skeletal information was not retraced and mapping was not successful. In this study, we solved the problem by performing human body recognition using OpenCV without using tracking of skeleton coordinates.

5. Conclusion

In this study, we used an object detector created with Opencv to create an interactive projection mapping whose mapping changes according to human movement**.** In previous study, when the subject disappeared from the Kinect field of view and entered the field of view again, there was a problem that the skeletal information was not retraced and mapping was not successful. In this study, we were able to solve these problems by using a method that does not use arbitrary thresholds. For example, in previous artist concerts, artists have been following the movement of projection mapping. However, with this method, it is possible to perform projection mapping that matches the movement of the artist. This will reduce the difference between each other's movements and make it possible to produce more realistic performances. In the future, we would like to realize interactive projection mapping using more accurate human body classifiers by increasing the number of sample images.

References

1. K.Ogasa, H.Katayose "TPPM (Take Part in Projection Mapping): Multiplayer participation type projection mapping application using tablet terminal" 2014
2. N.Koeda, E.Ueda, Y.Nakamura ″Introduction to image processing with OpenCV″ 2018.
3. Nobuyuki Nishiuchi, Takuma Suzuki, Kimihiro Yamanaka "Development of Work Posture Evaluation System Using Boosting" J Jpn Ind Manage Assoc 62,51-58,2011
4. Yoav Freund, Robert E Schapire "A Decision-Theoretic Generalization of On-Line Learningand an Application to Boosting" journal of computer and system sciences 55, 119139 (1997)
5. Ojala T, Pietikäinen M & Harwood D, “Performance evaluation of texture measures with classification based on Kullback discrimination of distributions.”, Proc. 12th International Conference on Pattern Recognition (ICPR 1994), Jerusalem, Israel. Vol I, 582-585. (1994)
6. Makoto Sakamoto, Takahiro Shinoda, Takahiro Ishizu, Masahiro Hori, Hazuki Watanabe, Amane Takei, Takao Ito "A Proposal of Interactive Projection Mapping Using Kinect" IEEE Xplore: 29 November 2018, DOI: 10.1109/ICT-ROBOT.2018.8549899

Fundamental Study on Control of CG Characters by Electroencephalography (EEG) Analysis

Kenji Sakoma, Makoto Sakamoto*, Takahiro Ishizu, Takahiro Shinoda and Amane Takei

Faculty of Engineering, University of Miyazaki,

Miyazaki-City, Miyazaki, Japan

Takao Ito

Graduate School of Engineering, Hiroshima University,

Higashi-Hiroshima, Hiroshima, Japan

E-mail: hm14002@student.miyazaki-u.ac.jp, fruits2000jp@yahoo.co.jp
**Corresponding Author*

Abstract

Virtual Reality (VR) technology is expected to develop in various fields such as medical, education, business and entertainment. In this study, we aim at more intuitive operation by focusing on troublesome mounting in VR. To achieve this goal, we propose manipulation method of CG character by electroencephalography (EEG) with relatively inexpensive equipment. In this paper, we ask five subjects to manipulate CG characters by EEG and evaluate this system by questionnaire.

Keywords: Virtual Reality (VR), manipulation method, CG character, electroencephalography (EEG)

1. Introduction

Since 2016, the so-called the VR (Virtual Reality) first year, VR technology has made remarkable progress around the world. Every year, higher-level VR devices are released, and although the price is somewhat high, it has become easy to purchase for us. As a result, VR technology is expected to develop in various fields such as medical care, education, business, and entertainment.

However, there are five major problems that must be solved in the development of VR technology. The first one is the problem of troublesomeness of wearing, the second the problem of specifications, the third the social and market problem, the fourth the problem at the production site, and the last the health problem. Among these, I focused on the first problem of troublesomeness of wearing. When using VR devices, it is necessary to set up many cables and sensors. Also, there must be a certain space around the user when using it. This is for the consideration of the surroundings or the safety of the user extent when performing any operation in the VR world. For this reason, there arises a problem that it is necessary to select a place when using a VR device. Recently, integrated (stand-alone) models that do not require any cables are gradually appearing. The integral type solves the problem of cables. However, space problems around the user have not been solved[1)].

Therefore, as a method for solving this problem, it is proposed to use an operation by electroencephalogram, that is, a method called BMI (Brain Machine Interface). BMI, also called BCI (Brain Computer Interface: Brain Computer Interface), is a study that measures brain waves with an instrument, extracts features from the brain waves, and causes the computer to react in some way according to the features. We think that space problems around the user may be solved by using this method[2,3].

2. Research Method

In this study, it is necessary to measure brain waves. Therefore, the above equipment was used in this research. Generally, the BMI research uses an international 10-20 electroencephalograph, but in this research, the equipment sold by Neurosky Inc. from the viewpoint of whether BMI can be realized inexpensively (see Fig. 1.) was used.

2.1. *Used equipment*

- Mind Wave Mobile

Fig. 1. Mind Wave Mobile.

2.2. *Development environment*

The development environment for constructing this system is as shown in Table 1 below.

OS	Windows 10
Programming language	C#
Measuring equipment	Mind Wave Mobile
Software	Visual Studio 2017 Unity 2017.3.1f1

Table 1. Development environment.

2.3. *Library*

- libStreamSDK
  Used to receive brainwave data from Mind Wave Mobile in Visual Studio 2017.
- MathNet.Numerics
- Accord
- Accord.MachinerLearning
- Accord.Math
- Accord.Statistics
  Used to process data received from Mind Wave Mobile in Visual Studio 2017.

2.4. *Brainwave measurement*

In this study, it is necessary to measure brain waves. In the general BMI research, the international 10-20 method is used, but in this research the brain wave measuring machine Mind Wave Mobile manufactured by Neurosky Inc., which is relatively inexpensive. because it is also that it is from the viewpoint of verifying whether BMI is possible even with a simple device. When receiving EEG data in Visual Studio 2017, I used a library called libStreamSDK.

In addition, when acquiring an electroencephalogram, the kind of electroencephalogram acquires a Row wave which is not processed at all. Then, extract features in Visual Studio 2017.

2.5. *EEG processing method*

2.5.1. *FFT (Fast Fourier Transform)*

It is a kind of DFT (Discrete Fourier Transform), and is an algorithm designed to execute DFT fast on a computer (Equation 1). When using this analysis method, I used the library of MathNet.Numerics[5].

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i\frac{2\pi}{N}} \qquad (1)$$

2.5.2. *Continuous wavelet transform*

It is possible to use a wavelet function as a basis function and leave temporal information that is lost in the Fourier transform. This is a method that attempts to represent the waveform of a given input by scaling, translating, and adding wavelet functions (see Equation 2).

$$WT(a,b) = \frac{1}{\sqrt{a}} \int f(t)\psi\left(\frac{t-b}{a}\right) \mathrm{d}t \qquad (2)$$

2.5.3. *Discreate wavelet transform*

The original signal is decomposed into the high frequency component and the low frequency component, and the decomposed low frequency component is also decomposed into the high frequency component and the low frequency component.

$$d_k^{(j)} = 2^j \int_R f_{(t)} \overline{\psi(2^j t - k)} \, \mathrm{d}t \qquad (3)$$

2.5.4. *EEG feature discrimination method*

As EEG feature discrimination method, SVM (Support Vector Machine: support vector machine) generally used in EEG feature discrimination was used. SVM is one of pattern recognition models using supervised learning (Equations 4 and 5). SVM has high discrimination performance against unlearned data. In this study, Gaussian kernel was used as the kernel function. In addition, the library of MathNet.Numerics Accord, Accord. MachinerLearnin, Accord.Math, and Accord.Statistics was used to execute SVM [4].

$$y_{(x)} = \mathrm{sgn}\left\{\sum_{n=SV} w_n K(x_n, x) + b\right\} \quad (4)$$

$$= \begin{cases} 0 & x \in classA \\ 1 & x \in classB \end{cases} \quad (5)$$

$$K(x_1, x_2) = \exp(-\|x_1 - x_2\|^2 / 2\sigma^2) \quad (6)$$

2.5.5. *Visual Studio 2017*

Visual Studio 2017 processes brain wave data sent from Mind Wave Mobile and performs feature detection. At this time, the brain wave processing method used FFT. Thereafter, the processed electroencephalogram data is subjected to SVM, and the electroencephalogram is discriminated such that 0 is output when relaxing, and 1 is output when user tries to move (equation 5).

2.5.6. *Unity 2017.3.1f1*

We created CG characters and programs that change in response to the features of the brain waves (Fig. 2). It was processed in Visual Studio 2017, and it received the feature which was judged by SVM, and made it to respond according to the feature. In addition, this character uses the one promulgated free from Unity's Asset Store.

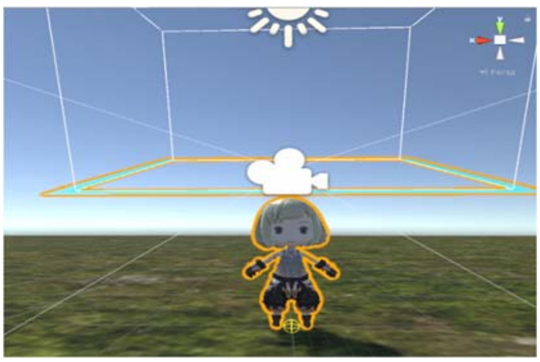


Fig. 1. The CG character which used in Unity.

3. Experimental method

The subject sits in a chair and starts measuring after wearing Mind Wave Mobile on the head. The brain wave processing method uses FFT. During the experiment, subjects are asked to relax moderately or move their body, and so on, to evaluate how CG characters in Unity move. In this research, CG characters move only in two types, and they are stopped when relaxing and move forward when trying to move their bodies.

4. Evaluation experiment

In this experiment, five university students became subjects, and they conducted evaluation experiments and asked them to answer questionnaires. The contents of the evaluation were in the form of free-form descriptions of "good points", "bad points" and "others".

5. Evaluation results

As a result of conducting the questionnaire, it became as follows.

"Good things" were "Simply fun", "I was interested because I wanted to move the body, so I was interesting", "CG character was cute" and so on.

"The bad point" were "It took time for the CG character to actually reflect the reaction after trying to relax / move the body," "Mind Wave Mobile is coming off," "Mind Wave Mobile was tightening my head and hurts me", "It moves even though it is relaxing. / The character freezes while trying to move the body." and so on.

"Other" were that "is it not compatible with other movements?" and "is there no other CG character?", etc.

6. Consideration

From the evaluation results, it was found from the “good points” that the control of CG characters by EEG which measured by inexpensively equipment, which is the purpose of this research, can achieve certain results with regard to only two simple patterns of forward and backward.

However, from the "bad point", it was found that the method of detecting the characteristic of the electroencephalogram and the discrimination method were still incomplete. We also found that there was a problem with the installation of Mind Wave Mobile. This is considered to be a problem that the size cannot be changed or cannot be selected. From the above "bad points", it was found that this research needs further improvement.

According to the opinion of "other", it was found that the operability and key CG characters in this research require diversity.

7. Conclusion

In this research, we focused on the "trouble of wearing" which is a problem in the development of VR technology. Above all, we focused on the problem of spatiality around the user, and researched that it could be solved by the following method. We have proposed a method in a field generally called BMI the electroencephalogram data is measured using an electroencephalograph, the electroencephalogram data is received by Visual Studio 2017, the electroencephalogram data is appropriately processed, and then the electroencephalogram features are discriminated by SVM, and the discriminant values sent to Unity 2017.3.1f1 and operates CG characters in Unity according to the determined value. In addition, although an electroencephalogram measurement machine conventionally uses the international 10-20 method, in this research, Mind Wave Mobile manufactured by NeuroSky was used because it verifies that it could be substituted by an inexpensive device. The result of the evaluation experiment shows that although the manipulation of CG character by electroencephalogram which is the purpose of this research can achieve a certain result, its accuracy is bad and there is much room for improvement. Future issues will include brain wave feature detection, improvement of brain wave feature discrimination accuracy, improvement of troublesomeness when wearing Mind Wave Mobile, and diversity of motion of CG characters.

8. References

1. 2017 Advanced Content Technologies, etc. Distribution Promotion Business Report, Part I, guidelines for using content production technology such as VR, 2018 [Online]. https://www.vipo.or.jp/u/I-1_SenshinContents_Guideline.pdf
2. Tomoki Hidaka, Yuki Shiraishi, Yukiko Kawai, Jiro Okuda , "Information control system by electroencephalogram using a portable terminal" , Information Processing Society of Japan, 2014.
3. Uchida Yu, Kikuro Fujimura, Yoshihiro Maegaki, "A Study on Numerical Processing of EEG Data for EEG Analysis" , Japan Society for Fuzzy Theory and Intelligent Informatics, 2014.
4. Takio Kurita , "Introduction to Support Vector Machine"[Online]. https://home.hiroshima-u.ac.jp/tkurita/lecture/svm.pdf
5. Jun Sakamoto, Yasuhiro Mori, Takayoshi Sekioka, "Probability Analysis Method By Discreate Fast Fourier Transform", J. Struct. Constr, Eng . , AIJ , No , 472, 39−45, Jun, , 1995.

Development of parallel microwave analysis code: ADVENTURE_FullWave

Amane Takei
Department of Electrical and System Engineering, University of Miyazaki, 1-1 Gakuen Kibanadai-Nishi Miyazaki, 889-2192, Japan
E-mail: takei@cc.miyazaki-u.ac.jp
www.miyazaki-u.ac.jp/english/

Abstract

With the expansion of electromagnetic field analysis using computers, large spaces that include complex shapes have also become an analysis target, and the development of a high-accuracy analysis is required for these problems. Therefore, in the present study, Berenger's PML, which is currently the most effective absorbing boundary condition, is applied to the parallel finite element method based on the domain decomposition method, which is an effective analysis method for the microwave band. As a basic study, we developed an analysis code: ADVENTURE_FullWave using a parallel finite element method based on the iterative domain decomposition method. In verifying the accuracy of the analysis code, we analyzed TEAM Workshop Problem 29, which is a benchmark problem, and confirmed that a highly accurate solution is obtained. Next, a model with Berenger's PML added to the dipole antenna model is used as an analysis object, and the absorption performance of the PML is evaluated using a reflection coefficient based on the S parameter. Moreover, the accuracy of the antenna analysis is evaluated by comparing the directivity of the dipole antenna with the theoretical solution.

Keywords: Iterative domain decomposition method, Parallel finite element method, Berenger's PML, Large-scale analysis, Microwave analysis.

1. Introduction

In the present study, Berenger's PML for the full-wave electromagnetic field analysis, which is currently regarded as the most effective absorbing boundary condition, is applied to the parallel finite element method, which is an analysis method that is capable of computing large-scale problems. We then show that the parallel finite element method is effective as an analysis method for a microwave band by calculating the problem dealing with an open domain and evaluating the solution's accuracy. First, as a basic study, we developed a three-dimensional electromagnetic field analysis solver using a parallel finite element method based on the iterative domain decomposition method. In order to verify the accuracy of the developed analytical solver, we calculate TEAM Workshop Problem 29, which is a benchmark problem, and evaluate the accuracy of the analytical solver. Next, a model with a PML added to the dipole antenna is used as the analysis object, and the absorption performance of the PML is evaluated using a reflection coefficient based on the *S* parameter. Moreover, by comparing the directivity of the dipole antenna with the theoretical solution, evaluations of the accuracy and performance of the antenna analysis are performed. As a result, we confirmed the effectiveness of this method for microwave analysis.

2. Numerical analysis

We analyze TEAM Workshop Problem 29[1] in order to verify the accuracy of the developed analytical solver.[2] TEAM 29 is a benchmark problem and involves a resonator model. The resonator is cylindrical and has a diameter of 1.9 [m] and a height of 1.45 [m]. In the analysis, a dielectric phantom having a relative permittivity ε_r of 80 and an electric conductivity σ of 0.52 [S/m] is positioned, and the resonance state is investigated. The analysis domain boundary is a perfect conductor. The analysis model is shown in Fig. 1. Table 1 shows the specifications of the TEAM 29 model. The highest calculation efficiency is achieved when the number of elements contained in one subdomain is approximately 170, and the number of partial domains is determined such that the number of elements contained in one subdomain is equal to 170.

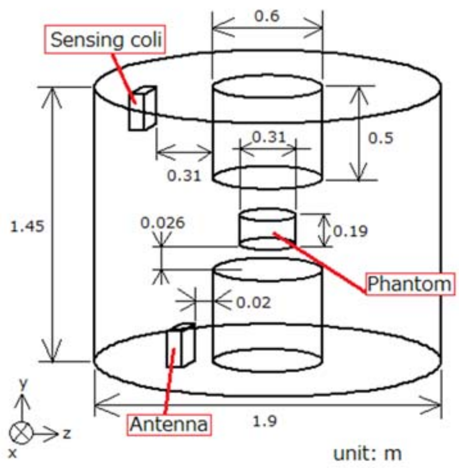


Fig. 1. TEAM 29 cavity resonator model

Table 1. TEAM 29 model data.

No. of Elements	DOFs	No. of Subdomains
121,277	149,668	100×7

The results of the analysis of TEAM 29 are subjected to frequency response analysis in order to confirm the accuracy of the developed analytical solver. In order to detect the resonant frequency and compare solutions with actual measurements, the frequencies of some range are analyzed. The frequency band of 60 [MHz] to 140 [MHz] is calculated for 2-MHz steps, and the response for every frequency step is investigated. In addition, calculations near the resonance frequency are performed in 0.4-MHz intervals. The computing environment used in the present study is a 25-PC cluster equipped with Intel Core i7-2600K multi-core CPUs (total: 100 cores) and 32 GB memory is used. The compiler used is the gcc. In addition, Message Passing Interface (MPI) is used for the parallelization library. The average calculation time per frequency step and the averagely used memory are shown in Table 2. Fig. 2 shows the frequency response of the magnetic field. The measured and calculated values are shown in Table 3.

Table 2. Specifications of the TEAM 29 model.

No. of Elements	DOFs	No. of Subdomains
121,277	149,668	100×7

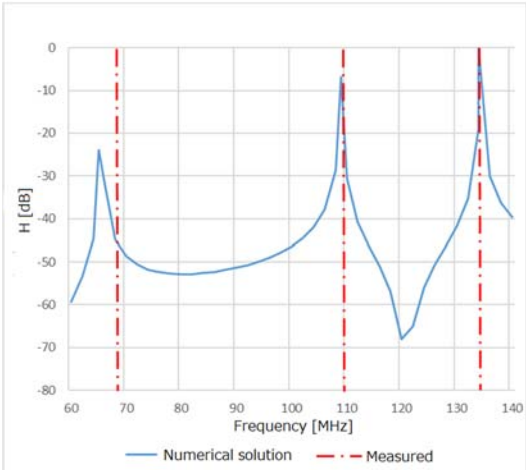


Fig. 2. Numerical and measured frequency response analysis results

Table 3. Resonance frequencies (Units: [MHz]. The error rate [%] between measured data and the numerical solution is shown in parentheses.)

Resonance mode	Measured data	FDTD 25-mm mesh	Result
1st	68.6	67 (2.33)	65.6 (4.37)
2nd	110	110	109.0 (0.91)
3rd	134	134	134.4 (0.30)

As shown in Fig. 2, a solution resonating around the resonance frequency of the actually measured value is obtained. In the comparison of the measured and calculated values, the error rate is 4.37 [%] in the 1st mode, 0.91 [%] in the 2nd mode, and 0.30 [%] in the 3rd mode. As the mode increases, the error rate decreases. However, it is the same tendency as the analysis result obtained by the FDTD method. Moreover, the error rate is less than 5 [%], and a solution with high accuracy is obtained.

Therefore, the solution obtained by the developed solver is proven to have a sufficiently high accuracy. Moreover, in the analysis of the dipole antenna applying the PML described in the following sections, the error tolerance index is defined as 5 [%] in order to evaluate the accuracy.

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. 3 shows a schematic diagram of Berenger's PML absorbing boundary.

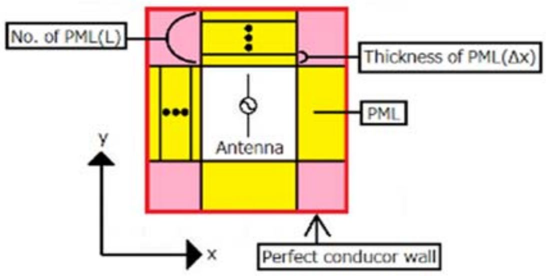


Fig. 3. PML absorbing boundary

In this paper the distribution of the electric conductivity for PML is expressed as follows:

$$\sigma = \sigma_{max}\left[\frac{\left(L-\hat{L}(x)\right)\Delta x}{L\Delta x}\right]^M \tag{1}$$

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:

$$|R(\phi)| \cong exp\left[-\frac{2\sigma_{max}L\Delta x}{(M+1)\varepsilon_0 c}\cos\phi\right] \tag{2}$$

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:

$$\sigma_{max} = -\frac{(M+1)\varepsilon_0 c}{2L\Delta x}\ln|R(0)| \tag{3}$$

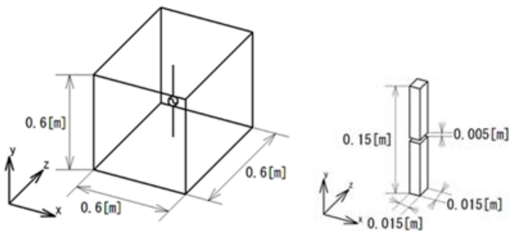
In the present study, we construct a PML using (1) through (3) with L = 9, M = 4, and Δx = λ/10. However, in order to reduce the analysis scale, we examine the optimum value of L in the next section.

3.2. *Numerical results*

We assign the PML to the dipole antenna model. The analysis domain is a cube of length 0.6 [m] so that the distance from the antenna to the innermost PML matches the wavelength. The current density is applied to the antenna as a current source as follows:

$$I(y) = I_0 \cos\left(\frac{2\pi}{\lambda} y\right) \quad : -l \leq y \leq l \tag{4}$$

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. 4. Dipole antenna model

We assign PMLs to the domain boundary as shown in Fig. 4(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 4 lists the number of elements and the degree of freedom of the analysis model.

Table 4. Number of elements and DOFs of the dipole antenna model

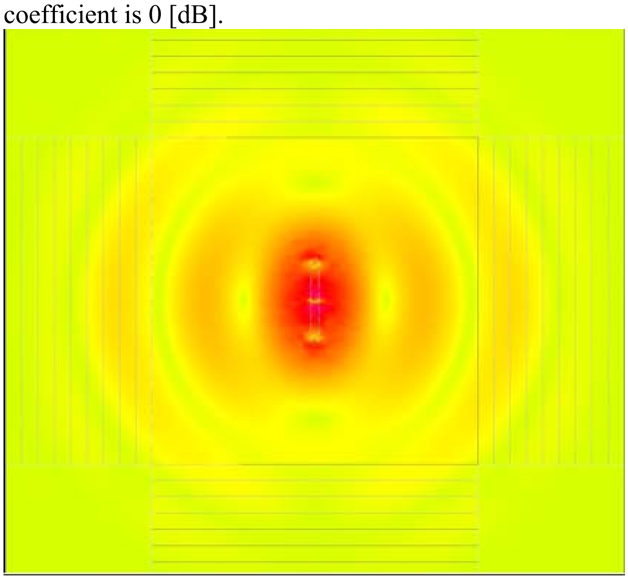
	PML(0): Perfect conductor wall	PML(9)
No. of Elements	4,669,759	26,899,669
DOFs	5,506,368	31,703,550

In (3), 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 (1). 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 is the same as in the section 2. Table 5 lists the reflection coefficient, the CPU time, and the memory size.

Table 5. 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].



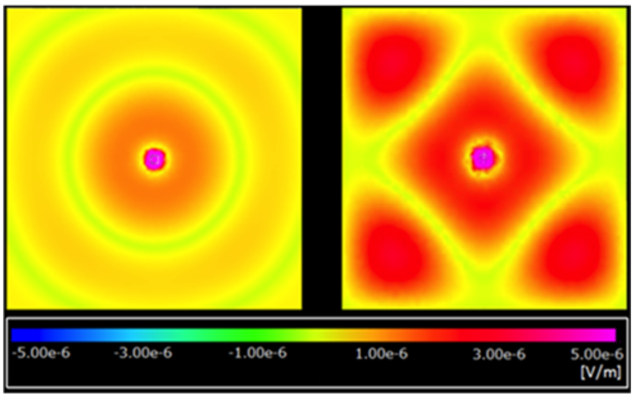


Fig. 5. **V**isualization of the analysis result (electric field) (Upper: Side view of with PML of PML(9), Lower-Left: Upper view of PML(9), L-Right: Upper view of PML(0))

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.

In Fig. 5, 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. 5 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:

$$E_\theta = j60I\frac{e^{-jkr}}{r}\cdot\frac{\cos\left(\frac{\pi}{2}\cos\theta\right)}{\sin\theta} \qquad (5)$$

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. 6 shows a plot of the numerical analysis solution e_θ and the theoretical solution E_θ in increments of 1 [deg].

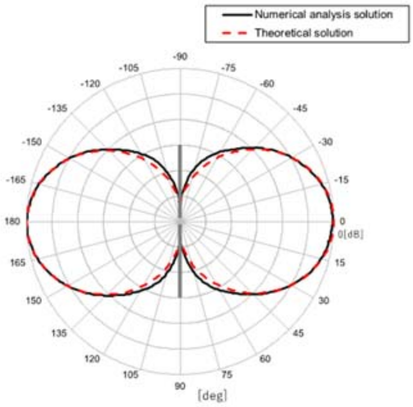


Fig.6. Numerical and theoretical solutions in the E plane

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 (6). 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 (7). As a result, the average error rate is 1.70 [%], and it is shown that a highly accurate solution can be obtained.

$$\arcsin\left(\frac{2\,l^2}{r\lambda}\right) + 90 \le \theta \le 90 - \arcsin\left(\frac{2\,l^2}{r\lambda}\right) \qquad (6)$$

$$E_{err} = \frac{\sum_{i=\theta_{Min}}^{\theta_{Max}} \frac{|e_i - E_i|}{E_i}}{\theta_{Max} - \theta_{Min} + 1} \times 100 \quad [\%] \qquad (7)$$

In the calculations shown in Fig.5, 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. The computing environment is the same as in the section 2. Table 6 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 6. 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 6, 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 for the full-wave electromagnetic field analysis and gives the average value of the electric conductivity of each layer at the corner of the model. Performance evaluation reveal 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: http://www.fujikura.co.jp/eng/

Behavioral Modeling Technique for Multiple Objects of Software Using Extended Place/Transition Nets with Attributed Tokens

Tomohiko Takagi
Faculty of Engineering and Design, Kagawa University
2217-20 Hayashi-cho, Takamatsu-shi, Kagawa 761-0396, Japan

Ryo Kurozumi
Graduate School of Engineering, Kagawa University
2217-20 Hayashi-cho, Takamatsu-shi, Kagawa 761-0396, Japan
E-mail: takagi@eng.kagawa-u.ac.jp, s19g457@stu.kagawa-u.ac.jp

Abstract

This paper shows EPNAT (Extended Place/transition Net with Attributed Tokens) and a modeling technique using it in order to address the problem of EPN. In software modeling using EPN, objects of which the software consists need to be defined individually as sub-EPN models, even if they have the same behavior. However, in software modeling using EPNAT, objects that have the same behavior can be integrated into one sub-EPNAT model, and therefore EPNAT models would be smaller than EPN models.

Keywords: software modeling, behavioral model, place/transition net, VDM

1. Introduction

EPN (Extended Place/transition Net)[1,2] is a formal modeling language to represent the behavior of software that consists of multiple objects. The objects mean modules, components, or subsystems in this study, and they interact each other to provide expected functionality of the software. EPN enables engineers to construct unambiguous and executable software specifications that can be used for systematic skeleton code and test case generation. However, in software modeling using EPN, objects need to be defined individually even if they have the same behavior, which causes an increase in model size. A large model that includes redundant definitions often leads to additional cost for design, implementation, test, and maintenance, due to its poor readability.

This paper shows a novel language called EPNAT (EPN with Attributed Tokens) and a modeling technique using it in order to address this problem. In EPNAT, objects are expressed as attributed tokens that are classified into types, and also the states and events of objects of the same type are expressed as places and transitions, respectively. Attributed tokens can pass through places and transitions that belong to the same types as theirs. EPNAT models are converted to VDM++ specifications, and allow engineers to refine software specifications, create programs, test cases, and so on.

The rest of this paper is organized as follows. Section 2 shows software modeling using EPN and its problem. In section 3, we propose EPNAT and a modeling technique using it. In section 4, we discuss its effectiveness and future work.

2. Related Work

2.1. *Software Modeling Using EPN*

Petri net including PN has been used in traditional software modeling and testing.[3,4] In previous study, we extended the PN by introducing VDM++ (a formal modeling language in VDM)[5] in order to enhance its representation power, and generate test cases systematically. The extended PN, that is, EPN[1,2] consists of the following four kinds of structural elements.

- Places to express states of each object
- Transitions to express events of each object
- Tokens to express a current state of each object
- Arcs to specify the flow of tokens between places and transitions

Details that cannot be represented in PN, such as the actions and pre-conditions (guards) of transitions, are formally written in VDM++. When a transition is fired, the values of variables that are defined for each object can be changed by its actions.

An EPN model consists of multiple sub-EPN models that represent the behavior of objects, and the sub-EPN models are connected by glue transitions that play an important role to specify the interaction among the objects. In an EPN model, the execution traces of software, that is, test cases are described as the sequences of successive markings (distribution of tokens on places), transitions, values of variables, and so on. An EPN model that gives engineers the overview of the behavior of software can be converted to a VDM++ specification that is used to refine the specification of the software.

2.2. *Problem*

In software modeling using EPN, objects need to be defined individually as sub-EPN models even if they have the same behavior, which causes an increase in model size. In general, a large model that includes redundant definitions often leads to additional cost for design, implementation, test, and maintenance, due to its poor readability.

We introduce an example given in Fig. 1 in order to discuss this problem. This simple load balancer (hereinafter, referred to as SLB) contains multiple servers and requests as objects. All server objects have the same behavior, and also all request objects have the same behavior. However, when the behavior of SLB is defined as an EPN model, each object is defined as a sub-EPN model that represents the behavior of each object. That is, if s and r are the maximum numbers of servers and requests that can be handled at the same time by SLB, an EPN model of SLB will contain s and r sub-EPN models for server and request objects, respectively.

This simple load balancer (hereinafter, referred to as SLB) can handle multiple servers and requests. Each server and request are uniquely identified in SLB. Also, each server and request have specific capacity and data size, respectively. It is assumed that the capacity is larger than the data size. A server uses the capacity of data size just during processing a request.

A new server can be added to SLB at any time, and it can start working after maintenance. If there are requests on a queue that is described below, SLB allocates a server to the request that waits on the head of the queue. The server to be allocated must be in a state of working, and must have enough remaining capacity to process the request. If SLB finds multiple servers that can be allocated, it selects an arbitrary one. SLB can allocate a server to multiple requests if the server has enough remaining capacity. After a server has finished processing a request, SLB deallocates the server. A server that is not allocated can stop working and go to a state of maintenance. A server in a state of maintenance can be removed from SLB at any time.

SLB can receive new requests at any time. When SLB receives a request, the request is added to the tail of the queue. If SLB finds a server for the request on the head of the queue, the request is removed from the queue.

Fig. 1. Software requirements of a simple load balancer.

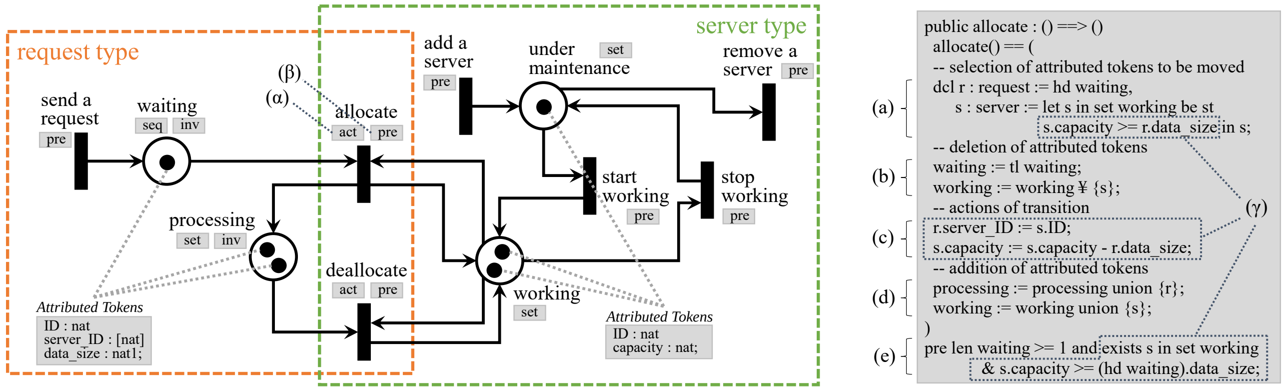
The objects that have the same behavior cannot be easily integrated into one sub-EPN model, since the variables for each object are defined as attributes of an EPN model or each sub-EPN model.

3. EPNAT and a Modeling Technique

In this section, we propose EPNAT and a modeling technique using it in order to address the above-mentioned problem.

3.1. *EPNAT*

EPNAT is the extension of EPN, and it includes all the structural elements of EPN. The most obvious difference is that tokens in EPNAT correspond to objects that are classified into types, and they have variables for the objects. In this paper, tokens that have variables for objects are called attributed tokens. For example, in SLB of Fig. 1, the server objects (that is, objects of server type) need to have variables such as ID and capacity, and thus attributed tokens of server type (that is, attributed tokens that correspond to the server objects) have those variables. Attributed tokens of the same type have the same variables but different values. For example, the attributed tokens at_a and at_b that correspond to the server objects a



(i) Overview of an EPNAT model

(ii) Details of the glue transition "allocate"

Fig. 2. EPNAT model of a simple load balancer.

and *b* in SLB have different values, such as at_a.ID=1 and at_b.ID=2, respectively.

Places and transitions that express the states and events of objects respectively are also classified into types, and they are shared by attributed tokens of the same type in EPNAT. Attributed tokens can pass through places and transitions that belong to the same types as theirs. For example, the attributed tokens of sever type can pass through the places and transitions of server type. Note that glue transitions belong to multiple types. Transitions can have actions for attributed tokens. When a transition is fired, the variables of attributed tokens that pass through it can be changed by its actions. Transitions can also have pre-conditions and post-conditions that need to be satisfied just before and after the fire of the transitions, respectively. Places can have invariants, that is, a condition that needs to be always satisfied. The variables of attributed tokens can be referred in pre-conditions, post-conditions, and invariants.

All structural elements of each sub-EPNAT model belong to the same type, and each sub-EPNAT model represents the behavior of all the objects of the type. For example, one sub-EPNAT model of server type covers all the objects (an arbitrary number of the objects) of server type. All sub-EPNAT models are connected by glue transitions in an EPNAT model that represents the behavior of software.

3.2. *Example*

Fig. 2 (i) shows the overview of an EPNAT model of SLB that is constructed based on Fig. 1. It consists of two sub-EPNAT models that represent the behavior of objects of server type and request type. There are two places and six transitions in server type, and two places and three transitions in request type. The transitions "allocate" and "deallocate" are glue transitions, and the two sub-EPNAT models are connected by them.

The labels "seq" and "set" indicate the collection type of places, that is, whether the order of arrival of attributed tokens on places should be kept or not. Either "seq" or "set" should be specified for each place by engineers. In VDM++ specifications that is discussed later, places with the labels "seq" and "set" are implemented as instance variables of sequence type and set type, respectively. The label "inv" means that a place has an invariant. Details of "inv" are written in VDM++ by engineers, but they are omitted in the overview of an EPNAT model for ease of readability.

The label "act" means that actions are given to a transition. In VDM++ specifications, transitions including their actions, incoming and outgoing arcs are implemented as operations. The labels "pre" and "post" mean that a transition has a pre-condition and post-condition respectively, but the latter does not appear in Fig. 2 (i). Details of the labels "act", "pre" and "post" are written in VDM++ by engineers, but they are omitted in the overview of an EPNAT model for ease of readability.

Attributed tokens of server type and request type have two and three variables, respectively. The term "nat" means natural number type in VDM++. In Fig. 2 (i), there are six attributed tokens. Note that it is a snapshot and there should be no attributed tokens in the initial state of

SLB. In VDM++ specifications, attributed tokens are implemented as record type.

Fig. 2 (ii) shows the details of the glue transition "allocate" written in VDM++. The statement (a) specifies the way of selection of attributed tokens to be moved. The statements (b) and (d) specify the way of deletion and addition of attributed tokens on the related places, respectively. The statements (c) specify the actions to be executed on the transition "allocate", and they correspond to the label "act" that is pointed by (α) in Fig. 2 (i). The statement (e) specifies the pre-condition that needs to be satisfied just before the fire of the transition "allocate", and it corresponds to the label "pre" that is pointed by (β) in Fig. 2 (i). The codes that are pointed by (γ) in Fig. 2 (ii) need to be considered and written based on software requirements by engineers.

In this paper, the details of the other structural elements in Fig. 2 (i) are omitted because of limitations of space. Many VDM++ codes for them can be semi-automatically generated by our coding pattern that is discussed in section 3.3.

3.3. *Modeling Technique*

The modeling technique using EPNAT consists of the following four steps. Note that these steps are not always separated clearly. If engineers find any problems in a current step, they can return to a previous step. Also, engineers can incrementally and iteratively proceed with these steps. The product that has been constructed in each step is executable, and thus should be tested in the end of each step.

Step 1: Construction of Sub-PN Models

Objects are identified and classified into types, and then the abstracted behavior of objects of the same type is defined as sub-PN models, as shown in Fig. 3.

Step 2: Integration of Sub-PN Models

All the sub-PN models are connected by glue transitions in order to complete a PN model that represents the abstracted behavior of software.

Step 3: Addition of Details to a PN Model

The following definitions, which are written in VDM++, are added to the PN model in order to complete an EPNAT model.

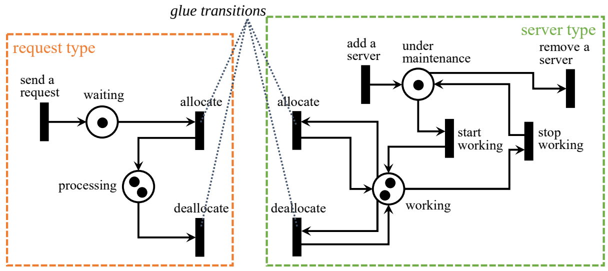


Fig. 3. Construction of sub-PN models.

- Variables of attributed tokens
- Collection type of places ("seq" or "set")
- Invariants for places
- Actions to be executed on transitions
- Pre-conditions and post-conditions for transitions
- Conditions to select attributed tokens to be moved on transitions

Step 4: Conversion to a VDM++ Specification

The EPNAT model is converted to a VDM++ specification. Fig. 4 shows our coding pattern to perform this conversion systematically. The codes shown in the form of <*terms*> are extracted from the EPNAT model, or are written by engineers.

The VDM++ specification is used to refine software specifications, create programs, test cases, and so on.

4. Discussion and Future Work

In the previous section, we proposed EPNAT and a modeling technique using it in order to address the problem of EPN. In software modeling using EPN, objects need to be defined individually as sub-EPN models even if they are the same type, which causes an increase in model size. However, in software modeling using EPNAT, the objects of the same type can be integrated into one sub-EPNAT model, and therefore EPNAT models would be smaller than EPN models. We discussed the problem and effectiveness by using an example of SLB.

The coding pattern proposed in this paper will not be suitable for extremely large and complex software requirements, since they are defined as one class in a VDM++ specification. For example, each sub-EPNAT model may be defined as a class in order to address this problem, which will be discussed in our future study.

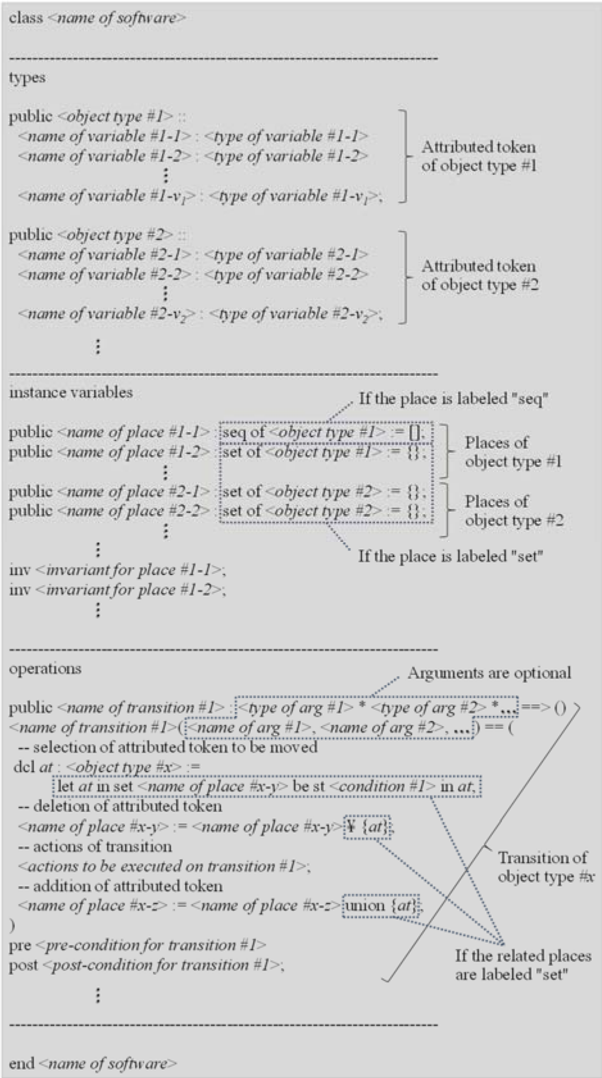


Fig. 4. Coding pattern to convert from an EPNAT model to a VDM++ specification.

Our technique should be supported by a tool, since engineers will need to spend a certain amount of time and effort to construct EPNAT models and VDM++ specifications. In our previous study,[1] a prototype tool was developed to edit EPN models and VDM++ specifications, and automatically convert between them. The new functions to edit EPNAT models and convert them to VDM++ specifications are being added to the prototype tool. We will apply it to non-trivial software requirements to evaluate the effectiveness of our technique further.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number JP17K00103.

References

1. T. Takagi and R. Kurozumi, Prototype of a Modeling Tool to Convert between Extended Place/Transition Nets and VDM++ Specifications, *Proceedings of International Conference on Artificial Life and Robotics*, Jan. 2019, pp.157-160.
2. T. Takagi, R. Kurozumi and T. Katayama, State Transition Tuple Coverage Criterion for Extended Place/Transition Net-Based Testing, *Proceedings of Pacific Rim International Symposium on Dependable Computing*, Dec. 2019, pp.29-30.
3. N.G. Leveson and J.L. Stolzy, Safety Analysis Using Petri Nets, *IEEE Transactions on Software Engineering*, Vol.13, No.3, Mar. 1987, pp.386-397.
4. H. Li, X. Ye, C. Wu, L. Liu and L. Wang, Modeling Interactive Property of MIPv6 with Petri Net for Interoperability Testing, *Proceedings of International Conference on Information and Computing Science*, 2009, pp.313-316.
5. J. Fitzgerald, P.G. Larsen, P. Mukherjee, N. Plat and M. Verhoef, *Validated Designs for Object-Oriented Systems*, Springer-Verlag London, 2005.

Learning Support Technique of Software Visual Modeling Using Place/Transition Nets

Yuki Ue
Graduate School of Engineering, Kagawa University
2217-20 Hayashi-cho, Takamatsu-shi, Kagawa 761-0396, Japan

Tomohiko Takagi
Faculty of Engineering and Design, Kagawa University
2217-20 Hayashi-cho, Takamatsu-shi, Kagawa 761-0396, Japan
E-mail: s18g459@stu.kagawa-u.ac.jp, takagi@eng.kagawa-u.ac.jp

Abstract

This paper shows a learning support technique of software visual modeling. The ideas of block and animated graphics are introduced into this technique in order that learners can intuitively understand the notation and behavior of models, and can be given guidance about the way to construct proper models. Place/transition net is selected as a modeling language in this paper, but other formal behavioral modeling languages based on directed graphs also can be introduced into this technique.

Keywords: software modeling, visual modeling, learning support, place/transition net

1. Introduction

Software modeling is important especially for the development of large and complex software. For example, UML (Unified Modeling Language), state machines, and Petri nets are well known as modeling languages to construct software models, and they have been widely used in the actual development processes of requirements definition, design, and testing. However, the quality of software models depends on the skill of engineers, and learning it costs time and effort.[1]

We propose a learning support technique of software visual modeling to address this problem. The ideas of block and animated graphics, which are well-known in the field of visual programming,[2] are introduced into this technique in order that learners can understand the notation and behavior of models intuitively, and can be given guidance about the way to construct proper models. PN (Place/transition Net) that is a kind of Petri net[3] is selected as a modeling language in this paper. However, other formal behavioral modeling languages based on directed graphs also can be introduced into this technique, since the formal behavioral modeling languages based on directed graphs including PN have the following common characteristics.

- The action of software is associated with nodes and/or arcs.
- The execution order of the action is determined by the arcs.

In this paper, we show the basic concept of visual modeling in section 2, and propose the learning support technique of the visual modeling using PN in section 3. Last section gives conclusion and future work.

2. Basic Concept of Visual Modeling

The visual modeling in this paper is developed from the existing visual programming languages of block type, such as Scratch and Blockly. In the visual programming, programs can be easily created by connecting visual objects that are called blocks. The block describes an instruction to perform specific action, such as to move a cartoon character, to make a sound, and so on. Thus the visual programming is one of effective learning methods for beginners in programing.[2]

Our visual modeling is aimed at supporting the intuitive understanding of the notation (syntax) and behavior (semantics) of PN models. PN is one of formal behavioral modeling languages that are based on directed graphs. In general, formal modeling languages are useful to construct unambiguous and executable definitions, but it will not be so easy for beginners to learn them. Therefore, we introduce the following two key ideas of the visual programming into our visual modeling.

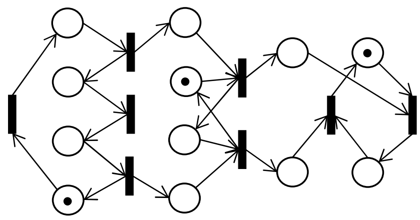
- The structural elements in a program are defined as blocks, and the blocks can be connected with each other only if they are not syntactically wrong. Therefore, it is expected that beginners who are not familiar with syntax can intuitively make their programs.
- The behavior of a program is visualized as animated graphics, and thus it is expected that beginners who are not familiar with semantics can intuitively understand the behavior of their programs.

In our visual modeling, the structural elements of a PN model (that is, places, transitions, tokens, and arcs) are defined as blocks in which the notation rules of PN models are implemented. For example, each place can hold one or more tokens, and it must be connected with one or more transitions by arcs, and so on. As in the case of the visual programming, it is expected that beginners who are not familiar with the notation rules of PN models can intuitively make their PN models.

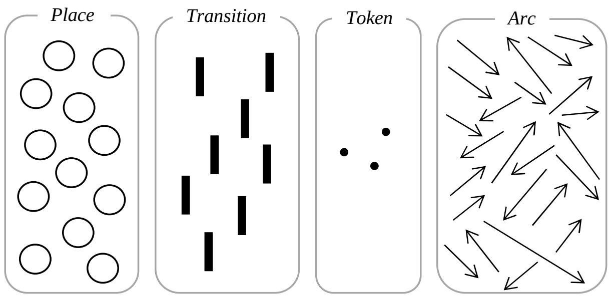
Animated graphics in our visual modeling are used to visualize the behavior of a PN model under construction. If the PN model is wrong (that is, it does not satisfy software requirements), its animated graphics do not move correctly. For example, the animated graphics of a wrong PN model of a crossing gate control system will show that a crossing gate does not work when a train has reached the grade crossing. As in the case of the visual programming, it is expected that beginners who are not familiar with the semantics of PN models can intuitively understand the behavior of their PN models.

3. Learning Support Technique

In this section, we propose the learning support technique of our visual modeling. This technique consists of (1) creation of exercises, (2) work on exercises, (3) review, (4) answer checking, and (5) demonstration. It is assumed that PN is used as a modeling language, and expected learners are beginners in software modeling using PN. This technique should be implemented as a Web application in order to enable instructors and learners to interact with each other.



(a) PN model



(b) A set of blocks that are classified into places, transitions, tokens, and arcs

Fig. 1. Example of a PN model and a set of blocks (abstracted illustration).

3.1. *Creation of Exercises*

Instructors who are expert in software modeling using PN create exercises for learners. Each exercise consists of software requirements written in natural languages, a correct PN model as an answer, a set of blocks of which the correct PN model consists, and animated graphics.

The instructors construct the correct PN model based on the software requirements, and then the set of blocks is constructed by disassembling the correct PN model, as shown in Fig. 1. The set of blocks in Fig. 1 (b) is constructed by disassembling the PN model in Fig. 1 (a). Each block corresponds to a place, transition, token, or arc, and it is given the notation rules. Also, blocks of places and transitions are labeled with state names and event names, respectively (Note that the labels are omitted in Fig. 1, since it is an abstracted illustration). Some wrong blocks, that is, blocks that should not be included in the correct PN model can be added to the set of blocks, in order to increase the difficulty of an exercise. On the contrary, if the difficulty needs to be decreased, a semi-finished PN model and a small set of blocks that are used to finish it are prepared instead of the full set of blocks.

Fig. 2 shows an example of animated graphics of a crossing gate control system. The animated graphics consist of multiple graphical components, and the blocks include the internal programs to control the detailed behavior of the graphical components. For example, the animated graphics of a crossing gate control system will contain the graphical components of a train, crossing gate bars, crossing alarms, and so on. The instructors need to prepare the graphical components, integrate them into the animated graphics, and write the internal programs. The internal programs usually need not to be written by learners, and are hidden from the learners. However, if an internal program includes some statements that are essential to specify the behavior of software, the instructors need to translate the statements into a compact sentence in a natural language. A block including an internal program including essential statements is labeled with their compact sentence, and thus learners can understand them. The instructors can introduce a blank problem style by removing an important word or value from the compact sentence.

A newly created exercise is added to the collection of exercises that is open to all learners.

3.2. *Work on Exercises*

Each learner selects an exercise from the collection of exercises based on the domain and difficulty. The domain includes information systems, embedded software, and so on. The difficulty can be determined based on the size of the given set of blocks, the complexity of the correct PN model (for example, the number of branches on the reachability graph of the correct PN model), and so on.

First a learner reads the given software requirements, and then tries to construct his/her PN model that satisfies them. The learner selects a block from the given set of blocks, and puts it on a workspace. The learner is allowed only the connection among blocks that satisfies the notation rules of PN models. The learner can remove an arbitrary block that has already been connected with other blocks from the workspace, and then return it to the set of blocks. However, the learner is not allowed to disassemble a semi-finished PN model that has been given by instructors of the exercise.

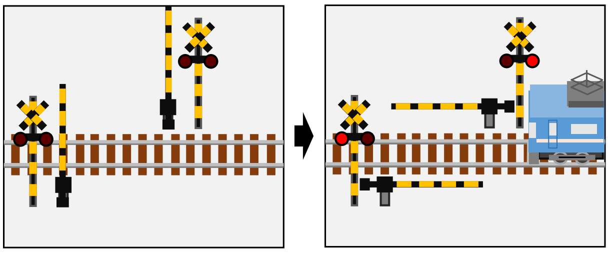


Fig. 2. Example of animated graphics.

The animated graphics are played according to the internal programs of blocks of which learner's PN model under construction consists, and give the learner a better understanding of its behavior. Additionally, if the learner cannot construct his/her PN model due to the high degree of the difficulty of the exercise, he/she can call in advisers, that is, ask instructors and other learners for their advice. The advisers can send their comments at any time, if the learner permits it.

3.3. *Review*

Before performing answer checking, the learner can call in reviewers, that is, ask instructors and other learners to review his/her PN model. The review comments from the reviewers will be useful not only to understand the modeling technique but also to maintain the motivation for learning. If the learner finds any errors in his/her PN model by the review comments, he/she tries to remove the errors. The reviewers can send their review comments at any time, if the learner permits it.

3.4. *Answer Checking*

When the learner has finished constructing his/her PN model, he/she performs answer checking, that is, his/her PN model is automatically compared to the correct PN model given by instructors. If they are not the same, it is wrong answer and the learner tries to remove his/her errors.

In general, a PN model consists of multiple sub-PN models that correspond to components, modules, or subsystems of software. For example, a PN model of a crossing gate control system will consist of sub-PN models that represent the behavior of a train, a crossing gate, and a controller. A sub-PN model that includes errors can be highlighted, in order to give the learner a clue to assist in removing the errors.

The learner can call in advisers or reviewers, and discuss with them.

3.5. *Demonstration*

If many learners cannot construct their PN models correctly in a specific exercise, an instructor who has created it gives a demonstration of how to construct a correct PN model for the learners.

4. Conclusion and Future Work

In this paper, we proposed a learning support technique of software visual modeling using PN in order that learners can intuitively understand the notation and behavior of a PN model. The ideas of block and animated graphics, which are well-known in the field of visual programming, are introduced into this technique. This technique consists of (1) creation of exercises, (2) work on exercises, (3) review, (4) answer checking, and (5) demonstration. It is assumed that learners are beginners in software modeling using PN. This technique should be implemented as a Web application in order to enable instructors and learners to interact with each other. The instructors' workload will not be small especially in the creation of exercises, and should be decreased by additional techniques and tools.

This technique is also related to our previous study. Ref. 4 shows a training support method for removing an error from a faulty EPN (Extended PN) model. Same as this technique, the training support method in Ref. 4 introduces animated graphics to illustrate the behavior of an EPN model, and thus this technique can be extended so as to support EPN. In EPN, a text-based formal modeling language is used to represent the essential and detailed behavior of software instead of a natural language, and visual programming of block type will be applicable to it.

In future work, we plan to develop a prototype tool to support our technique, and will evaluate its learning effect by using it.

References

1. M. Petre, UML in practice, *Proceedings of International Conference on Software Engineering*, May 2013, pp.722-731.
2. D. Bau, J. Gray, C. Kelleher, J. Sheldon and F. Turbak, *Learnable Programming: Blocks and Beyond*, Communications of the ACM, Vol.60, No.6, June 2017, pp.72-80.
3. N.G. Leveson and J.L. Stolzy, Safety Analysis Using Petri Nets, *IEEE Transactions on Software Engineering*, Vol.13, No.3, Mar. 1987, pp.386-397.
4. T. Takagi, S. Morimoto, Y. Ue and Y. Imai, Animated Graphics-Based Training Support Method and Prototype Tool for Bug Fixing of Extended Place/Transition Nets, *Journal of Robotics, Networking and Artificial Life*, Vol.5, No.4, Mar. 2019, pp.278-282.

Test Cases Redundant Elimination on Code Coverage Uses Distance and Correlation Measurement Method

Mochamad Chandra Saputra*, Tetsuro Katayama*, Yoshihiro Kita†,
Hisaaki Yamaba*, Kentaro Aburada*, and Nanoubu Okazaki*
**Interdisciplinary Graduate School of Agriculture and Engineering, University of Miyazaki,*
1-1 Gakuen-Kibanadai Nishi, Miyazaki, 889-2192 Japan
†School of Computer Science, Tokyo University of Technology, Hachioji, Tokyo 192-0982, Japan
Email: chandra@earth.cs.miyazaki-u.ac.jp, kat@cs.miyazaki-u.ac.jp, kitayshr@stf.teu.ac.jp,
yamaba@cs.miyazaki-u.ac.jp, aburada@cs.miyazaki-u.ac.jp, oka@cs.miyazaki-u.ac.jp

Abstract

The Euclidean distance and correlation measured by comparing the line executed by the test case to find the similarity of the test cases. The test cases have the lowest value of distance means highly similar and possible executing a similar line or path. The research tries to eliminate redundant test cases based on that similarity. Several redundant test cases are eliminated to get the best test cases. By Euclidean distance, the research can eliminate a similar test case on the test suite.

Keywords: Euclidean distance, redundant test case, test suite, code coverage

1. Introduction

The definition of the test case by IEEE is a set of test inputs, execution conditions, and expected results developed for a particular objective, such as to exercise a particular program path or to verify compliance with a specific requirement[1]. Test case repository or test suite may contain redundant, ambiguous, vague, and unfit test cases[2]. Using the test case similarity check on the test suite will decrease the redundant test cases.

The Euclidean distance is used to calculate many cases of similarity. The distance similarity between the test cases with the same length can be calculated by summing the ordered point-to-point distance[3]. The problem with the similarity of the test cases used the distance measurement is not checking the code coverage of the test case. The code coverage needs to check to confirm that each line of code executed on the testing process[4].

The white-box testing consists of analyzing the source code in order to guide the selection of test data uses several methods and concerns the internal logic of source code. Basis path testing guarantees that the test case executing every statement in the program at least one time during testing[5]. Using a given test suite on the source code throughout the software testing process helps to understand the code coverage [6].

This research enhances the previous research on test case redundancy[7] which combines preprocessing on test case similarity. In the previous research, the redundant test cases are investigated by path coverage using all the test cases on the test suite. The duplicate test case must be eliminated before investigating the redundant test case by path coverage to reduce the number of test cases on the test suite. The elimination of redundant test cases is to reduce the number of test cases on the test suite by using the Euclidean distance score. After that, code coverage on test cases is investigated using path coverage to eliminate redundant test cases and the best test cases on the test suite are acquired to achieve 100% code coverage.

2. The Principle of Redundant on Test Case

The consequence of redundant test cases is many test cases with no use. The redundancy will increases the

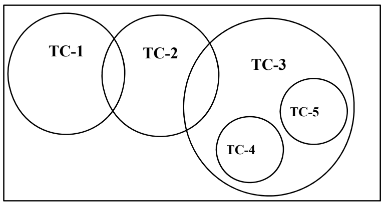


Figure 1. Concept of redundant test cases

the testing effort, cost and time of testing[8]. The fundamental concept of a redundant test case on code coverage will illustrate uses a union of sets on a mathematical approach. The test cases in the test suite may cover one or more code coverage. The research will identify redundant test cases based on coverage information using the line of code executed by the test cases.

The Venn diagram will be used to illustrate the concept of redundancy in Figure 1. The example, there are 5 test cases (TC-1, …, TC-5) that intersect with each other. The intersect means that some or all coverage from the test case containing another test case.

In Figure 1, there are 3 test cases (TC-1, TC-2, TC3) that have an intersection with other test cases. There are 2 redundant test cases because the coverage of the test case is already covered by another test case[9]. Finding and eliminating redundant test cases will increase the effectivity of the testing process.

3. Euclidean Distance

The experiment using the Euclidean distance calculation to find the similarity among the test cases on the test suites. The distance score on Euclidean is fiddling around with distance measures for some time especially with regard to profile comparison methodologies[10].

$$d = \sqrt{\sum_{i=1}^{p}\left(V_{1_i} - V_{2_i}\right)^2} \qquad (1)$$

The equation 1 is the Euclidean formula for calculating the distance between the two variables. Euclidean distance can calculate the similarity of the objects such as two person's profiles, a person and a target profile, in fact basically any two vectors taken across the same variables[10].

Identifying the similarity of a test case on the test suite is critical. Similar test cases in the test suite might repeat the same testing phase. The experiment tries to identify similar test cases by the distance score and then removing the several test cases. The lowest score of the similarity means that the test cases have highly similar.

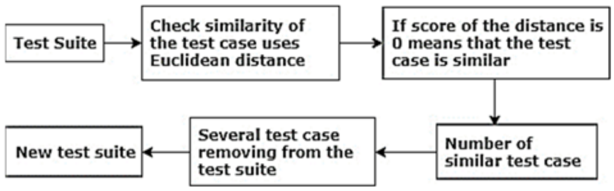


Figure 2. Step on similarity check on the test cases

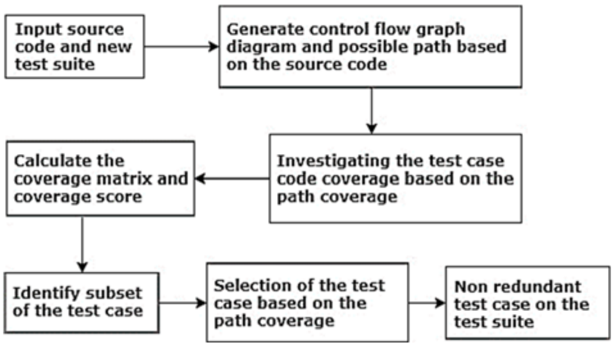


Figure 3. Step by step to find redundant test cases based on the path coverage

4. Methodology

The experiment suggests using distance and correlation measurement as preprocessing to minimize the number of test cases on the test suite. The distance and correlation measurement on the pre-processing is to calculate the similarity among the test case that already executed on the source code. The step by step preprocessing describes in Figure 2. The result from preprocessing is a new test suite without the similarities on the line executed by the test case.

The next step is to find the redundant test cases based on path coverage to ensure that the test suite achieves 100% code coverage. The path coverage is used to find the code coverage of the test case which already cover or under coverage by another test case. The coverage matrix and coverage score used to know the relation between the test case and path. The step by step to find redundant test cases based on the path coverage describes in Figure 3.

5. The Experiment

The experiment examines code coverage on the java source code with if-condition as shown in Figure 4.

```
package nestedif;
import java.util.Scanner;
public class IFCond {
    private static Scanner input;
    public static void main(String[] args) {
        int score;
        input = new Scanner(System.in);
          System.out.println(" Please Enter you Score: ");
          score = input.nextInt();
        if(score>=90 && score<=100){
            System.out.println("Your Grade is A");
        } else if (score<90 && score>=80){
            System.out.println("Your Grade is B");
        } else if (score <80 && score >=70){
            System.out.println("Your Grade is C");
        } else if (score <70 && score >=60){
            System.out.println("Your Grade is D");
        } else{
            System.out.println("You Faild on this class, try
next year");
        }
    }
}
```

Figure 4. Java source code with if-condition.

Table 1. Example of the result of the test case executed

Test case name	Result of the test case
TC-1	0,0,0,0,0,0,1,1,1,1,1,0,0,0,0,0,0,0,0,0,1,0;
TC-2	0,0,0,0,0,0,1,1,1,1,0,1,0,1,0,1,0,1,1,0,1,0;

Table 2. Example of the test case distance measurement by Euclidean distance

Test case distance	Distance score
1 & 2	2.4494897
1 & 3	2
1 & 4	2

The given test suite consists of 10 test cases that already executed on java source code. The result of executing the test cases used in the experiment is line executed.

Table 1 shows the data from the test case, TC-1 is the name of the test case. The value of the test case result consists of the lines executed by the test case. The 0 value means that the line on the code is not inspected by the test case and 1 means that inspected.

The first process of the experiment is to find a similar test case uses Euclidean distance from the test suite to reduce the number of test cases. The example of the result is shown in Table 2. The distance score defines the similarity among the test case. The value of the distance score is 0 means that the test case is definitely similar. The experiment can remove several test cases to eliminate a similar test case.

The new test suite without a similar test case will be used on the next step. Before investigating the redundant test cases based on the code coverage using the path coverage method, the experiment generated

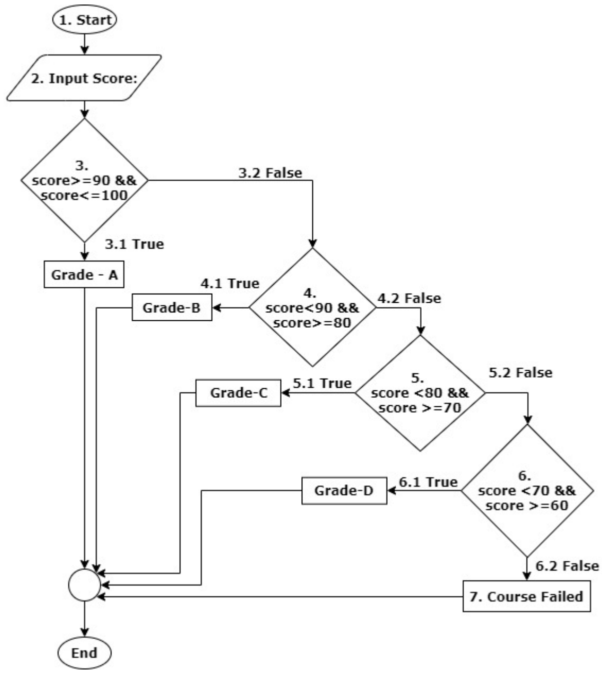


Figure 5. Flow graph of the if-condition java source code

Table 3. Result of coverage matrix

Path	Subset
P1	S1=(TC-1)
P2	S2=(TC-8)
P3	S3=(TC-4)
P4	S4=(TC-5)
P5	S5=(TC-2, TC-6, TC-7)

the control flow graph and the possible path of the if-condition java source code as shown in Figure 5.

Based on the flow graph, the path of the source code is as follow:

- Path-1 (P1): 1-2-3-3.1-End
- Path-2 (P2): 1-2-3-3.2-4-4.1-End
- Path-3 (P3): 1-2-3-4-4.2-5-5.1-End
- Path-4 (P4): 1-2-3-4-5-5.2-6-6.1-End
- Path-5 (P5): 1-2-3-4-5-6-6.2-7-End

The new test suite is investigated to the path and the result using on coverage matrix and coverage score calculation. The coverage matrix describes the possibility of the test case on the path. On the other hand that one path may be investigated by no one or more than one test case such as shown in Table 3.

The coverage score is calculated by the coverage matrix result. The coverage score describes the number of paths executed by the test case divided by all number of paths. The example result of a coverage score below contains coverage score (CS) on TC-n and then using the coverage score calculation.

1. CS(T1)= 1/5 = 0.2
2. CS(T2)= 1/5 = 0.2

Identifying the subset of the test cases used the coverage matrix and coverage score to guide on selecting the final test suite without a redundant test case.

6. Result and Discussion

The experiment uses the pre-processing to find a similar test case based on the score of the distance. The result of distance calculation using Euclidean distance with distance score 0 are TC-2 & TC-6,TC-2 & TC-7, TC-3 & TC-9, TC-3 & TC-10, TC-6 & TC-7 and TC-9 & TC-10.

Several test cases must be eliminated to find a unique test case. The experiment will select and delete several test cases by random. This is the step to select the test cases.

(i) TC-2==TC-6, TC-6 Deleted, Result: TC-2
(ii) TC-2==TC-7, TC-7 Deleted, Result TC-2
(iii) TC-3==TC-9, TC-9 Deleted, Result TC-3
(iv) TC-3==TC-10, TC-10 Deleted, Result TC-3
(v) TC-6==TC-7, already deleted.
(vi) TC-9==TC-10, already deleted.

The result of the selection of similarity test cases is TC-2, TC-3 and test case deleted is TC-6, TC-7, TC-9, TC-10. The test case elimination by the distance score is not considered the test case code coverage. The result of the selection test case similarity then uses to check the path coverage of the test case to ensure the code coverage. The new test suite consists of TC-1, TC-2, TC-3, TC-4, TC-5, TC-8. The result after the elimination of a similar test case, only 6 test cases need to verify the redundancy and TC-3 is redundant. The redundant on this condition because of the TC-3 is covered by another test case. The result of the new test case can achieve 100% code coverage is TC-1, TC-2, TC-4, TC-5, TC-8.

7. Conclusion

The research confirms the number of redundant test cases on the test suite is reduced because of the similarity. The current research eliminates a similar test case before checking the redundancy of the test case based on path coverage. In the experiment, the new test suite selected by Euclidean distance that must be verified on redundancy by path coverage is reduced by 40% and also satisfy 100% code coverage with 1 test case redundant that analyzes by path coverage.

Elimination of test cases based on coverage information does not guarantee to keep the fault detection capability of a given test suite. Future research is needed to consider the similarity in fault detection capability of test cases for the purpose of redundancy detection.

References

1. The Institute of Electrical and Electronics Engineers - IEEE. *IEEE Standard Computer Dictionary - A Compilation of IEEE Standard Computer Glossaries*. 1990.
2. Mohanty, H., Mohanty, J. R. & Balakrishnan, A. *Trends in Software Testing*. (Springer Singapore, 2017). doi:10.1007/978-981-10-1415-4.
3. Cassisi, C., Montalto, P., Aliotta, M., Cannata, A. & Pulvirenti, A. *Similarity Measures and Dimensionality Reduction Techniques for Time Series Data Mining*. in Advances in Data Mining Knowledge Discovery and Applications (InTech, 2012). doi:10.5772/49941.
4. Saputra, M. C. & Katayama, T. *Code Coverage Visualization on Web-Based Testing Tool for Java Programs*. J. Robot. Netw. Artif. Life 2, 2015 , pp. 89.
5. Roger S. Pressman, P. D. *Software Engineering: A Practitioner's Approach*. (McGraw-Hill, 2010).
6. Heed, P. & Westrup, A. *Automated platform testing using input generation and code coverage*. Department of Computer Science, Lund University, 2009.
7. Panday, A., Gupta, M., Singh, M. K. & Ali, N. *Test Case Redundancy Detection and Removal Using Code Coverage Analysis*. MIT Int. J. Comput. Sci. Inf. Technol. 3, 2013 , pp. 6–10.
8. Alian, M., Suleiman, D. & Shaout, A. *Test Case Reduction Techniques - Survey*. Int. J. Adv. Comput. Sci. Appl. 7, 2016 , pp. 264–275.
9. Koochakzadeh, N., Garousi, V. & Maurer, F. *Test Redundancy Measurement Based on Coverage Information: Evaluations and Lessons Learned*. in 2009 International Conference on Software Testing Verification and Validation (IEEE, 2009). , pp. 220–229 doi:10.1109/ICST.2009.8.
10. Barrett, P. *Euclidean Distance Raw, Normalized, And Double-Scaled Coefficients*. The Technical Whitepaper Series 6, 2005.

The Measurement of Class Cohesion using Semantic Approach

Bayu Priyambadha*, Tetsuro Katayama*, Yoshihiro Kita†, Hisaaki Yamaba*, Kentaro Aburada*, Nanoubu Okazaki*
**University of Miyazaki, 1-1 Gakuen-kibanadai nishi, Miyazaki, 889-2192, Japan*
†Tokyo University of Technology, 1404-1 Katakuramachi, Hachioji City, Tokyo 192-0982, Japan
E-mail: bayu@earth.miyazaki-u.ac.jp, kat@cs.miyazaki-u.ac.jp, kitayshr@stf.teu.ac.jp, yamaba@cs.miyazaki-u.ac.jp, aburada@cs.miyazaki-u.ac.jp, oka@cs.miyazaki-u.ac.jp

Abstract

The cohesion is one of the design quality indicators that can be measured by focusing on the correlation between elements inner the class. In Direct Distance Design Class Cohesion (D_3C_2) metrics, elements are assumed correlated if they have a similar type. In this research, the correlation between elements' inner class is not only based on type similarity but also on the meaning of the element's name. The results confirm that the value of cohesion increases in line with the similarity of meaning.

Keywords: Software Engineering, Software Quality, Design Quality, Cohesion Metric, D_3C_2 Metric.

1. Introduction

Software engineering aims to provide a way or method to build a qualified software system[1]. The guarantying the quality of the software not only focuses on one phase, but it should be maintained at every phase so that the resulting software that has high quality. The design phase aims to produce a description of the structure of the software, data models, data structures, interfaces between system components, and the algorithms used[1]. The quality of design affects the final result of the software. Cohesion is one of the indicators for assessing the quality of a result of design[2–4].

Cohesion is a level of relatedness between elements within a component[2]. The higher the value of cohesion in a component, then the better the design[5,6]. For example, in the development of software with an object-oriented approach, there is a class which is a component, an element that is in the class include the attributes and methods[2]. The high cohesion can increase the maintainability of classes that exist in the system. The changes in one class will not affect the other class. The maintenance of the system can be easily and focus on problem[6].

The process of measuring the value of the cohesion of a class will be very useful in maintaining the quality of the design[2,3]. Measurement of cohesion at the design phase aims to provide information about the quality of the design as soon as possible. Knowing it can save the cost and effort of developers to perform maintenance.

One metric that can be used to measure the value of cohesion that considers the interrelationships between attributes and methods is the Distance Design-based Direct Class Cohesion and then called D_3C_2 [2]. D_3C_2 metric considering the relatedness between elements by seeing the similarity of parameter type in the method of class and the unique attribute type. A method assumed has high relatedness if the parameter type of method is the same as the attribute type of the class. At the design phase, the source code of the system has not been defined yet. The certainty that an attribute is manipulated by a method is low. This condition raises a thought that similarity parameter types and attribute types do not always indicate that an attribute associated with the method.

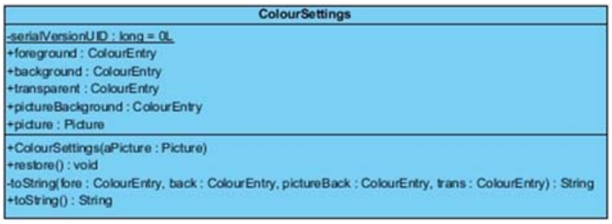


Fig. 1. Class ColourSettings

On the other hand, some attributes are manipulated by a method which does not have the same parameter types as attribute types.

This study tries to explore the information that can enhance the degree of certainty of the relationship between methods and attributes of the class. Based on the theory, the maintainability of software is affected by the naming of a variable (attributes and method)[1]. One of the purposes of this research is to find the correlation between attributes and methods not only looked at the similarity of type but also the name. The similarity of names is not only seen from the similarities of syntax but also views of the similarity of meaning (semantics). Then, the experimental results applied to the D_3C_2 metrics to calculate the value of cohesion in a real case. The result using D_3C_2 cohesion metrics with the semantic approach is used to compare with the previous approach.

2. Cohesion Metric

Direct Attribute Type (DAT) is a matrix used to map the relatedness between method and attribute[2]. DAT is created as the basis for the calculation of the D_3C_2 metric. DAT matrix prepared by comparing the method's type and the attribute's type. The example of the DAT matrix from class ColourSettings (Figure 1) is described in Table 1. Based on the DAT matrix, the following four metrics have to calculate before D_3C_2 will calculate.[2]

2.1. *Method-Method through Attributes Cohesion (MMAC)*

Method-Method Attributes through Cohesion (MMAC) is a metric to calculate an average value of cohesion of all pairs of methods. MMAC is formally written as follows:

$$MMAC(C) = \begin{cases} 0, & if\ k = 0\ or\ l = 0, \\ 1, & if\ k = 1, \\ \frac{\sum_{i=1}^{l} x_i(x_i - 1)}{lk(k-1)}, & otherwise. \end{cases} \quad (1)$$

Table 1. DAT Matrix using Previous Research

	long	ColourEntry	ColourEntry	ColourEntry	ColourEntry	Picture
ColourSettings	0	0	0	0	0	1
restore	0	0	0	0	0	0
toString	0	1	1	1	1	0
toString	0	0	0	0	0	0

Where x is the number of the value of 1 in one column (j), k is the number of methods, and l is the number of attributes in the class.

2.2. *Attribute-Attribute Cohesion (AAC)*

Attribute-Attribute Cohesion (AAC) is a metric to calculate the average value of cohesion of pair of attributes. AAC is formally written as follows:

$$ACC(C) = \begin{cases} 0, & if\ k = 0\ or\ l = 0, \\ 1, & if\ l = 1, \\ \frac{\sum_{i=l}^{k} y_i(y_i - 1)}{kl(l-1)}, & otherwise. \end{cases} \quad (2)$$

Where y is the number of the value of 1 in one row (i).

2.3. *Attribute-Method Cohesion (AMC)*

Attribute-Method Cohesion (AMC) is a metric to calculate an average value of cohesion based on the interaction of attributes and methods. AMC is formally written as follows:

$$AMC(C) = \begin{cases} 0, & if\ k = 0\ or\ l = 0, \\ \frac{\sum_{i=1}^{k} \sum_{j=1}^{l} m_{ij}}{kl}, & otherwise. \end{cases} \quad (3)$$

Where i is the number of rows in the matrix, j is the number of columns in the matrix, k is the number of methods in the matrix, and k is the number of the attribute in the matrix.

2.4. *Distance Design-based Direct Class Cohesion (D_3C_2)*

This process can't be defined if a class does not have class methods and attributes. D_3C_2 metrics used to calculate the final summation of the result of MMAC, AAC, and AMC. D_3C_2 is formulated as follows:

$$D_3C_2(C) = \begin{cases} 0, & if\ k = 0\ and\ l = 1, \\ 1, & if\ k = 1\ and\ l = 0, \\ \frac{k(k-1)MMAC(C) + l(l-1)ACC(C) + 2lkAMC(C)}{k(k-1) + l(l-1) + 2lk}, & otherwise. \end{cases} \quad (4)$$

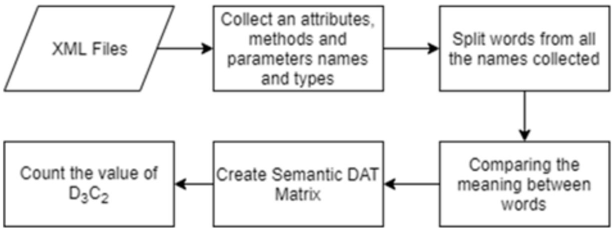


Fig. 2. The Process of Calculation

3. Semantic Similarity

Dictionary or repository of words that can be used to assist in the identification of words that mean the same thing has been developed and used in several studies [7,8]. A WordNet is a dictionary that has been prepared based on the relation of synonyms, antonyms, hyponyms, and hypernyms, meronyms, troponin, and entailment relationships[9]. Wu and Palmer formulate a way of comparing the meanings of the two words by considering the proximity of the relations in the word's dictionary[10].

In a study conducted by Dijkman[11], Dijkman defines a formula for calculating the similarity between the two labels or sentence by considering the similarity of meanings (synonyms). Semantic similarity can be calculated using the following formula:

$$\frac{2.wi.|w_1 \cap w_2| + ws.(|s(w_1,w_2)| + |s(w_2,w_1)|)}{|w_1| + |w_2|} \quad (6)$$

Where w_1 and w_2 are the collections of a word from every compared sentence. $s(w_1, w_2)$ or $s(w_2, w_1)$ is the number of words that has a synonym relationship between two sentences. wi and ws are the weight that defined for a similar word and the word that has semantic similarities (synonym). Dijkman defines the value of $wi = 1$ and $ws = 0.75$ [11].

4. Methodology

Figure 2 shows the flow of the automatic calculation system. The first step in this research is receiving the XML files. Those files are generated from design tools named Visual Paradigm for UML. The calculation of the value of cohesion performed by using prototype software that developed using Java language. The results will be stored in storage, which will then be analyzed.

In D_3C_2 metric calculation process, it is necessary first to calculate metrics MMAC, ACC, and AMC.

Table 2. Proposed DAT Matrix (Semantic DAT)

	(serialVersionUID) long	(foreground) ColourEntry	(background) ColourEntry	(transparent) ColourEntry	(pictureBackground) ColourEntry	(picture) Picture
ColourSettings	0	0	0	0	1	1
restore	0	1	0	0	0	0
toString	0	1	1	1	1	0
toString	0	0	0	0	0	0

Before calculating the MMAC, ACC, and AMC, the DAT matrix must be prepared to describe the relationship between the methods and attributes. In previous studies, the DAT matrix is formed by giving the value 1 for the pair method and attribute which has the same type. In this study, the DAT matrix is prepared by comparing the method of the semantic similarity of the name of the method and attributes besides considering the similarity of type. If there is no type recognized, then the semantic similarity is considered. The system has to be able to split the words inner both names to make a comparison semantically between words from the name of method and attribute.

After split the name of methods and attributes, then every word is comparing semantically using Eq. (6) And a threshold of 0.5. WordNet will enrich the Eq. (6) to calculate the similarity of words. If the score is above the threshold, then it is considered as similar semantically. The example of the semantic DAT matrix is described in Table 2.

5. Result and Discussion

It needs several things to test the prototype. The input is the class diagram that exported as an XML file by using Visual Paradigm. The data selected for this testing is the source code of the application jDraw version 1.1.5 available on the website www.sourceforge.net.

5.1. *Result*

The result of the calculation will be compared with the manual observation to get the conformance between the approach. Calculating the conformance between approach is using the Kappa Coefficient. The Kappa coefficient between the previous approach and the semantic approach is descript in Table 3 and Table 4.

The result shows that there is an increment of value of Kappa from 0.055954 to 0.29474 between the previous and semantic approaches. Based on the level of Kappa, it is also increasing from slight agreement to fair agreement.

Table 3. Kappa of the Previous Approach

KOHENS KAPPA

	Observer		
System	**Y**	**N**	**Total**
Y	49	343	392
N	123	1401	1524
Total	172	1744	1916
Po	0.756785		
Pc	0.7423695		
Kappa	**0.055954**	slight agreement	

Table 4. Kappa of Semantic Approach

KOHENS KAPPA

	Observer		
System	**Y**	**N**	**Total**
Y	54	82	136
N	118	1662	1780
Total	172	1744	1916
Po	0.895616		
Pc	0.851992		
Kappa	**0.29474**	fair agreement	

5.2. *Discussion*

Examples of cases of jDraw are a class named ColourSettings, as illustrated in Figure 1. Table 5 shows the result of calculation D_3C_2 using the previous approach and the semantic approach.

D_3C_2 metric final results showed an increase in the value of 0.144 into 0.2 (0.056 difference). In a previous study, attributes pictureBackground has no connection with the method ColourSettings (constructor), due to ColourEntry data types not contained in the type parameter or returns data type of a method ColourSettings. pictureBackground has close meaning with the method ColourSettings and parameter aPicture using the semantic approach. In the source code of method ColourSettings, the method manipulates attributes pictureBackground. It shows that there is a connection between the pictureBackground attribute with a constructor method ColourSettings. Figure 3 shows the presence of pictureBackgound in the body of the ColourSettings method.

6. Conclusion

Cohesion calculation at the design phase has challenges because of the lack of information is provided by the design artifact, such as class diagram. The application of the semantic approach in calculating D_3C_2 metrics can increase the conformance of the calculation results. In the future, it is essential if there is any additional information considered other than class diagrams.

Table 5. Comparison D_3C_2 Value

Previous Approach				Semantic Approach			
MMAC	ACC	AMC	D_3C_2	MMAC	ACC	AMC	D_3C_2
0.0	0.1	0.21	0.14	0.05	0.11	0.29	0.2

```
public ColourSettings(Picture aPicture) {
    picture = aPicture;
    Palette pal = picture.getCurrentPalette();
    foreground = pal.getColour(picture.getForeground());
    background = pal.getColour(picture.getBackground());
    pictureBackground = pal.getColour(picture.getPictureBackground());
    final int t = picture.getTransparent();
    if (t == -1) {
        transparent = null;
    }
    else {
        transparent = pal.getColour(t);
    }
}
```

Fig. 3. The Code of Method ColourSettings

The process flow in the method described in the flow diagram or pseudo code in the design phase is worth considering.

References

1. Sommerville, I. *Software Engineering*. (Addison-Wesley Professional, 2011). DOI:10.1111/j.1365-2362.2005.01463.x.
2. Al Dallal, J. *A design-based cohesion metric for object-oriented classes*. Int. J. Comput. Electr. Autom. Control Inf. Eng. **1**, 2007 , pp. 195–200.
3. Al Dallal, J. & Briand, L. C. *An object-oriented high-level design-based class cohesion metric*. Inf. Softw. Technol. **52**, 2010 , pp. 1346–1361.
4. Chowdhury, I. & Zulkernine, M. *Using complexity, coupling, and cohesion metrics as early indicators of vulnerabilities*. J. Syst. Archit. **57**, 2011 , pp. 294–313.
5. Pressman, R. S. *Software Engineering : A Practitioner's Approach*. (McGraw-Hill, 2015).
6. Chen, Z., Zhou, Y., Xu, B., Zhao, J. & Yang, H. *A Novel Approach to Measuring Class Cohesion Based on Dependence Analysis*. Int. Conf. Softw. Maint. , 2002 , pp. 377–384 DOI:10.1109/ICSM.2002.1167794.
7. Wei, T., Lu, Y., Chang, H., Zhou, Q. & Bao, X. *A semantic approach for text clustering using WordNet and lexical chains*. Expert Syst. Appl. **42**, 2015 , pp. 2264–2275.
8. Gao, J. B., Zhang, B. W. & Chen, X. H. *A WordNet-based semantic similarity measurement combining edge-counting and information content theory*. Eng. Appl. Artif. Intell. **39**, 2015 , pp. 80–88.
9. Miller, G. a. *WordNet: a lexical database for English*. Commun. ACM **38**, 1995 , pp. 39–41.
10. Wu, Z. & Palmer, M. *Verb semantics and lexical selection*. 32nd Annu. Meet. Assoc. Comput. Linguist. , 1994, pp. 133–138 DOI:10.3115/981732.981751.
11. Dijkman, R., Dumas, M., van Dongen, B., Käärik, R. & Mendling, J. *Similarity of business process models: Metrics and evaluation*. Inf. Syst. **36**, 2011 , pp. 498–516.

Proposal of an Algorithm to Generate VDM++ by Using Words Extracted from the Natural Language Specification

Yasuhiro Shigyo*, Tetsuro Katayama*, Yoshihiro Kita†,
Hisaaki Yamaba*, Kentaro Aburada*, and Naonobu Okazaki*
**University of Miyazaki, 1-1 Gakuen-kibanadai nishi, Miyazaki, 889-2192, Japan*
†Tokyo University of Technology, 1404-1 Katakura, Hachioji, 192-0914, Japan
E-mail: shigyo@earth.cs.miyazaki-u.ac.jp, kat@cs.miyazaki-u.ac.jp, kitayshr@stf.teu.ac.jp,
yamaba@cs.miyazaki-u.ac.jp, aburada@cs.miyazaki-u.ac.jp, oka@cs.miyazaki-u.ac.jp

Abstract

The natural language includes ambiguous expressions. VDM is one of methodology on the formal methods to write the specification without ambiguity. Because VDM++ is written by strict grammar, it is difficult to write a VDM++ specification. This research attempts to generate a VDM++ specification automatically from a natural language specification by machine learning. To generate a VDM++ specification, it is necessary to extract words which consist of predicate corresponding to the function and nouns corresponding to variable from the natural language specification. This paper proposes an approach to generate a VDM++ specification from the extracted words list. Identifiers are generated from the extracted words list, and then the VDM++ specification can be generated by converting them into VDM++ grammar.

Keywords: natural language specification, VDM++, automatic generation, formal method, formal specification

1. Introduction

Software bugs affect a huge impact on our society[1,2]. Most of them can be caused by the general use of a natural language in the upstream process of software development because the natural language contains ambiguity. The ambiguity can lead programmers to misinterpret the specification[2]. As a result, the programmer might embed some bugs into the program.

One way of solving this problem is to design software in the upstream process by using a formal method. In development by using it, specifications are written in a formal specification language based on mathematical logic. They can be proved the mathematically by using theorem proving and mechanical checking[3]. It allows for precise design without the ambiguity.

VDM (Vienna Development Method)++ is a formal specification language that can handle object-oriented modeling[4]. VDM++ has strict grammar. The strict grammar makes them difficult to describe because they are needed to write data types and system invariants that are not in a natural language specification.

This study attempts to generate a VDM++ specification automatically from a natural language specification by machine learning. The automatic generation of a VDM++ specification, it is necessary to extract words that consist of predicates and nouns from the natural language specifications. The predicates are corresponding to functions and nouns are corresponding to variables, respectively.

However, it is difficult to use the extracted nouns and predicates along the grammar of VDM++. A VDM++

Table 1. The definitions in VDM++.

VDM++ keyword	Number of elements	Syntax
values	2	$1^{st} = 2^{nd}$;
types	2	public $1^{st} = 2^{nd}$;
instance variables	3	1^{st}: 2^{nd}: = 3^{rd} ;
operations	5 or more	public 1^{st} : 2^{nd} ==>3^{rd} pre 4^{th} post 5^{th} ;

specification cannot be described from the only extracted nouns and predicates.

This paper proposes an approach to generate VDM++ specification from the extracted nouns and predicates.

2. VDM++

VDM is one of the formal methods. VDM++ is an object-oriented extension of the VDM-SL language and is currently most used in VDM[4].

Table 1 shows the definitions in VDM++ targeted in this paper. The keywords contain several elements and syntax information. The syntax is the set of description rules in VDM++, and each element corresponds to ordinal numbers in the syntax in Table 1[4].

3. The Proposed Approach

The proposed approach automatically generates a VDM++ specification. Fig.1 shows the proposed approach.

1. Extract a word list from a natural language specification. It is referred to as a word list as the extracted word list.
2. Classify word list by machine learning from the extracted word list. It is referred to as a word list as the classified word list.
3. Insert an identifier into the classified word list.
4. Get the identifier from the word list into which the identifier is inserted.
5. Put the elements of each word list into the VDM++ specification syntax corresponding to the gotten identifier.

This paper focuses on 3-5 steps (lower frame of Fig.1). Here, the generation of a classified word list by machine learning from the natural language specification in 1-2 steps is a future issue.

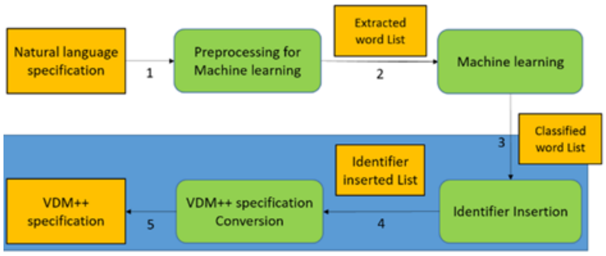


Fig. 1: The flow of the proposed approach

3.1. *The Data structure*

The proposed approach defines two data structures: classified word list and identifier.

3.1.1. *Classified Word List*

A classified word list is a line-by-line list of elements required to generate a syntax for a variable or function. Fig. 2 shows the flow of the identifier insertion process described in Section 3.3. It also shows an example of a classified word list at the upper left.

Each column of the classified word list is explained as follows. The element in the 1st column (Column A) is a temporary ID, and it is inserted into an identifier later. The element in the 2nd column (Column B) is the variable name or function name described in the VDM++ specification.

The elements in the 3rd column (Column C) and the subsequent columns differ depending on each target: variables, instance variables, and functions. The elements for each target are shown below.

- In the case of the variables, the element in column C is real values.
- In the case of the instance variable, the columns after column C consist of the type name and initial value, which may not be specified.
- In the case of the function, the columns after column C consist of a pre-condition, post-condition, function argument, delimiter "#", and return type.

3.1.2. *Identifiers*

Identifiers are generated based on the number of elements in the classified word list corresponding to the keywords of the VDM++ in Table 1. The insertion condition of each identifier in the classified word list is described as follows. The "values" keyword indicates that the number of elements in the second and subsequent columns of the

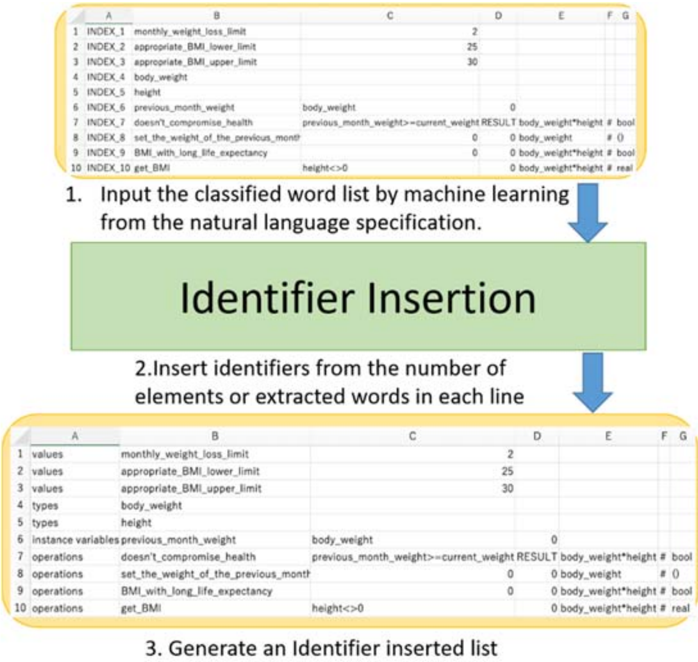


Fig. 2: The flow of the identifier insertion process

classified word list is 2, and the third element is a value. The "types" keyword indicates that the number of elements in the second and subsequent columns of the classified word list is 2, and the third element is a type definition. If the number of elements is 1, the element in the third column is real. The "instance variables" keyword has a row with 3 elements. The "operations" keyword has greater than equal to 5 elements.

3.2. *The Approach Process*

In order to generate VDM++ specification, it is proposed two algorithms: Identifier Insertion Process and VDM++ Specification Conversion Process.

3.2.1. *Identifier Insertion Process*

Fig. 2 shows the flow of the identifier insertion process. The identifier insertion process replaces a temporary ID in the 1st column (Column A) with an identifier for each row of the classified word list following the conditions described in Section 3.1.2.

3.2.2. *VDM++ Specification Conversion Process*

Fig. 3 shows the flow of the VDM++ specification conversion process. The VDM++ specification conversion process describes the VDM++ specification based on the identifier inserted list.

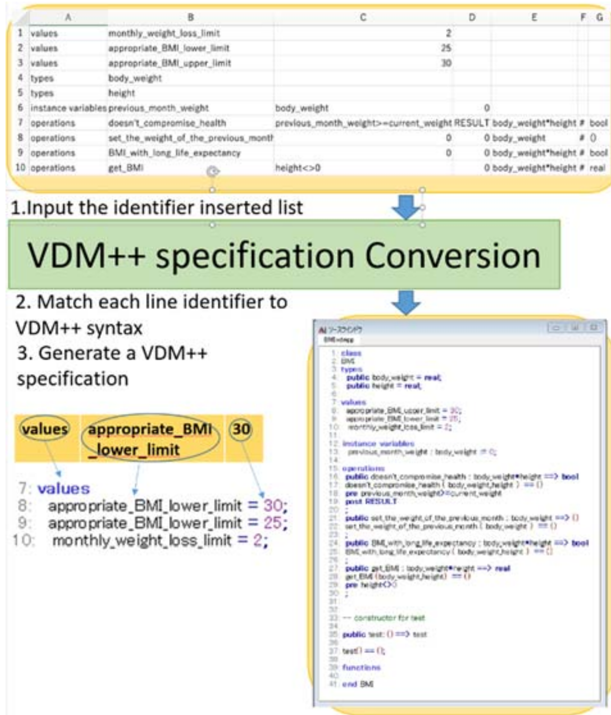


Fig. 3: The flow of VDM++ specification conversion process

How to describe the VDM++ specification is explained. Firstly, the "class <class-name>" is described on the first line. The word "<class-name>" is described the class name of the specification. Next, each VDM++ keyword is described. Statements are written after the keyword according to the syntax. Finally, the "end <class-name>" is described to the end of the line.

As an example, the third row of the list in which the identifier is inserted in Fig.2 is described. The first step reads the identifier of the first column and finds the "values". The second step, according to its syntax " 1st = 2nd ;", associates "appropriate_BMI_upper_limit" in the 2nd column (Column B) and "30" in the 3rd column (Column C) of the identifier inserted list, respectively. Final step writes the statement "appropriate_BMI_upper_limit = 30 ;" after keyword "types" that in the VDM++ specifications.

4. Application Examples and Discussion

We have developed two tools: an inserter and a converter, which are implemented the two proposed process. They have applied a classified word list to the inserter and an identifier inserted list to the converter.

Fig. 2 shows the input and output of the inserter. For each row in the output list, we verify that the identifier can be inserted correctly as follows. Rows 1-3 in the classified word list represents a variable with an initial value. Because the number of the elements is two and the third column (Column C) is a number, the inserter inserts an identifier “values”. Rows 4-5 represents a variable that defines only the types. Because the number of the elements is 1, the inserter inserts an identifier “types”. Row 6 represents the instance variable. Because the number of the elements is following or more, the inserter inserts an identifier “operations”. Thus, the inserter can generate an identifier inserted list correctly.

The upper left of Fig. 3 shows the input of a converter, and the right of Fig. 3 shows the output of the converter. The output of the inserter in Fig. 2 is the same as the input of the converter in Fig. 3. We verify that the converter generates a VDM++ specification correctly based on the syntax in Table 1 as well as the inserter. We confirm that the converter can generate the VDM++ specification that consists of statements corresponding to the syntax in Table 1.

Furthermore, we verify that the VDM++ specification generated by the tools based follows the VDM++ syntax. Fig.4 shows an output of the converter to describe the VDM++ specification. VDM++ Toolbox with syntax and type checker displays no warning for the generated VDM++ specification. Therefore, we have confirmed that the proposed process can automatically generate the VDM++ specification corresponding to the VDM++ syntax from the classified word list.

5. Conclusion

This paper has proposed the approach to aim at automatically generating a VDM++ specification from the classified word list. It means this paper proposes 3-5 steps (lower frame of Fig. 1). As a constraint, it assumes that a classified word list is generated by the machine learning from a natural language specification by the steps 1-2. As a result, the VDM++ specification can be automatically generated from the list, which will be able to classify from the natural language specification by machine learning.

We confirmed that the VDM++ specification is generated by applying the classified word list to the proposed approach. From this result, it can be useful by the automatic generation of VDM++ specifications.

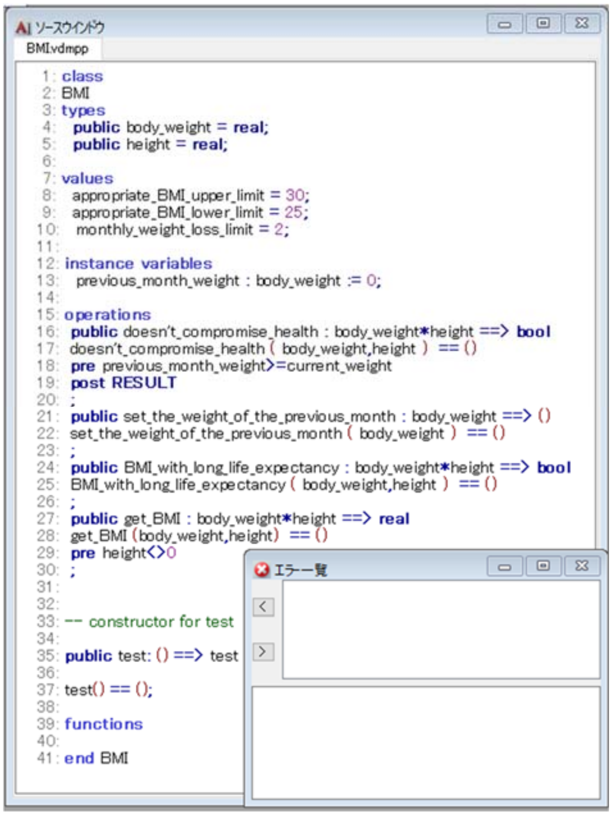


Fig. 4: The VDM++ specification displayed by VDM++ Toolbox

Future issues are as follows:

- Generating a classified word list by machine learning from a natural language specification
- Corresponding to other syntax in VDM++

References

1. E. Marcus and H. Stern, “Blueprints for high availability”, John Willey and Sons, 2000.
2. National Institute of Standards and Technology (NIST), Department of Commerce, “Software errors cost U.S. economy $59.5 billion annually”, NIST news release 2002-10, 2002.
3. M. Fähndrich, M. Barnett, and F. Logozzo. “Embedded contract languages”, In the 2010 ACM Symposium on Applied Computing (SAC 2010), pp.2103–2110, ACM, 2010.
4. International Organization for Standardization, ISO/IEC JTC 1/SC 22/WG, “Information technology--Programming languages, their environments and system software interfaces--Vienna Development Method--Specification Language--Part 1: Base language”, 1996.

Customization of Contents for Acquisition of Skills of FPS without Trainer

Masao Kubo, Takeshi Ueno, Hiroshi Sato
Computer Science, National Defense Academy, Hashirimizu-10-20
Yokosuka, Kanagawa 239-8686, Japan
E-mail: masaok@nda.ac.jp, hsato@nda.ac.jp

Abstract

In this paper, we conducted a study to develop self-study materials with the ability to provide more suitable materials for students' skills. A first person shooter (FPS) type simulator-based teaching material is created to acquire knowledge. Conventional games resume from a predetermined scene regardless of trainee's skill contributes to no increases in knowledge and skill. In this paper, we propose a re-spawning point suitable for each trainee by a recommendation algorithm which tries to find good game scenes by trial and error.

Keywords: Reinforcement Learning, E-Learning, Game AI, Virtual Reality.

1. Introduction

The role of E-learning is growing wider and wider every year, so better quality E-learning system is required for a wide variety of classrooms. The main issue of update of quality of E-learning is the quality of recommendation of contents which fit the student's skill. Some of E-learning curriculums need human support staff to re-offer the contents because the machine-selected contents are unfitted to the students. Therefore, self-study using by such a poor grade of E-learning curriculums is difficult.
We focus on the respawn point as the automatic content generation/recommendation of the video game based E-Learning. Response point is a game resuming point.
Generally speaking, there are some games which a player has to start the game from the beginning whenever a player fails the game along the way. In such a game, in sometimes a content that does not contribute to an increase in knowledge and skill is forced to be played. This paper attempts to propose a respawn point suitable for students' skill by using a recommended AI that gains experience while proposing a scene through trial and error.
This paper is organized as follows. In the next section, the FPS game as our text to learn is explained. We have to design the game like text in which whoever can learn new knowledge regardless of any skill and knowledge. So in this game-like text we adopt zombies, which everyone already know attack the player but there is no one who knows actual behavior. In the third section the algorithm to recommend a better response point for a student to expect to acquire the necessary knowledge and skill for its ability is proposed. The final result of the simple experiment by 20 university students is shown.

2. The E-learning text of FPS style which it can learn how to escape from zombie

2.1. *The overview of the E-learning*

This FPS text is made by Unity, a commercial 3D game design tool. Fig.1 illustrates the game field. Buildings of a high school are simulated in the cyber space.

Fig.1 The FPS-like lesson field in cyber space

The site of the school is surrounded by walls and there is the only one gate to go to the outside from the inside of the school. 50 zombies are set around the gate. The avatar of the student starts from the farthest point from the gate and it aims to reach the gate. Zombies can know the location of the avatar precisely and they go toward to the avatar's current location. If a zombie collides the avatar a trial is over.

2.2. *"Respawn" and "Respawn point"*

As shown in Fig.2 a student uses this lesson material by First Person Shooter Style (FPS).

Fig.2 A scene of lesson

The operation of its avatar is done by moving the keyboard and changing the viewpoint by moving the mouse. One trial is to reach the gate of the entrance or get caught by the Zombie. When replaying, call the respawn to change the position of the player and zombies and start the game. The point or its corresponding scene is called the respawn point.

3. The proposed respawn algorithm

Here, the successful and unsuccessful attempts collected in advance are made into a database, and a respawn point that is easy to learn is generated. The respawn point is automatically generated in three stages after a play failure.(Step 1) First, select the teacher player by comparing the record of the player registered in the database and the student's play history. (Step 2) Next, select a successful play closest to the failed play from the history of the selected teacher player, and select a promising point as the respawn point from the initial state at the next trial. (Step 3) Update the value of the respawn point adopted according to the result of the trial.
The following briefly describes each of these steps.

3.1. *Step1: Selection of teacher player*

Compares the player-level data recorded in the database with the student data, and selects a teacher player. There are many methods can be thought. In our current system we adopt the maximum information gain approach. The teacher who can expect the largest information gain if the student can behave as it does is selected. The information gain uses KL divergence as follows.

$$KL(P||Q) = \sum_i P(i)\, log \frac{P(i)}{Q(i)} \qquad (1)$$

where P(i={success, failure}) is the ratio at which a given player has reached state i in the past from a given state and Q is the same ratio of the current student.
Figure 3 shows an example of this step of the selection of the teacher player, where 5 trial data each of players A and B are found in the database near the student's trials. In that case, therefore PA (i = success) = 3/5, PB (i = success) = 2/5, where the probability of A as the teacher of the student is high.

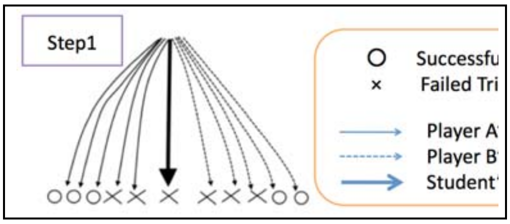


Fig.3 Step1 of Respawn algorithm

3.2. *Step2: Selection of Teacher's Successful Record*

In the step 2 firstly calculates the core point C. The core point is the intersection closest to the state of the student at the end of the play in the set of successful trials of the teacher player. Fig.4 shows an example.

Suppose now that the n-th play failed and ended. Here, it is assumed that C1 and C2 exist during the teacher's trials near the point E (n) at the end of the student's play. In this case C2 is the core point C(n) because the closest intersection C1 is belong to a failure trial.

3.3. *Step3:Learing of Content Recommendation Skills*

In the Step 3 the value of the selected respawn point is updated by using Temporal Difference Learning (TD learning) [2].

$$V\big(R(n)\big) \leftarrow V\big(R(n)\big) \\ +\alpha[r_{n+1} + \gamma V(C(n+1) - V(R(n))] \quad (2)$$

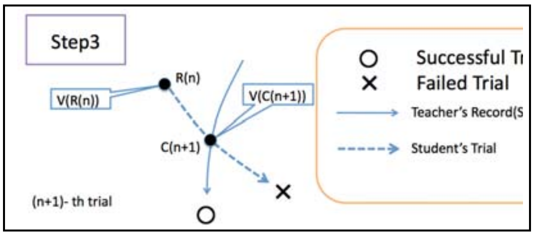


Fig.5 Step3 of Respawn algorithm

In Fig. 5, the value V(·) of the respawn point R (n) is updated at the (n + 1)-th play.
The point itself has the value V. This V will be higher if it contributes to the improvement of students' skills, and it will have a lower value otherwise. By learning this value through trial and error, it can be expected to select a respawn point where learning can be performed more efficiently.

4. Experiments

We actually played a simulator equipped with the proposed respawn algorithm.

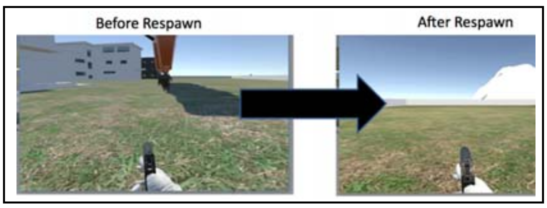


Fig.6 Result of the experiment

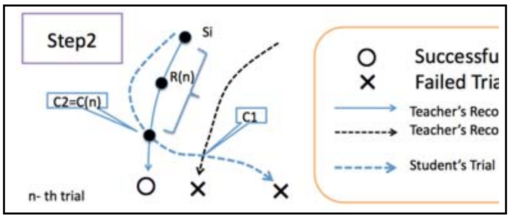


Fig.4 Step2 of Respawn algorithm

Fig. 6 shows the situations: the left figure shows the scene at the end of a trial after colliding with zombies and the right figure illustrates the scene of the respawn point recommended by the proposed algorithm. We confirmed the basic procedure of the algorithm well working.

5. Conclusion

In this paper, we developed a video game like teaching material using virtual space. We proposed a respawn algorithm that supports the efficient learning experience to acquire skills of the player.

References

1. G.N. Yannakakis, J. Togelius, Artificial Intelligence and Games, Springer, 2018.
2. H. Y anagimoto, S. Omatu, Information Filtering Using Kullback-Leibler Divergence,pp1147-1152, IEEJ Trans. EIS Vol.125, No.7, 2005.
3. R. S. Sutton, A. G. Barto, Reinforcement Learning: An Introduction, A Bradford Book,1998.

Visual Classification of Malware by Few-shot Learning

Kien Tran, Masao Kubo, Hiroshi Sato
Department of Computer Science, National Defense Academy of Japan, 1-10-20 Hashirimizu, Yokosuka, Kanagawa 239-8686, Japan
E-mail: ed17005@nda.ac.jp, masaok@nda.ac.jp, hsato@nda.ac.jp
www.mod.gov.jp/nda

Abstract

The extent of damage by malware has been multiplying. Many techniques are proposed for detecting malware. However, the usual pattern matching method does not work because when the new malware appeared, many variants are created very soon. In order to catch the new malware, we have to detect and classify them from very few samples. In this paper, we propose a machine learning mechanism that can learn from very few samples of the image of the malware.
Keywords: Few shot Learning, Malware Classification, Matching Network, and Visualization Classification.

1. Introduction

According to a report from Kaspersky Lab, in 2017, at least 360,000 new malicious files were detected every day in 2017 — an 11.5% increase from the previous year[1]. For example, in May of 2017, a new type of ransomware called WannaCry and its variance spread quickly through a number of computers of companies around the world, encrypted files on the PCs and caused substantial financial damage to these companies.

In the report published by Symantec Corporation in February 2019, a decrease in ransomware activity during 2018 had been observed the first time since 2013, with the overall number of ransomware infections on client sides dropping by 20%. However, ransomware such as WannaCry continued to inflate infection figures, and the number of ransomware infections has been shifted toward enterprises with 81% of all infections[2]. These harmful programs are more devastating than others due to their spreading speeds and its functionalities. This trend leads to a need for methods, which are strong enough to detect and classify these kinds of malware as soon as they start spreading widely.

Nowadays, the automatic defense systems can respond to the malware threats by keeping up with the speed of the malware development, but true success depends on strategic insight as well as the speed of the response. Therefore, the most effective defense requires both intelligent (machine learning-led) programs as well as human expertise. Thanks to the huge development of AI technology, more and more fast and reliable methods have been being developed to detect quickly as well as classify new types of malware.

However, there is not always enough samples for the system to learn to recognize new types of malware as the number of new malware has been increasing every day. This is also one of the most key challenges of machine learning solutions is the number of collected samples. Normally, in machine learning, or deep learning, in particular, the more data we collect, the better the accuracy we get.

In this case, an idea of learning object class from only a few data called one-shot/few-shot learning is widely used. Many one-shot learning algorithms have been proposed to deal with the problems of "data-hungry". Therefore, we adopt few-shot learning approach and try to classify malwares from very few samples.

Some Meta-Learning based concepts are introduced in this paper such as a variation of Neural Turing Machine for one-shot learning tasks introduced by A. Santoro et al.[3], Matching Network by O.Vinyals et al.[4], or Siamese Network by G. Koch et al.[5]. These meta-learning models are capable of well adapting or generalizing to new tasks with unknown data using their learned meta-knowledge during training time.

Generally, most of the few-shot learning algorithms have been shown efficiency in the computer vision area. In this paper, a new approach to malware classification with few-shot learning algorithms is introduced. This approach takes advantage of static analysis in which each malware binary code is treated as a greyscale image, then using some state-of-the-art few-shot learning algorithms to classify them.

2. Related works

There are several method to classify malwares by using image of executable binary image. We introduced some examples here. However, these methods need many samples to train their classifiers. This could lead to data hungry problem.

2.1. *Malware Images: Visualization and Automatic Classification*

Using image processing techniques, L. Nataraj et al.[6] proposed an effective method to classify malware. He represented a malware executable as a binary string of zeros and ones. This vector then is shaped into a matrix and viewed as an image. The results showed significant visual similarities in image texture for each malware belonging to the same class. Compare to other traditional classification methods, this approach does not require either malware disassembly or executions, but still showed significant improvement performance. With these images, the authors use GIST to project them into lower dimensions, k-nearest neighbors with Euclidean distance for classification. Since then, some other researches are also introduced. They followed this idea too and used CNN to deal with the classification tasks. Our approaches are also inspired by the work of this image-based malware analysis.

2.2. *Deep learning at the shallow end: Malware classification for non-domain experts*

This research is another approach that considers a malware file as a gray image. This is similar to the image representation of a raw binary file as the work of L.Nataraji et al.[6], but it is simpler. His conversion method preserves the sequential order of the by code in the binaries. He then applies the Convolutional Neural Network in combination with Bi Long Short Term Memory architect (CNN - Bi-LSTM). His approach is applied to the Microsoft Malware Classification Challenge dataset and achieves very good results.

3. Image-based Unknown Malware Classification with Few-Shot Learning Models

In this section, we introduce a meta-learning model to solve the problems of classifying malware classes with very few known samples.

In this approach, since the purpose is to deal with malware classification problems, real behaviors of malware are not necessary to understand. Hence, via static analysis, the contents of the malware file are quickly scanned and visualized as plain pictures. Then, the few-shot models are adapted to classify them in the malware classification problems. The proposed approach is summarized as illustrating in Fig. 1.

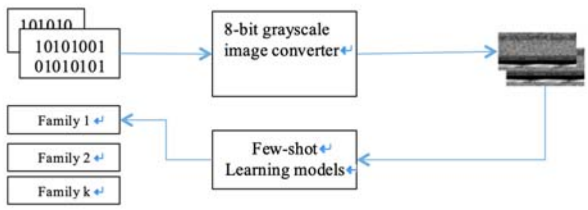


Fig. 1. The proposed approach uses malware binary as an 8-bit grayscale image as input features of few-shot learning tasks.

For Few-shot learning models in Fig.1, we adopt the following two models – Matching Network with external Memory and Weighted Prototypical Class. Both are explained below.

3.1. *Matching Network with external Memory*

To address a challenge of K-shot N-way classification tasks, the proposed models apply embedding learning methods that embed $x \in X \subseteq Rd$ to a smaller embedding space $z \in Z \subseteq Rm$. Using these new spaces, it is easy to identify similar and dissimilar pairs of support samples and test samples. Currently, these methods have three main functions: function *f*(.) embeds sample *xtest* ∈ *Dtest* to *Z*, function *g*(.) embeds *xsupport* ∈ *Dsupport* to *Z* and a similarity measure *s*(.,.) calculates the similarities between the output of f(.) and each output of *g*(.) in the new space *Z*. An overview of this architect is illustrated in Fig. 1.

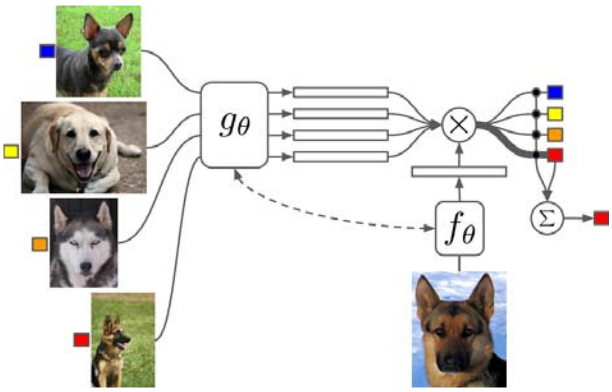


Fig. 2. Illustration of embedding learning methods for few-shot learning classification tasks (excerpt from Ref. 4).

This proposed model draws inspiration from the architecture of the Matching Network model as well as the MANN model for one-shot learning tasks. Thus, its strategy is to enhance an embedding space with memory components, help accordingly recognizing unseen objects based on the content located in these memory matrices.

3.2. *Weighted Prototypical class*

For the N-way K-shot classification tasks, Snell et al.[7] proposed a prototype that computes a representation ck of class k (k=1..N) based on an average calculation of instances of that class.

$$c_k = \frac{1}{K}\sum_{i=1}^{K} g(x_{k,i}) \quad (1)$$

In some cases, the class distribution is skewed. That is, some samples could locate outside the range of major samples in the class. Appling a prototype from eq (1) in such a situation could lead to a biased mean sample of the class. One way to overcome this is to treat those samples unequally based on their weights. These weights, which are used to determine the relative importance of each data point, are considered as the distance of a point to other points in the same class. So, the contributions of the points to the representative point of their class are proportional to the distance between them and the others.

4. Experiments

In this chapter, we will perform 5-way 1-shot tasks and 5-way 4-shot tasks using the approach with two few-shot models (i.e., Matching Networks and Prototypical Networks) on the datasets called MalImg.

We based on meta-learning methods to classify malware. First, we assume some malware classes in MalImg are already known. These samples are trained with few-shot learning models. Then, they are tested with the rest classes which are assumed as never-seen-before classes. We also use the results extracts from the experiment of Kang et al.[8] as the baselines for our comparison.

4.1. *Dataset*

To demonstrate his approach, Nataraj et al.[6] introduced a large dataset of 25 families with more than 9400 malware. The provided samples are stored as greyscale images with different dimensions according to their original file size. The detail of this dataset is summarized in Table 1.

Table 1. MalImg dataset Description

Family	Samples	Family	Samples
Allaple.L	1591	Alueron.gen!J	198
Allaple.A	2949	Malex.gen!J	136
Yuner.A	800	Lolyda.AT	159
Lolyda.AA1	213	Adialer.C	122
Lolyda.AA2	184	Wintrim.BX	97
Lolyda.AA3	123	Dialplatform.B	177
C2Lop.P	146	Dontovo.A	162
C2Lop.gen!g	200	Obfuscator.AD	142
Instantaccess	431	Agent.FY!	116
Swizzot.gen!!	132	Autorun.K	106
Swizzor.gen!E	128	Rbot!gen	158
VB.AT	408	Skintrim.N	80
Fakerean	381		

Regarding malware families, malware authors usually develop new malware based on their previous codes. Only small parts of the old malware are rewritten or removed. By visualizing all parts of the binary file as an image, the analysts themselves by empirical observation could easily recognize the differences between the malware belonging to different classes.

To visualize a malware as an image, we transform a given malware binary file as vectors of 8-bits unsigned integer. Each vector represents a pixel value of the target greyscale image, which is in a range of 0 and 255. Finally, based on the malware file size indicated in the works of Nataraj et al.[6], the resolution of the image is decided. Hence, if a small portion of code is changed, the overall patterns of the malware families may not be affected.

Fig. 3 illustrates the images of specific families of malware. As can be seen from Fig. 3, various malware families have distinct visual characteristics.

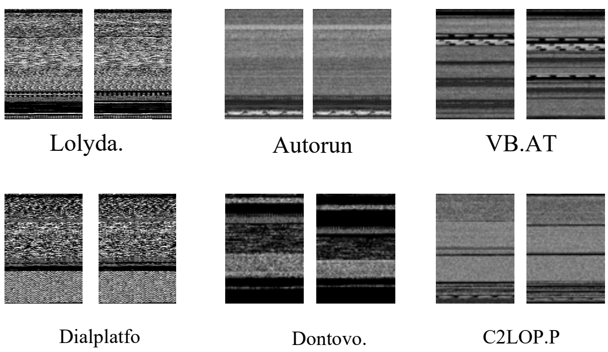


Fig. 3. Some visualization examples extracted from various families. These images from the MalImg dataset indicate that the difference between classes could be easily distinguished by observation.

4.2. *Setting*

In these experiments, since we want to simulate the situation in which the models could classify malware into certain families with a few knowledge of them, The dataset is split into three parts, one part is used for training the model with 11 families, other 7 families are for the validating, and the rests are for testing model. Among available classes, we randomly select five classes then perform 5-way classification tasks.

Two experiments are implemented. The first scenario is only one sample per class provided. The models have to guess the class of the test sample. This scenario is called a 5-way 1-shot task. Another experiment is a 5-way 4-shot task in which instead of using one sample per class, four random samples are picked out of five random classes. For every malware inputs belonging to those five classes, the models have to guess their family. These experiments are implemented with both Matching Networks and Prototypical Networks.

To begin the experiments, we resize all samples into an 84x84 scale despite their varied sizes. These samples are extracted the necessary information and embedded into feature space via a simple yet powerful Convolution Neural Network (CNN). This CNN consists of four stacked blocks of {3×3-convolutional layer with 64 filters, batch-normalization, 2×2 max-pooling, leaky-relu, drop out layer with rate 0.3}. The output is passed through a fully connected layer resulting in a 64-dimensional embedding output.

The 64-dimensional vectors are then fed to function $f(.)$ and function $g(.)$ in the Matching Networks (MNets) model to compute the Cosine distance between the support samples and test sample.

In the case of the Prototypical Networks (ProtoNets), the prototype of each supported class is computed as an average of all 64-dimensional vectors of that support class. Those calculated prototypes are used to classify the test samples based on the L2 Euclidean distance between them and the test samples.

4.3. *Result*

The experiment results reported in Table 2 are the averages of 1,000 times of testing. The baselines are extracted from the works of Kang et al.[8].

To classify malware to certain families based on one or a few known samples, we use the visual similarity of malware images with some few-shot models, which have been well studied. The two few-shot learning models used in this approach are Matching Networks[4] and its extended variation, Prototypical Networks[7].

It is seen that Matching Networks and Prototypical Networks overcome the model using Memory Augmented Neural Network and the other baselines in both 1-shot and 4-shot tasks. Moreover, the differences between these results are very high. While the MANN could perform only 66.2% accuracy on 1-shot task and 79.4% on 4-shot task, both the MNets and the ProtoNets are all reaching over 86% and 89% respectively. Especially, with the ProtoNets, even only one sample is provided, the model could determine the correct class among five classes of test samples with 92.4% accuracy. If four samples of each class are known, the model is superb with 95.3% certainty.

Table 2. MalImg dataset's Classification Results

Model	1 shot	4 shots
FeedForward Network	37.9%	30.0%
CNN	42.6%	50.6%
LSTM	56.2%	64.1%
MANN	66.2%	79.4%
Matching Networks	86.3%	89.7%
Prototypical network	**92.4%**	**95.3%**

5. Conclusion

This paper proposed another malware classification approach that takes advantage of the developments of few-shot learning algorithms to introduce a novel way of malware classification. It helps malware analysts to quickly classify malware into the correct groups even with only a few known samples. In this approach, the effectiveness of classifying malware based on Image Processing in combination with other few-shot learning models (Matching Networks and Prototypical Networks) has been proven. The accuracies of classifications are outstanding, even with only one provided sample. Furthermore, the accuracies of the classifications could be improved by adjusting the hyper-parameters of the embedding network (CNN) as well as Image Processing procedures.

As future work, it is necessary to dig deeper into this method to reduce the effect of noise samples, and improve final results as well as compare them to existing methods in case of few-shot learning tasks. We will also take a more in-depth look at some other one-shot learning algorithms or other simpler methods such as NCC (Normalized Cross-Correlation) to find more suitable methods for malware analysis to improve the accuracies of our approach. More benign programs will also be collected along with different kinds of malware; hence, we could re-evaluate better our methods.

References

1. Kaspersky, "Kaspersky Lab Report", https://www.kaspersky.com/about/press-release/2017kaspersky-lab-detects-360000-new-malicious-files-daily, accessed Nov 10, 2019.
2. Department of Health and Care, Securing cyber resilience in health and care Progress update October 2018, Oct 2018, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/747464/securing-cyber-resilience-in-health-and-care-september-2018-update.pdf, accessed Dec 12, 2019.
3. A. Santoro, S. Bartunov, M. Botvinick, D. Wierstra, and T. Lillicrap. "Meta-learning with Memory Augmented Neural Networks." In Proceedings of The 33rd International Conference on Machine Learning, pp. 1842–1850, vol. 48, 2016. R. Loren and D. B. Benson, Deterministic flow-chart interpretations, *J. Comput. System Sci.* **27**(2), 1983, pp. 400–433.
4. O. Vinyals, C. Blundell, T. Lillicrap, k. Kavukcuoglu, and D. Wierstra. "Matching Networks for one-shot learning," In Advances in Neural Information Processing Systems, pp. 3630– 3638, vol. 29, 2016.
5. G. Koch, R. Zemel, and R. Salakhutdinov, "Siamese neural networks for one-shot image recognition," In ICML 2015 Deep Learning Workshop, 2015.
6. Zhenguo Li and Fengwei Zhou and Fei Chen and Huang Li, "Meta-SGD: Learning to Learn Quickly for Few-Shot Learning", ArXiv, abs/1707.09835, 2017
7. J. Snell, K. Swersky, and R. Zemel."Prototypical networks for few-shot learning," In Advances in Neural Information Processing Systems, pp. 4077–4087, vol. 30, 2017.
8. Min Chul Kang, Huy Kang Kim, "Rare Malware Classification Using Memory Augmented Neural Networks," (Korean) Journal of The Korea Institute of Information Security & Cryptology, vol. 28, No.4. pp. 847–857, Aug 2018.

Detecting Pedestrians and Their Moving Directions by a MY VISION System

Kenta Hori *
Graduate School of Engineering, Kyushu Institute of Technology
Kitakyushu, Fukuoka 804-8550, Japan†

Seiji Ishikawa

Joo Kooi Tan
Faculty of Engineering, Kyushu Institute of Technology
Kitakyushu, Fukuoka 804-8550, Japan
E-mail: etheltan@cntl.kyutech.ac.jp

Abstract

This paper proposes a pedestrian detection method using a MY VISION system, which is an image processing system employing an ego-camera. First, a flow area different from the camera movement is extracted on the images from a user's head-mounted camera. The area is a candidate area where pedestrians may exist. Second, MSC-HOG features are calculated to detect pedestrians. HOF feature is then computed to recognize the moving direction of the detected pedestrians. Experimental results show effectiveness of the proposed method.

Keywords: Optical Flow, MSC-HOG, Pedestrian Recognition, Direction Recognition.

1. Introduction

When a visually impaired person goes out, there is a risk of colliding with obstacles, in particular, moving pedestrian. Most of the conventional pedestrian detection methods are a combination of Histograms of Oriented Gradients (HOG) and Support Vector Machine (SVM) [1] or Light Detection and Ranging (LiDAR) which can measure the distance to an object with high accuracy. However, these methods cannot prevent collision because they only detect pedestrians or are used for vehicles.

This paper proposes a method of detecting a moving pedestrian around a person particularly a visually impaired person who uses a MY VISION system. MY VISION is the image provided from a camera mounted on the head of a user. A MY VISION system analyzes MY VISION to obtain certain useful visual information. The system can therefore function as a substitute for a user's eyes and a brain. Proposed method is introduced in Section 2. In section3,4,5, experiments, discussion and conclusion are respectively introduced. The proposed method is examined its effectiveness experimentally.

2. Proposed Method

2.1. *Outline of the proposed method*

The outline of the proposed method is described in the following. Using MY VISION video, the location is extracted where the optical flow is different from the camera movement. The area is regarded as a pedestrian candidate area, and features of the area are extracted. Thereafter, the feature of the region is applied to a pedestrian classifier to recognize whether or not there is a pedestrian, and if so, his/her traveling moving direction is also recognized. For feature extraction, Multiple Scale Cell-Histograms of Oriented Gradients (MSC-HOG) [2]

---

*Sensui-cho 1-1, Tobata-ku, Kitakyushu, Fukuoka 804-8550, Japan

features and Histograms of Optical Flow (HOF)[3] features using Total Variation L1 (TV-L1) Optical Flow[4] are employed. A classifier is defined by Random Forest.

2.2. *MSC-HOG feature*

MSC-HOG features expanded from HOG features are used as the features for pedestrian detection. In the MSC-HOG feature, the position and size of the cell are variable, and the cells are arranged exhaustively along the outline of the person obtained from a human model. In addition, since the cell size is not constant, cells that match the characteristics of the human body shape, e.g., long in a specific direction, can be placed, so that more effective features can be obtained. **Fig. 1** shows an example of the cells that are arranged exhaustively for person detection using MSC-HOG features.

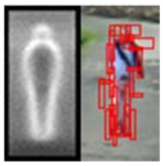


Fig. 1. Cells and their distribution in MSC-HOG features: (a) A human model, (b) an example of cell layout (red rectangle).

2.2.1. *Human model*

A large number of person images are collected in order to create a human model. A brightness gradient image is created for the collected images, and an average of all the brightness gradient images is calculated to create a human model.

2.2.2. *Calculation of the feature*

(i). *Calculation of luminance gradient direction and intensity*

The gradient vector is calculated for each pixel in the placed cell. The gradient strength $m(x,y)$ and gradient direction $\theta(x,y)$ are calculated as follows;

$$m(x,y) = \sqrt{f_x(x,y)^2 + f_y(x,y)^2} \tag{1a}$$

$$\theta(x,y) = \tan^{-1}\frac{f_y(x,y)}{f_x(x,y)} \quad (0° \le \theta(x,y) < 180°) \tag{1b}$$

$$f_x(x,y) = f(x+1,y) - f(x,y) \tag{1c}$$

$$f_y(x,y) = f(x,y+1) - f(x,y) \tag{1d}$$

Here, $f(x,y)$ is the luminance value of the input image, and $f_x(x,y)$ and $f_y(x,y)$ are the first derivative values of $f(x,y)$ in the x and y directions, respectively.

(ii). *Making a gradient histogram*

Using the gradient strength calculated by Eq.(1a), a gradient histogram is created with each cell. The gradient direction $(0° \le \theta(x,y) < 180°)$ is separated into 9 directions by 20 °, and the gradient vector of each pixel are assigned to the 9 directions.
The gradient strength of each pixel in a cell is added with each direction, and a histogram showing the frequency in each direction is created. The histogram value $hist(j)$ in each direction in the cell is calculated by

$$hist(j) = \sum_{(x,y)\in cell_i \wedge dir(x,y)=j} m(x,y) \tag{2}$$

Here, $cell_i$ represents the i-th cell.

(iii). *Making MSC-HOG features*

Using Eq. (3), a histogram of each cell $cell_i (i = 1,2,\dots,N)$ is calculated. For each cell, a 9-dimensional feature vector $c_i{}'(i = 1,2,\dots,9)$ is defined by Eq. (3). Vector $c_i{}'$ is normalized using Eq. (4) to obtain c_i whose L2 norm is 1. Finally, the feature vectors of all the cells are combined and an MSC-HOG feature vector $\boldsymbol{x}$ is created by Eq. (5).

$$\boldsymbol{c}_i' = (h_{i,1}, h_{i,2}, \dots, h_{i,9}) \tag{3}$$

$$\boldsymbol{c}_i = \frac{\boldsymbol{c}_i'}{\|\boldsymbol{c}_i'\|_2} \tag{4}$$

$$\boldsymbol{x} = (\boldsymbol{c}_1, \boldsymbol{c}_2, \dots, \boldsymbol{c}_n) \tag{5}$$

Here, $h_{i,j}$ is the value of $hist(j)$ in the i-th cell and n is the total number of cells in the detection window.

2.3. *Histograms of Optical Flow*

HOF feature is employed for recognizing a pedestrian's motion direction. HOF is a histogram created by dividing the local area into blocks and describes the flow of two successive frames with each block, based on the direction and intensity of the flow.
In this research, we use TV-L1 Optical Flow of dense optical flow. The constraint condition of TV-L1 Optical Flow is shown in Eq. (6). The first term allows flow discontinuity, and the second term shows the difference in luminance value between frames. The calculation is repeated so that the sum of these terms is minimized.

$$\min_{\boldsymbol{u}} \left\{ \int_{\Omega} |\nabla \boldsymbol{u}| dx + \lambda \int_{\Omega} \left| I_1(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x})) - I_0(\boldsymbol{x}) \right| dx \right\} \tag{6}$$

Here, I_0 and I_1 are two successive images, $\boldsymbol{u}(\boldsymbol{x})$ is a change amount at coordinates$\boldsymbol{x} = (x, y)$, λ is a weight, and$\Omega \to \mathbb{R}^2$. The visualized TV-L1 Optical Flow is shown in **Fig. 2**.

2.3.1. *Calculation of the HOF feature*

(i). *Making a Histogram*

A HOF is created for each cell calculated in section 2.3. The flow direction $(0° \le \theta(x, y) < 360°)$ of each pixel $I(x, y)$ in the arranged cell is separated into 18 directions by 20°, and the flow of each pixel in the cell is assigned in 18 directions. The flow allocation direction and the histogram value in each direction in the cell are calculated using Eqs. (2) and (3) as in MSC-HOG.

(ii). *Making HOF feature*

If the feature vector of each cell is $\boldsymbol{a} = \{a_1, a_2, \dots, a_{18}\}$, the feature vector of block $k(k = 1,2, \dots K)$ can be represented by $\boldsymbol{v}'_k = \{a_{k1}, a_{k2}, \dots, a_{k18}\}$. The feature vector $\boldsymbol{v}_k$ is obtained by normalizing $\boldsymbol{v}'_k$ using Eq. (7). Finally, the normalized feature vectors of all the blocks are combined, and the HOF feature vector $\boldsymbol{x}_{flow}$ is created using Eq. (8).

$$\boldsymbol{v}_k = \frac{\boldsymbol{v}'_k}{\|\boldsymbol{v}'_k\|_2} \qquad (7)$$

$$\boldsymbol{x}_{flow} = (\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_K) \qquad (8)$$

Here, Fig. 2(d) shows an example of cell and block arrangement.

2.4. *Extracting a pedestrian candidate area*

Pedestrian candidate areas are extracted using TV-L1 Optical Flow. First, the flow is calculated from two successive frames obtained from the camera. Second, the flow angle and intensity of each pixel are visualized using the HSV color system to create a flow image. Third, the flow image is first-order differentiated in the x direction. This is to separate the pedestrian flow from the camera flow and to generate an edge image of the flow. Finally, Edge images are combined into a single region by morphological closing process. A candidate area is then obtained as shown in **Fig. 3**.

2.5. *Recognition*

For recognition, Random Forest[5] is used as a classifier. In the recognition process, the MSC-HOG feature is calculated on the input image and it is fed to the random forest classifier to judge the existence of a pedestrian. If the judgment is affirmative, the HOF feature is calculated from the input image and its next image. The resulting HOF feature $\boldsymbol{x}_{flow}$ is then fed to the random forest classifier to recognize the pedestrian direction.

(a) (b) (c) (d)

Fig. 2. Visualization of the flows between two successive frames: (a) frame at time t, (b) frame at t+1, (c) visualized flows and (d) example of cell (black rectangle) and block (red rectangle) arrangement.

(a) (b) (c) (d) (e)

Fig. 3. Finding a candidate area: (a) frame at time t, (b) frame at t + 1, (c) flow image, (d) flow edge image, (e) candidate area.

3. Experimental Results

3.1 *Experimental setup*

In order to confirm the effectiveness of the proposed method, we conduct two kinds of experiments: The first experiment is a pedestrian recognition experiment using test data from INRIA Person Dataset[6] (1126 positive and 3000 negative static images). This is a simple experiment: MSC-HOG features are calculated from still images of test data to determine whether a person or the background. INRIA Person Dataset (2416 positive and 5000 negative static Images) and our own collected dataset (7378 positive static images) are employed to create a classifier for pedestrian recognition. MSC-HOG features are calculated using this data set to create the classifier.

The direction recognition classifier uses two consecutive frames data set (front left: 618 [309 image pairs], front: 544 [272 pairs], front right: 618 [309 pairs]). HOF features in each direction are calculated from this data set to create the direction recognition classifier.

3.2. *Evaluation method*

Recall, *Precision* and *F-measure* are used for evaluation of pedestrian detection. On the other hand, *Accuracy* is used for the evaluation of a pedestrian's moving direction recognition. The formulas are shown below;

$$Recall = \frac{TP}{TP + FN} \times 100[\%] \qquad (9)$$

$$Precision = \frac{TP}{TP + FP} \times 100[\%] \qquad (10)$$

$$F - measure = \frac{2 \times Recall \times Precision}{Recall + Precision} \qquad (11)$$

$$Accuracy = \frac{N_{direction}}{N} \qquad (12)$$

TP is the number of frames in which the pedestrian area was identified as the Positive class, *FN* is the number of frames in which the pedestrian area was identified as the Negative class, and *FP* is the number of frames in which the non-pedestrian area was identified wrongly as the Positive class. $N_{direction}$ is the number of frames with which the direction of the pedestrian was correctly recognized. *N* is the number of frames used for pedestrian recognition.

3.3. *Results*

Table 1 shows the results of pedestrian recognition experiment using the test data of INRIA Person Dataset, and **Table 2** shows the result of pedestrian detection and direction recognition experiment for the three videos. **Table 3** shows the breakdown (in %) of all the frames in the videos with which pedestrian direction recognition was done. Values of *Recall*, *Precision* and *F-measure* were 94.05, 75.97 and 84.05, respectively, in experiment 1, On the other hand, in experiment 2, the average values of *Recall*, *Precision*, *F-measure* and *Accuracy* were 76.28, 72.64, 74.15 and 68.72, respectively.

4. Discussion

From Table 2, we can see that the recall of scene 2 and scene 3 is lower than that of scene 1. This is probably because the candidate area including the pedestrian could not be obtained well due to many frames with small pedestrian flow in scene 2 and scene 3. We consider that it can be improved by changing the frame interval for creating flow. Some examples of the visualized flow of each scene are given in **Fig. 4**.
Table 3 tells that the recognition rate of front is the lowest. This is thought that, because the front flow is similar to both the front right and the front left, the false recognition rate increases. We believe that the recognition rate can be increased by adding a feature on shape (for example, LBP) to the direction recognition feature.

5. Conclusion

In this paper, we proposed a pedestrian detection method including his/her moving direction information from MY VISION. The MSC-HOG feature and the HOF feature by TV-L1 Optical Flow were introduced, and pedestrian recognition experiments using a standard dataset and three videos were performed to confirm the effectiveness of the proposed method. Considering the feature on shape may as well contribute to better performance of the proposed method.

Acknowledgement
This work was supported by JSPS KAKENHI Grant Number 25350477.

6. Reference

1. Dalal, Navneet, and Bill Triggs. "Histograms of oriented gradients for human detection." 2005.
2. Ehara, Y., Tan, J. K., & Ishikawa, H. K. S.," Human Detection from an Image Employing Reduced Number of Weak-Classifiers", Proceedings of the IEEE Image Electronics and Visual Computing Workshop,2012
3. Dalal, Navneet, Bill Triggs, and Cordelia Schmid. "Human detection using oriented histograms of flow and appearance." European conference on computer vision. Springer, Berlin, Heidelberg, 2006.
4. Pérez, Javier Sánchez, Enric Meinhardt-Llopis, and Gabriele Facciolo. "TV-L1 optical flow estimation", *Image Processing On Line*, pp.137-150, 2013.
5. Breiman, Leo. "Random forests", Machine learning 45.1 (2001): 5-32.
6. http://pascal.inrialpes.fr/data/human/

Table 1. Confusion matrix of experiment 1.

		Predicted class	
		Positive	Negative
Actual class	Positive	1059	67
	Negative	335	2665

Table 2. Evaluation of experiment 2.

	Recall	*Precision*	*F-measure*	*Accuracy*
Scene1	93.75	78.95	85.71	80.00
Scene2	68.42	76.47	72.22	46.15
Scene3	66.67	62.50	64.52	80.00
Average	76.28	72.64	74.15	68.72

Table 3. Confusion matrix of experiment 2 [%].

		Predicted class		
		Left Front	Front	Right Front
Actual class	Left Front	80	13.33	6.67
	Front	23.08	46.15	30.77
	Right Front	10	10	80

(a) (b) (c)

Fig. 4. Examples of visualized flow images (two in each direction): (a) left front, (b) front and (c) right front

Human Motion Recognition Using TMRIS

Cao Jing
Graduate School of Engineering, Kyushu Institute of Technology
Kitakyushu, Japan

Youtaro Yamashita
Graduate School of Engineering, Kyushu Institute of Technology
Kitakyushu, Japan

Joo Kooi Tan
Faculty of Engineering, Tan Laboratory, Kyushu Institute of Technology
Kitakyushu, Japan
E-mail: etheltan@cntl.kyutech.ac.jp

Abstract

With the aggravation of the aging trend in Japan, the number of elderly people living alone has gradually increased, and the development of elderly people care system has begun to receive attention. Considering this trend, the paper proposes a MHI(Motion History Image)-based method called TMRIs (Triplet Motion Representation Images) that solve in recognition, a self-occlusive motion problem particularly in the depth direction using a single camera. The performance and effectiveness of the method are verified by experiments.

Keywords: Elderly care; focus of expansion; MHI; motion recognition; Triplet Motion Representation Images.

1. Introduction

In modern society, Japan's aging trend is getting more and more serious. The increase in the number of elderly people living alone and 65 years of age or older is remarkable for both men and women. .In 1980, approximately 190,000 men and about 690,000 women among total population of the elderly, 4.3% of men and 11.2% of women, respectively, were living alone. On the other hand, about 1.92 million men and about 4 million women among the total population of the elderly, 13.3% for men and 21.1% for women, respectively, live alone in 2015.[1] Therefore, it is necessary to establish an elderly support/care system. According to Ministry of Health,[2] "fracture and fall" is the fourth leading cause of the need for protection for the elderly.

Nowadays, the Japanese elderly care system is gradually improving. There are many care systems in the current market, for example, specialists regularly visit elderly homes to see his/her health status: they use contact sensors to monitor the usages of home appliances or even gas in the home to know the activities of a person in the house: and 24-hour monitoring systems are also popular. In this paper, aiming at the development of an elderly care system which employs an intelligent robot, we propose a novel human motion description and recognition method employing a single camera. In conventional researches, Andrade et al.[3] proposed flow vectors for establishing optical flow distributions in regions of interest. Bobick, Davis[4] proposed MHI to describe historical motion using a single camera. Michelson, Hilton[5] proposed three-dimensional restoration of motion. However, most of the methods of motion recognition using a single camera, such as MHI assume that the motion for recognition is on the plane perpendicular to the optical axis of the camera, and the motion toward the camera or leaving from the camera is not dealt with because of its self-occlusive nature. To describe a self-occlusive motion, Tan et al.[6] proposed a reverse MHI method. In this paper, we propose a method which can deal with the motions to and from a camera by expanding the MHI. The method is named TMRIs (Triplet Motion Representation Images) which represents depth information, i.e. approaching or leaving from a camera, by three characteristic images. The performance of the proposed method is shown experimentally.

2. Proposed Method

The proposed method is mainly separated into four steps. The first step is a foreground extraction. Since a background may change over time, an extraction of human region is performed using a background model that can cope with changes in the background. In the second step, the optical flow is used to represent the human movement. If the lines representing optical flows are finally concentrated at a point, it is considered to find a FoE (Focus of Expansion), indicating that the motion contains the movement toward an observing camera or away from the camera. If no FoE is found, the motion is considered to be in a plane perpendicular to the optical axis of the camera. The third step is to extend the traditional MHI method into TMRIs to represent a human motion. In the final step, Hu moments are calculated as the feature vector of the TMRIs, which is used for learning and recognition of motions.

2.1. *Foreground extraction*

In order to extract a human region from the temporal videos, successive background image inference is performed using a GMM (Gaussian mixture model)[7] for each pixel constituting the image. The influence of the external environment such as the light mutation in the background can be effectively ignored by the background image inference. In this study, the EM algorithm is used to determine the means and variances of the initial parameters of GMMs in the background video that does not contain a moving object.

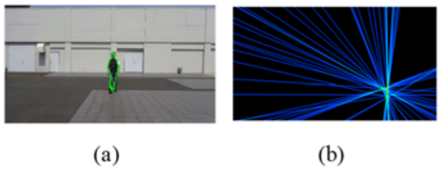


Fig. 1. Extension of optical flow: (a) Optical flows of a human motion, (b) FoE on a voting plane.

2.2. *Computation of a FoE*

In order to keep the detection of the Focus of Expansion (FoE) stable, feature points are equally spaced on the outline of the extracted human region. In this study, LK tracker[8] is used to calculate the optical flow between successive frames. Fig.1(a) shows an example of optical flows of human movement. The optical flows are extended on the voting plane to find a FoE, as shown in Fig.1(b). If the number of lines intersecting a point exceeds a predetermined threshold, the point is considered as a FoE.

Because of the possible wrong detection of a FoE due to the deviation of the voting, this paper adopts a weighted voting method. It gives weight to the extended optical flow line. A larger weight is given to the pixels close to the line, whereas smaller weight is given to the pixels distant from it. A FoE exists if the result of the final vote is greater than a predetermined threshold.

2.3. *Description of motion: TMRIs*

In this paper, we extend the original 2D MHI, so that it can describe a motion in a 3D way. To realize this, we propose a motion description method using three kinds of motion history images called TMRIs. They are newness, indicating the original 2D MHI, density, indicating the frequency of appearance in the past τ frames, and depth, showing the movement of the object in the depth direction. The specific mathematical definition of each image is as follows;

$$H_\tau^{new}(x,y,t) = \begin{cases} \tau & if\ D(x,y,t)=1 \\ \max(0, H_\tau^{new}(x,y,t-1)-1) & otherwise \end{cases}. \tag{1}$$

$$H_\tau^{den}(x,y,t) = \sum_{i=0}^{\tau} D(x,y,t-i). \tag{2}$$

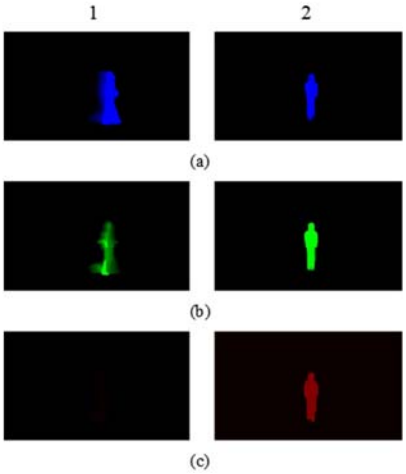


Fig. 2. Examples of TMRIs: (a) newness, (b) density, (c) depth: 1– Walk to the right, 2– Walk to the back.

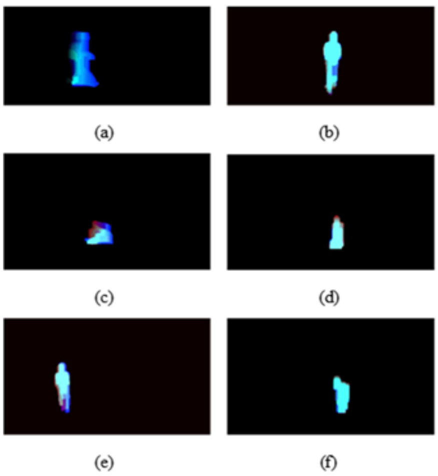


Fig. 3. Examples of motions described by TMRIs: (a) walk right, (b) walk back, (c) fall right, (d) fall back, (e) walk right front, (f) crouch.

$$H_{\tau}^{dep}(x,y,t)=\sum_{i=0}^{\tau}\left\{N_{layer}(t-i)\times D(x,y,t-i)\right\}. \quad (3)$$

Here $H_{\tau}(x,y,t)$ is the gray value at (x,y,t) and D is a binary image showing foreground region. Among them,

$$N_{layer}=\begin{cases}\gamma\times L_{ave} & if\, V_{\max}>T_{vote}\\ 1 & otherwise\end{cases}. \quad (4)$$

Here $V_{\max}$ is the maximum value with voting results: T_{vote} is the voting threshold: $N_{layer}(t)$ is the number of layers to be overlapped at time t : γ is a constant used to determine the value of N_{layer} according to the size of optical flows: L_{ave} is the average length of optical flows.

The TMRIs w.r.t. walking are shown in the Fig.2.

2.4. *Feature extraction*

Since Hu moments are invariant to image scaling, rotation and flipping, the feature of TMRIs are described using the method. Hu moments are defined by seven invariant features v_i (i =1,2,…,7). Since they have large difference in scale, they are changed to logarithmic values by $s_i=\log(sign(v_i)\cdot v_i)$. The feature vector is then defined as $\boldsymbol{s}=(s_1,s_2,\cdots,s_7)$.
Since TMRIs contain three motion history images and each image is described by Hu moments, they are integrated to form a 21-dimensional vector as

$$V^{TMRI}=(\boldsymbol{s}^{new},\boldsymbol{s}^{den},\boldsymbol{s}^{dep}). \quad (5)$$

2.5. *Motion recognition*

Motion recognition is performed by the k-nearest neighbor method according to the 21-dimensional Hu moments vector obtained in Section 2.4, as shown in Eq.(6). The class of motion l is denoted by C^l, the jth learning data in C^l by v_j^l, and the input unknown motion is by v.

$$l^*=maj_l\left\{arg_{v_j^l}\,k\min\left\{I(\boldsymbol{v},\boldsymbol{v}_j^l)\forall l,\forall j,\boldsymbol{v}_j^l\in C^l\right\}\right\}. \quad (6)$$

Here l^* represents the recognized class: $maj_l\{S\}$ returns the class that appears the most in set S: $k\min\{T\}$ is the k minimum numbers in the set T: I is the dissimilarity defined by $I(\boldsymbol{v},\boldsymbol{w})=\|\boldsymbol{v}-\boldsymbol{w}\|$.

3. Evaluation

In order to verify the performance of the proposed method, it is necessary to evaluate it by experiments. Experiments were carried out on the detection of FoE and the accuracy of motion recognition. The experimental video contains 13 different actions: walk left/right/front/back, walk left-front/right-front, walk left-rear/right-rear, fall left/right/front/back, and crouch. Each of the motions is done once by four people (students of our lab).

3.1. *Detection of a FoE*

When the maximum count in the voting result exceeds a threshold (=20 in this experiment), a FoE is detected. The detection accuracy is given by the following equation;

$$Accuracy=\frac{F_{TP}+F_{TN}}{F_{ALL}}\times 100[\%]. \quad (7)$$

Here F_{ALL} is the total number of frames for which detection of a FoE was performed: F_{TP} is the number of frames in which the FoE is detected correctly for motion in depth direction: F_{TN} is the number of frames for which the FoE is not detected correctly for other motions. The detection accuracy of FoEs with all the actions is 64.3% in average.

3.2. *Motion recognition*

As shown in Fig.3, the motions are described by Eqs.(1)-(4) which superimpose the three images (newness, density and depth) into a single represented by R, G and B. Each image is also described in the form of a feature vector by Eq.(5).
Each of the 4 individuals performed the 13 motions, and 80 TMRIs were yielded from each video, totaling 52 videos and 4160 TMRIs. (Note that one motion produces one video.) Leave-one-out cross validation was performed to evaluate the accuracy of motion recognition. With respect to the k-nearest neighbor algorithm, *k* is set to 3. The obtained motion recognition rate is shown in Table 1. For comparison, motion recognition using the traditional MHI[4] and the same motion set was done. The result is given in Table 1.

4. Conclusions

With most of the researches in human motion recognition, the dealt motions are almost perpendicular to a camera's optical axis. TMRIs proposed in this paper can describe the motion toward the depth (optical axis) direction under a single camera. An arbitrary motion is described first by the TMRIs, then they are transformed into 21-dimensional Hu moments vector and the unknown motion is recognized in the 21-D feature space employing 3-nearest neighbor method. The accuracy of motion recognition by TMRIs is verified by experiments and satisfactory results are obtained.
Since the TMRIs have two other kinds of information about density and depth on the basis of MHI, the recognition rate of each action is significantly higher than the traditional MHI. The average recognition rate of TMRIs is much higher than that of MHI by nearly 29%.
Although the recognition accuracy is acceptable, the proposed method requires further improvements to make it have higher and more stable performance.

Table 1. Recognition rate of motions.

Motion	Recognition rate [%]	
	MHI	TMRIs
Walk left	51.25	88.75
Walk right	39.69	74.69
Walk front	68.44	95.63
Walk back	70.63	94.38
Walk left front	34.06	72.50
Walk right front	50.31	82.81
Walk left rear	33.13	74.38
Walk right rear	53.75	80.31
Fall left	42.50	74.38
Fall right	46.25	84.38
Fall front	76..25	90.31
Fall rear	67.19	90.00
Crouch	87.19	94.06
Average	55.43	84.35

5. Acknowledgements

This study was supported by JSPS KAKENHI Grant Number 16K01554.

References

1. Cabinet Office: 2017 White Paper on Aging Society (Overall version), 2017, pp.2-6. (in Japanese)
2. Ministry of Health, Labor and Welfare: 2016 National Life Basic Survey, 2016. (in Japanese)
3. E. Andrade, R. Fisher, S. Blunsden, Detection of emergency events in crowded scenes, *Proceedings of IEEE International Symposium on Imaging for Crime Detection and Prevention*, 2006, pp. 528-533.
4. A. Bobick, J. Davis, The recognition of human movement using temporal templates, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(3), 2001, pp. 257-267.
5. J. Michelson, A. Hilton, Simultaneous pose estimation of multiple people using multiple-view cues with hierarchical sampling, *Proceedings of British Machine Vision Conference*, 2003, pp. 1-10.
6. J. K. Tan, S. Okae, Y. Yamashita, Y. Ono, A method of describing a self-occlusive motion – A reverse Motion History Image, *International Journal of Biomedical Soft Computing and Human Sciences*, 24(1), 2019, pp. 1-7.
7. C. Stauffer, W. E. L. Grimson, Adaptive background mixture models for real-time tracking, *Proceedings of Conference on Computer Vision and Pattern Recognition*, 2, 1999, pp. 246-252.
8. B. D. Lucas and T. Kanade, An Iterative Image Registration Technique with an Application to Stereo Vision, in *Proc. of Int. Joint Conf. on Artificial Intelligence*, 1981, pp. 674-679.

Automatic Extraction of Abnormalities on Temporal CT Subtraction Images Using Sparse Coding and 3D-CNN

Yuichiro Koizumi, Noriaki Miyake, Huimin Lu, Hyoungseop Kim
Department of Mechanical and Control Engineering, Kyushu Institute of Technology, 1-1 Sensui-cho, Tobata-ku, Kitakyushu-shi, Fukuoka, 804-8550, Japan

Takatoshi Aoki
University of Occupational and Environmental Health, 1-1, Iseigaoka, Yahatanishi, Kitakyusyu-shi 807-8555, Japan

Shoji Kido
Osaka University, 2-2, Suita, Osaka 565-0871, Japan

Abstract

In recent years, the rate of death from cancer has tended to increase in Japan. Also, there is a concern that increasing the performance of CT will increase the burden on doctors. Therefore, by presenting a "second opinion" in the CAD system, the burden on doctors can be reduced. We develop a CAD system for automatic detection of lung cancer. In this paper we propose a method to detect abnormalities based on temporal subtraction technique, sparse coding and 3D-CNN. We obtain the result that sparse level contributed most to the score.

Keywords: Temporal Subtraction Technique, Sparse Coding, Convolutional Neural Network

1. Introduction

In Japan, the proportion of deaths from cancer tends to increase, especially the number of deaths from lung cancer is also increasing. In the number of cancer deaths by site in 2017, lung cancer is ranked first in men and second in women, and is considered to be a serious disease [1]. In addition, early detection and early treatment are required since the cancer progresses quickly. To detect abnormalities, CT device are mainly used for precise examination of lung cancer. Because the CT device can acquire detailed information in the body, improvement of the accuracy of the diagnosis is expected. However, as the performance of the CT device is improved, the frequency of CT scans increases and the number of images also increase. This increases the burden on doctors who interpret CT images. Therefore, concerns are raised such as overlooking the lesion due to physician fatigue and overlooking by an inexperienced doctor.

In recent years, computer aided diagnosis (CAD) system has attracted attention in order to reduce burden on interpreting doctors. The CAD system is a system that performs image processing and image analysis using a computer and presents the result to a doctor as a "second opinion". By using the CAD system, improvement of interpreting speed and image interpreting accuracy is expected, since it is possible to reduce the burden on interpreting doctor as well as variation in diagnostic accuracy.

As one of the CAD systems, there is a temporal subtraction technique. The technique can emphasize only a lesion part that changes with time by performing difference calculation processing between the two images of a same patient [2]. The temporal subtraction image in the chest X ray image is considered effective, and in particular, it was recognized that the correct examination rate was improved, so it is considered more effective for the doctor with less experience [3].

There is a technique, which is called sparse coding. This is a method of reconstructing an image by representing a given image with a linear sum of a relatively small number of bases. This will eliminate useless information. Sparse representation is applied to many vision tasks such as face recognition, image super-resolution, image segmentation, image classification and so on [4]. In addition, as the field of artificial intelligence, deep learning has been introduced effective tool. This is a generic name of a multilayer of machine learning. High identification rate was recorded in 2012 general image recognition contest "ImageNet large scale visual recognition challenge (ILSVRC)" [5]. These techniques are also often used for CAD.

CAD methods for lung cancer diagnosis have been developed [6-8]. Teramoto et al. [6] proposed a method for automated detection of nodule in PET/CT images. They reduce false positives by using CNN. Seito et al. [7] proposed CAD systems for nodule using multi-view convolutional networks (ConvNets). They used architecture that comprises multiple streams of 2D ConvNets. Tanaka et al. [8] performed nodule detection using sparse coding and SVM. They classified the final candidate nodule using SVM method based on coefficient matrix, which are obtained by the sparse coding.

Many methods analyze 2D images, but few methods use 3D images. In this paper, we propose a CAD scheme that automatic detection of lesion candidate region (nodule and GGO) from 3D images generated by sparse coding and deep learning technique.

2. Method

The scheme which our proposed has three steps: temporal subtraction, sparse coding, and CNN. In the first step, lesion candidate regions are extracted using temporal subtraction technique. In the second step, the image is reconstructed by sparse coding for the extracted region. In the final step, in order to classify the abnormalities, 3D-CNN identification using reconstructed images is performed. We also conducted experiments to confirm valid parameters used for sparse cording.

2.1. *Temporal subtraction technique*

Kondo et al. [9] proposed a method for making a temporal subtraction image. They determined lesion candidate regions using super voxel techniques and graph cut algorithms. The position of the lesion candidate region was determined by this method.

2.2. *Sparse coding*

Sparse coding techniques were used to reconstruct the lesion candidate regions obtained in the previous section. An image can be represented by a base model that is represented by a combination of several basic bases. This can be expressed by the following equation.

$$\boldsymbol{x} \cong D\boldsymbol{c}. \tag{1}$$

D shows a dictionary in which bases are arranged, and $\boldsymbol{c}$ is a coefficient vector. Also, by resolving the minimization problem of residual sum of squares and regularization term (See Eq. (2)), reconstruction using sparse coding becomes possible. As the regulation term, especially in the case of $0 \leq p \leq 1$ in the l_p norm of the coefficient vector c is often considered.

$$\min_{D,C} \|X - DC\|_2^2 + \lambda \sum_{i=1}^{m} \|\boldsymbol{c}_i\|_p^p \;, \lambda \geq 0. \tag{2}$$

As a result of this optimization problem, the sparse representation is reproduced in the coefficient matrix C. We use $\lambda = 1$ and $p = 1$. In this paper, dictionary learning by K-singular value decomposition (K-SVD) method and coefficient selection by orthogonal matching pursuit (OMP) method were performed [10, 11].

Aharon et al. [10] proposed base learning algorithms, which called K-SVD algorithm. The basic idea is to consider the error of approximating the signal without using $\boldsymbol{d}_l$ when updating the l-th base $\boldsymbol{d}_l$ and to make the basis expressing this error $\boldsymbol{d}_l$ new. The K-SVD algorithm shows good performance against real problems, coupled with its mathematical simplicity and high computational efficiency. It is used as a base learning method for sparse representation in image processing and speech signal processing.

The OMP method [11] is one of greedy algorithms. The greedy algorithms do not guarantee the stability of the solution, but often gives a good approximate solution. That is, it is an algorithm for finding the index set S of non-zero coefficients called "support" from the index set of coefficients used for approximation of observed signals.

Based on the position coordinates obtained in Section 2.1, a lesion candidate region is cut out as a 48 × 48 × 48 sized 3D image. First, the 3D image is divided into small areas to form a collection of small cubes. Note that the cube is assumed that paved without gaps. Next, image reconstruction by the sparse coding method is performed on the small cube using the created dictionary and the OMP method. This operation is performed on all the small cubes to reconstruct the entire image. We call the length of one side of a small cube as base size "a", the number of bases include in dictionary D as base number "b", the index set as coefficient matrix C, the number of nonzero coefficients per column in C as sparse level "s". As shown in Fig.1, by combining these parameters, many reconstructed images are created from one 3D image. We use these a, b, and s parameters $a = 2, 3, 4, 6$, $b = 50, 100, 150, 200$, $s = 1, 3, 5$ respectively.

2.3. *3D-CNN*

CNN is one of multilayered feedforward neural networks that can extract topological features of images. The learning algorithm is designed to recognize human visual patterns, so it can handle extremely changing patterns like handwritten letters.

In this paper, we use a network based on AlexNet [12] and ResNet [13]. This network consists of an input layer, five convolutional layers, three pooling layers, two normalization layers, six residual block and one fully connected layer. We also use Chainer [14] as a framework for deep learning.

We classified using the image acquired in the previous section as input. Incidentally, in order to confirm effective sparse coding parameters in CNN, an image for each parameter was used as an input image. However, since classification is performed for each reconstructed image, it is necessary to determine the original lesion candidate class. Therefore, the original lesion candidate class was determined by using the likelihood of each image.

3. Results

In this paper, we evaluated the performance of the constructed network based on the 5-fold cross validation. Also, we evaluated the proposed method by obtaining true positive rate (TP), false positive rate (FP) and area under the curve (AUC) for each data and calculating the average. TP and FP are defined by the Eq. (3), (4) based on Table 1. In Table 1, abnormal includes nodule and GGO, and normal includes normal tissue such as lung wall and blood vessel. Additionally, test positive (nodule and GGO) and test negative (normal tissue) show the classification result of 3D-CNN.

$$\mathrm{TP} = \frac{a}{a+b} \times 100[\%], \tag{3}$$

$$\mathrm{FP} = \frac{c}{c+d} \times 100[\%]. \tag{4}$$

In the experiment, we applied this method to 31 cases. The obtained results are shown in Table 2. It turned out that sparse level s affects the score more than other parameters.

4. Discussion

As shown in Table 2, sparse level more affect score than other parameters. For this result, we think base size, base number affect image roughness and sparse level affects information volume of image. In this model, we think not image roughness but information volume needs to be similar. Therefore, we think that better results can be obtained by extending the input to 3 channels using these parameter images.

Also, the dataset used in our experiments is more included normal tissue than abnormalities. We think this imbalance data will adversely affect the results. Therefore, it is considered effective to change the dataset or introduction of a loss function considering data imbalance.

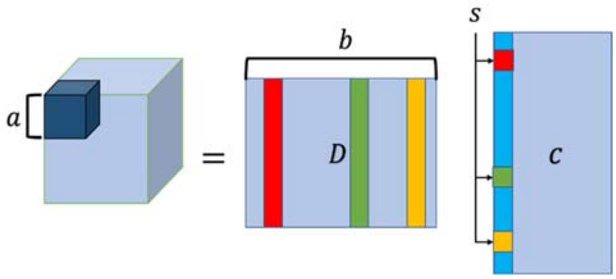


Fig. 1 Sample of image reconstruction by sparse coding. Combining base size a, base number b, and sparse level s. We generated multiple images like the one shown.

Table 1 Performance evaluation method

	Abnormal	Normal
Test Positive	a	c
Test Negative	b	d

Table 2 Experimental results

parameter		AUC
Base size	**2**	0.5651
	3	0.6686
	4	0.6224
	6	0.7463
Base number	**50**	0.6839
	100	0.5566
	150	0.5911
	200	0.5253
Sparse level	**1**	0.7436
	3	0.7107
	5	0.6627

5. Conclusion

In this paper, we developed a CAD system for automatic extraction of lesion candidate region from chest CT images. Our proposed method has three steps, which are temporal subtraction technique, sparse cording and Convolutional neural network. We evaluated our method by 5-fold cross validation method and obtain the result that sparse level s affects the score more than other parameters. Future works include reviewing the dataset and improving the model and introducing other loss functions.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 17H02110.

References

1. Cancer Information Service, https://ganjoho.jp/reg_stat/statistics/stat/summary.html (In Japanese)
2. A. Kano et al., "Digital image subtraction of temporally sequential chest images for detection of interval change", Medical Physics, Vol.21, No.3, pp.453-461, 1994.
3. T. Aoki et al., "Usefulness of computerized method for lung nodule detection on digital chest radiographs using similar subtraction images from different patients", Academic Radiology, Vol.18, No.18, pp.1000-1005, 2011.
4. J. Wright et al., "Sparse Representation for Computer Vision and Pattern Recognition", Proceedings of the IEEE, Vol.98, No. 6, pp.1034-1044, 2010.
5. http://www.image-net.org/challenges/LSVRC/2012/
6. A. Teramoto et al., "Automated detection of pulmonary nodules in PET/CT images: Ensemble false-positive reduction using a convolutional neural network technique", Medical Physics, Vol.43, No.6, pp.2821-2827, 2016.
7. A. Setio et al., "Pulmonary Nodule Detection in CT Images: False Positive Reduction Using Multi-View Convolutional Networks", IEEE Transactions on Medical Imaging, Vol.35, No. 5, pp.1160-1169, 2016.
8. T. Tanaka et al., "Detection of lung nodules on temporal subtraction images Using 3D sparse coding", Control, Automation and Systems, pp.1455-1457, 2017.
9. M. Kondo et al., "A temporal subtraction technique from thoracic MDCT images based on image registration method", Proceeding of the International Conference on Information and Communication Technology ROBOTICS, 2016.
10. M. Aharon et al., "K-SVD: An algorithm for designing over complete dictionaries for sparse representation", IEEE Transactions on Signal Processing, Vol.54, No.11, pp.4311-4322, 2006.
11. Y. C. Pati et al., "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition", Annual Asilomar Conference on Signals Systems and Computers, Vol.1, pp.40-44, 1993.
12. A. Krizhevsky et al., "ImageNet Classification with Deep Convolutional Neural Networks", Neural Information Processing Systems, Vol.1, pp.1097-1105, 2012.
13. K. He et al., "Deep Residual Learning for Image Recognition", IEEE Conference on Computer Vision and Pattern Recognition, pp.770-778, 2016.
14. Chainer, https://chainer.org/

Design of a Data-Driven Multi Controllers Using VRFT and Ensemble Learning

Takuya Kinoshita
Graduate School of Engineering, Hiroshima University,
1-4-1, Kagamiyama, Higashihiroshima city, Hiroshima, Japan

Yuma Morota
Graduate School of Engineering, Hiroshima University,
1-4-1, Kagamiyama, Higashihiroshima city, Hiroshima, Japan

Toru Yamamoto
Graduate School of Engineering, Hiroshima University,
1-4-1, Kagamiyama, Higashihiroshima city, Hiroshima, Japan
E-mail: {takuya-kinoshita, morota-yuma, yama}@hiroshima-u.ac.jp
http://www.hiroshima-u.ac.jp

Abstract

Data-driven control has been proposed for directly calculating control parameters using experimental data. Specifically, the virtual reference feedback tuning (VRFT) has been proposed for linear time-invariant systems. In the field of machine learning, the ensemble learning was proposed to improve the accuracy of prediction by using multiple learners. In this study, a design scheme of data-driven controllers using the ensemble learning and VRFT is newly proposed for linear time-varying systems. The ensemble learning can divide the linear time-varying system into some sections that can be regarded locally as linear time-invariant systems.

Keywords: Data-driven control, PID control, Ensemble learning.

1. Introduction

Data-driven control schemes have been proposed for directly designing a controller using experimental data. Virtual Reference Feed-back Tuning (VRFT)[1] and Fictitious Reference Iterative Tuning (FRIT)[2] have been proposed as data-driven control scheme. According to these schemes, control parameters can be directly calculated using a set of experimental data. However, VRFT and FRIT work well for linear time-invariant systems, and it is difficult to obtain good control performance for linear time-varying systems.

On the other hand, the effectiveness of the deep learning and machine learning have been demonstrated in the field of image recognition. The ensemble learning is a machine learning scheme, and it has been proposed to improve the accuracy of prediction by using multiple learners.

In this paper, a design scheme of data-driven controllers using ensemble learning and VRFT is newly proposed for linear time-varying systems. Specifically, a linear time-invariant system is firstly divided into some linear systems by applying ensemble learning based on decision tree learning. Second, VRFT is applied to each divided linear system for obtaining multiple linear controllers.

2. Design of Data-Driven Controller Using Ensemble Learning

Fig. 1 shows the schematic figure of the proposed data-driven control system where multiple controllers

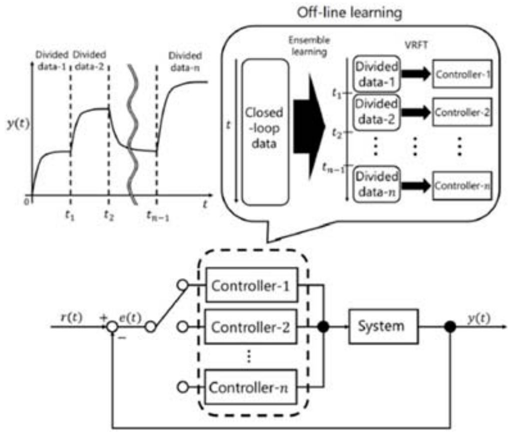


Fig. 1. Schematic figure of the proposed control system.

(icontroller-1, controller-2, ⋯, controller-n) are designed using ensemble learning using the initial closed-loop data. First, the closed-loop data is divided into some linear time-invariant systems based on time t ($t_1, t_2, \cdots, t_{n-1}$). Next, control parameters are calculated using VRFT for each divided data (Divided data-1, Divided data-2, ⋯, Divided data-n).

2.1. *VRFT*

Fig. 2 shows the block diagram of VRFT. $P(z^{-1})$ is a system plant and $C(z^{-1},\theta)$ shows the controller. θ is a control parameters. The optimization problem is adjusting θ so that colsed-loop transfer function $W(z^{-1})$ closed to the following desired reference model $G_m(z^{-1})$.

$$G_m(z^{-1}) = \frac{z^{-1}T(1)}{T(z^{-1})} \tag{1}$$

$$T(z^{-1}) = 1 + t_1 z^{-1} + t_2 z^{-2}, \tag{2}$$

where

$$\begin{cases} t_1 = -2\exp\left(-\dfrac{\rho}{2\mu}\right)\cos\left(\dfrac{\sqrt{4\mu-1}}{2\mu}\rho\right) \\ t_2 = \exp\left(-\dfrac{\rho}{\mu}\right) \\ \rho = \dfrac{T_s}{\sigma} \\ \mu = 0.25(1-\delta) + 0.51\delta \end{cases}. \tag{3}$$

T_s is the sampling time, σ and δ are the user-specified parameters related to the rise characteristic and attenuation characteristic of control system.

In order to match $W(z^{-1})$ to $G_m(z^{-1})$, one-shot experimental input $u_0(t)$ and output $y_0(t)$ is obtained. Next, a pseudo reference input $\tilde{r}(t)$ is calculated as follows:

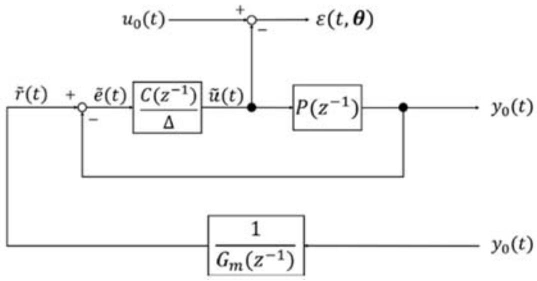


Fig. 2. Block diagram of VRFT.

$$\tilde{r}(t) = \frac{1}{G_m(z^{-1})} y_0(t). \tag{4}$$

The output $\tilde{u}(t)$ is denoted as Eq. (5).

$$\tilde{u}(t) = \frac{C(z^{-1},\theta)}{\Delta}\tilde{e}(t) \tag{5}$$

$$\tilde{e}(t) = \tilde{r}(t) - y_0(t) \tag{6}$$

The following equation denotes the objective function J.

$$J = \frac{1}{2}\sum_{t=1}^{N}\{\varepsilon(t,\theta)\}^2 \tag{7}$$

$$\varepsilon(t,\theta) = u_0(t) - \tilde{u}(t) \tag{8}$$

In this paper, I-P controller is employed and the control parameter θ which minimizes the evaluation function J. From Fig. 2, $\tilde{u}(t)$ becomes given by

$$\tilde{u}(t) = -K_P y_0(t) + K_I\frac{\tilde{e}(t)}{\Delta} \tag{9}$$

By putting $\tilde{e}(t)/\Delta = x(t)$, Eq. (9) is rewritten as follows.

$$\begin{aligned} \tilde{u}(t) &= -K_P y_0(t) + K_I x(t) \\ &= \theta\varphi(t) \end{aligned} \tag{10}$$

Here, the following equations denote θ and $\varphi(t)$.

$$\theta = [K_P\ K_I]^T \tag{11}$$

$$\varphi(t) = [-y_0(t)\ x(t)] \tag{12}$$

The control parameter θ^* which minimizes the objective function J is obtained as Eq. (13).

$$\theta^* = (\Phi^T\Phi)^{-1}\Phi^T U \tag{13}$$

$$\Phi = [\varphi(1)\ \varphi(2)\ \cdots\ \varphi(N)]^T \tag{14}$$

$$U = [u_0(1)\ u_0(2)\ \cdots\ u_0(N)]^T \tag{15}$$

2.2. *Design of learner to divide closed-loop data*

In this paper, a learner is designed to divide closed-loop data into n systems in Fig. 3 is designed. Decision tree learning is a scheme of assigning a data set allocated to a parent node to a child node according to a split function h^* and creating a decision tree such as Fig. 4. The split function h^* is obtained as follows.

$$h^* = arg\ max\ I \tag{16}$$

I is an evaluation function showing the degree of variation of the class in the child node which divided the parent node. Based on the division scheme of the decision tree, the optimum time to split closed-loop data $t^*(t_1^*, t_2^*, \cdots, t_{n-1}^*)$ is obtained using the evaluation function I as follows.

$$t^* = arg\ max\ I(t) \tag{17}$$

Here, the evaluation function I is defined as follows.

$$I(t) = -\sum_{k=1}^{n} J_k^* \tag{18}$$

J_k^* is the minimum value in Eq. (7) of the kth system obtained by dividing the closed-loop data at time t.

2.2.1. Ensemble Learning

Ensemble learning is a scheme of integrating multiple learners to generate one learner. If the accuracy of each learner is higher than 50%, it is known that higher accuracy is achieved than one learner since the learner that estimates erroneously becomes a minority.
In this paper, 'fminsearch.m' in MATLAB & Simulink Ver. 9.4.0813654 (R2018a), Optimization Toolbox is used to find t^*.

2.3. *Controller selection method based on the controller's plane*

Eq. (10) is as follows when the closed-loop transfer function at the top of Fig. 2 and the transfer function $G_m(z^{-1})$ are equal.

$$u_0(t) = -K_P y_0(t) + K_I \frac{e(t)}{\Delta} \tag{19}$$

$$-K_P y_0(t) + K_I \frac{e(t)}{\Delta} - u_0(t) = 0 \tag{20}$$

The following equation defined the equation of the plane passing the origin with x, y, z as axes.

$$ax + by + cz = 0 \tag{21}$$

From Eq. (20) and Eq. (21) can be regarded as an equation of a plane passing the origin with $y, e/\Delta$, and u as axes. From the equation (20), a linear time-invariant system can be represented by a plane and the coefficients consist of PI gains. From this property, input and output data are plotted in a space with $y, e/\Delta$, and u as axes, and a controller corresponding to a plane closest to the plotted points is sequentially selected.

3. Numerical example

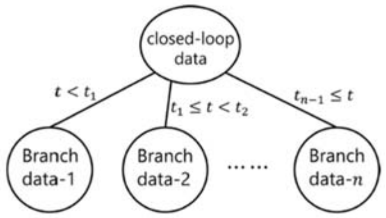


Fig. 3. Learning instrument.

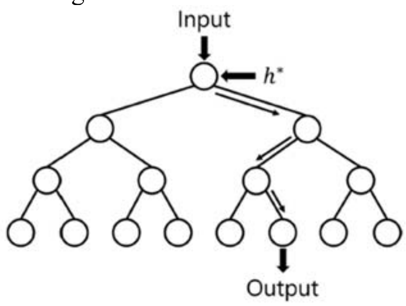


Fig. 4. Decision tree.

3.1. *Control object and setting parameters*

In this numerical example, the following system is discussed.

$$G(s) = \begin{cases} \dfrac{3}{1+20s} & (0 \le t < 200)(22) \\ \dfrac{2}{1+30s} & (200 \le t < 400)(23) \\ \dfrac{1}{1+10s} & (400 \le t \le 600)(24) \end{cases} \tag{22}$$

The reference signal $r(t)$ is given as follows.

$$r(t) = \begin{cases} 100 & (0 \le t < 200)(26) \\ 150 & (200 \le t < 400)(27) \\ 50 & (400 \le t \le 600)(28) \end{cases} \tag{23}$$

White Gaussian noise with zero mean and a variance of $(1/3)^2$ is added to the controlled object. The reference model $G_m(z^{-1})$ was set as Eq. (24).

$$G_m(z^{-1}) = \frac{0.0392z^{-1}}{1 - 1.6057z^{-1} + 0.6449z^{-2}} \tag{24}$$

3.2. *Simulation result*

The control result with fixed PI gains which are calculated by VRFT is shown in Fig. 5. Control results using the fixed PI controller. In Fig. 5. Control results using the fixed PI controller., the overshoot has occurred on the output with fixed PI gains and good control results are not obtained.

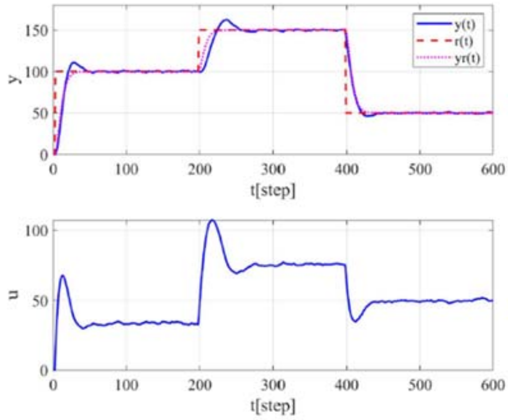


Fig. 5. Control results using the fixed PI controller.

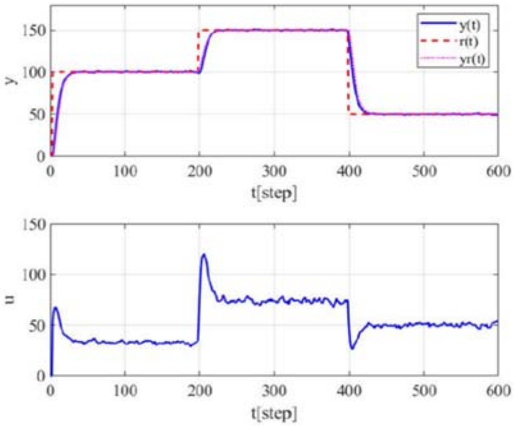


Fig. 6. Control results using the proposed scheme.

Fig. 6. Control results using the proposed scheme. shows the control result applying the proposed scheme. Here, the number of weak learners is 10 . Fig. 7. Estimation result of ensemble learning. and Fig. 8. Trajectories of PI gains. respectively show the estimation result of the system change step in the weak learners and the transition of PI gains in the proposed scheme. Fig. 7 shows that good control result can be obtained using the proposed scheme. In addition, Fig. 8 and Fig. 9 show that the system change step is estimated accurately and PI gains are changed at the proper timing.

4. Conclusions

In this paper, a new control scheme has been proposed. It is a scheme of dividing a linear time-varying system into multiple linear time-invariant systems and applying VRFT to calculate multiple linear controllers. In the numerical example, the effectiveness of the proposed scheme has been shown.

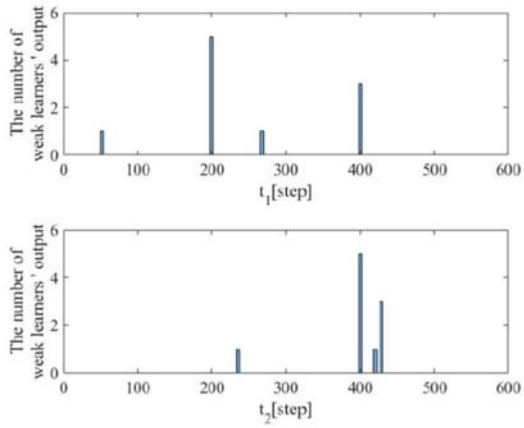


Fig. 7. Estimation result of ensemble learning.

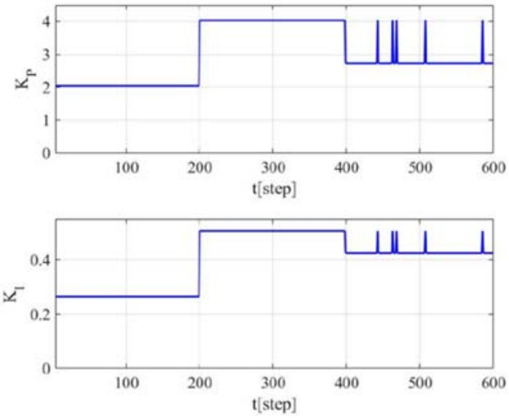


Fig. 8. Trajectories of PI gains.

In the future works, it is necessary to optimize the number of divisions for an unknown system. In addition, the least squares method is susceptible to noises other than normal distribution. Therefore, further consideration is needed such as incorporating bagging[3] to suppress the effects of noise.

References

1. M.C. Campi, A. Lecchini and S.M. Savaresi, “Virtual Reference Feedback Tuning: A Direct Method for the Design of Feedback Controllers”, *Automatica*, **38**(8) (2002), 1337-1346.
2. S. Soma, O. Kaneko and T.Fujii, “A New Approach to Parameter Tuning of Controllers by Using One-Shot Experimental Data: A Proposal of Fictitious Reference Iterative Tuning (in Japanese)”, *Trans. of the Institute of systems, Control and Information Engineers*, **17**(2) (2004) 528–536.
3. L. Breiman, “Bagging Predictors”,*T. Machine Learning*, **24**(2)(1996), 123–140.

Efficient detection device for wafer physical defects

Jianyong Chen [1], Xiaoyan Chen [1]*, Chundong Zhao [1]

[1] *College of Electronic Information and Automation, Tianjin University of Science and Technology, China;*

*E-mail: * cxywxr@tust.edu.cn*
Www.tust.edu.cn

Abstract

Wafer defect detection is an important part of semiconductor manufacturing. In order to improve the efficiency of semiconductor wafer defect detection, this paper designs an efficient visual inspection device. The device uses programmable logic controller (PLC) as controller for the transmission mechanism and uses servo motor as drive device. A CMOS camera is used to capture wafer images, a computer is used for wafer image processing, and results are displayed on a graphical user interface. Camera calibration is implemented by integrating the mapping relationship between the pixel coordinate system and the world coordinate system, the internal and external parameters of the camera and the distortion coefficient. The device proposed in this paper is low in cost and the detection process is stable and reliable. It provides a new solution for wafer defect detection.

Keywords: Semiconductor manufacturing, Machine vision, Camera calibration, Programmable logic controller,

1. Introduction

Silicon wafer processing is the basis of semiconductor manufacturing. Wafer defect detection is one of the most important steps in semiconductor wafer manufacturing. Semiconductor wafer defect detection methods are mainly manual inspections, which are affected by human factors such as visual fatigue and are inefficient [1-3].

Research on the automatic wafer defect detection device was carried out to solve the problem of manual limitations in wafer defect detection. At present, with the development of machine vision technology, a large number of machine vision technologies have been used to detect chip shape, size, physical defects and other aspects [4-8].

The performance of wafer defect detection device directly affects the quality of semiconductor products and the efficiency of semiconductor testing [9-10]. In order to improve the detection efficiency and positioning accuracy of the automatic wafer defect detection device, an optical system is designed and constructed in this paper to obtain a more complete wafer image. Camera calibration is used to eliminate the errors caused by the distortion of CMOS camera and improve the identification accuracy of the detection device. The main structure, motion mechanism, camera bracket and bearing platform of the device are designed and assembled. The Siemens s7-200 PLC is used as a controller to implement the control of all parts of the detection device.

2. Hardware Design

2.1. *Dual-axes motion platform*

The dual-axes(x−y) motion platform is selected for wafer defect detection. The horizontal ball screw pair of the platform is mounted on the longitudinal ball screw pair through the bracket, the longitudinal ball screw pair is directly mounted on the base surface, and the wafer stage is mounted on the horizontal ball screw pair through the bracket. The lead screw pair is supported by both ends, and the drive motor and ball screw pair are directly connected. The assembly model is shown in Fig. 1.

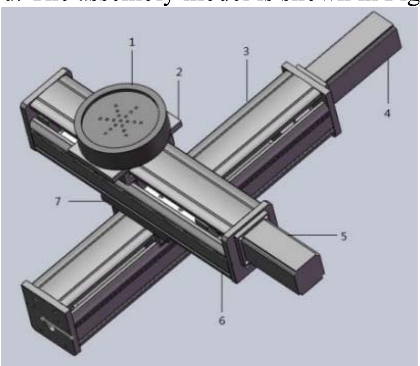


Fig. 1 Dual-axes motion platform. 1. Wafer stage 2. Bracket 3. Y-axis ball screw pair 4. Y-axis servo motor 5. X-axis servo motor 6. X-axis ball screw pair 7. Bracket

2.1.1. *Ball screw pair*

The ball screw pair is a transmission component that converts rotary motion into linear motion with high precision, high rigidity and high efficiency. Due to its low frictional resistance, ball screw pairs are widely used in a variety of industrial equipment and precision instruments.

The ball screw pair is rotationally driven by balls mounted between the spiral groove of the lead screw and the nut. According to the corresponding function, the ball screw pair is divided into four components, which are the screw, the nut, the ball and the reverser. In order to meet the requirements of positioning accuracy, feed rate, rapid response and stability of the motion platform, it is indispensable to choose a ball screw pair reasonably. The Y-axis ball screw pair is located in the lower layer with a larger load and higher precision requirements. Therefore, the Y-axis ball screw pair design parameters are used as the overall ball screw pair design index.

(i) Lead

The lead P_h is generally determined according to the maximum speed of the feed V_{max}, the maximum speed of the servo motor N_{max} and the transmission ratio i between the motor and the screw, which can be calculated by formula (1).

$$P_h = \frac{V_{max}}{i \times N_{max}} \tag{1}$$

In this paper, V_{max} =2400mm/min, P_h =5mm, N_{max}=480r/min.

(ii) Load and speed of ball screw pair

Equivalent load F_m is the actual axial force exerted on the ball screw by the transmission device. This device has no cutting operation, so equivalent load is equal to static friction force. F_m can be calculated by formula (2).

$$F_m = \mu \times (M_1 + M_2) \tag{2}$$

μ is the friction coefficient, the value is 0.04. M1 is the X-axis ball screw counter gravity, about 80N. M2 is the bearing table and the carrier table carrier gravity, about 20N. According to formula (2), F_m can be calculated as 0.4N.

(iii) Rated dynamic load

$$\mathrm{L}_d = \left(\frac{C_{am} \times f_a \times f_c}{F_m \times f_W}\right) \times P_h \tag{3}$$

According to formula (3), rated dynamic load C_{am} can be calculated as 110N. L_d is the expected operating distance, 2.5×10^8m. F_w is the load coefficient, set 1.0. f_a is the precision coefficient, set 1.0. f_c is the reliability coefficient, which is 0.62.

(iv) Ball screw subbottom diameter

Ball screw subbottom diameter d_{2m} can be calculated by the formula (4).

$$d_{2m} \geq 10\sqrt{\frac{10F_0 \times L}{\pi \times \partial_m \times E}} \tag{4}$$

F_0 is the static friction of the guide rail. L is the distance between the support shafts at both ends. ∂_m is the maximum allowable axial deformation of the ball screw. We figured out that d_{2m} is at least 0.43mm.

2.1.2. *Servo motor*

The device uses servo motor as the driving power system. The performance of the servo motor largely determines the positioning accuracy of the moving platform.

(i) Pulse equivalent

Pulse equivalent δ is the displacement of the actuator for each output pulse of the servo motor. It can be calculated from formula (5).

$$\delta = \frac{P_h \times i_m}{4 i_n \times q} \tag{5}$$

i_m is the electronic gear ratio of servo motor, the value is 2.5. i_n is the transmission ratio. Since the motor shaft and lead screw are directly connected through the coupling, the value of i_n is 1. q is the resolution of servo motor encoder, the value is 2500 P/R.

(ii) Capacity and torque

When the motor is running, in order to ensure the stable operation of the system, the moment of inertia of the full equivalent load on the motor J should match the moment of inertia of the motor rotor J_m.

$$J_1 = \sum J_i \left(\frac{n_i}{n_m}\right)^2 + \sum m_1 \left(\frac{v_1}{2\pi \times n_m}\right)^2 \quad (6)$$

$$\mathrm{J} = J_m + J_1 \quad (7)$$

Capacity and torque can be calculated from formula (6), (7). J_i and n_i are respectively the moment of inertia and speed of each rotating part. m_1 and v_1 are the mass and speed of each linear moving part respectively. J_m and n_m are the moment of inertia and speed of the motor respectively.

2.2. *PLC controller*

It is very important to select reliable and stable controller as the core of control system to ensure the stable and safe operation of transmission device. Programmable logic controller (PLC) is selected as the controller according to the requirements of the controlled object, so as to realize the effectiveness of the system. This paper follows the following PLC selection principles

(i) Under the condition of meeting the requirement of the controlled object, the PLC controller can be used effectively.

(ii) The control system has simple structure and strong maintainability.

(iii) A certain amount of storage space and I/O points are reserved for subsequent adjustments and expansions.

According to the above requirements, Siemens s7-200 PLC, 224XP CN DC/DC/DC CPU were adopted in this paper.

3. Image acquisition system

The image acquisition system designed in this paper consists of CMOS industrial camera, lens and ring light source. CMOS industrial cameras transmit data in parallel for faster speeds. The purpose of a lens is to image an object optically onto a sensor. The diameter of the ring light source is 100mm-140mm, which is used to make the image receiving light uniform and clear. The structure of the image acquisition system is shown in Fig.2

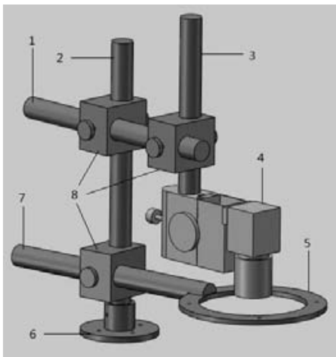


Fig. 2 Image acquisition system. 1. Camera shaft 2. Spindle 3. Camera shaft 4. CMOS camera 5. Light source fixing ring 6. Flange base 7. Light source shaft 8. Fixture

4. Camera Calibration

The main role of camera calibration is to correct image distortion and determine the positional relationship between the pixel coordinate system and the world coordinate system.

4.1. *Calibration principle*

(i) Camera internal and external parameters

As shown in formula (8), (9), The default value of Z plane is 0. φ is the scale factor. A is the internal parameter matrix. f_x,f_y is the scale factor of the u-axis and v-axis of the pixel coordinate system. γ represents the scale deviation of pixel points in the image coordinate system. u_0 and v_0 are the center of the image plane. r_1, r_2 and t are rotation matrix and translation vectors respectively.

$$\varphi \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = A[r_1, r_2, t] \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (8)$$

$$\mathrm{A} = \begin{bmatrix} f_x, & \gamma, & u_0 \\ 0, & f_y, & v_0 \\ 0, & 0, & 1 \end{bmatrix} \quad (9)$$

As shown in formula (10), (11), (12), H is a homography matrix, and the basic constraint condition of internal parameter matrix can be deduced according to the properties of rotation matrix. The homography matrix can be obtained by the least square method with the pixel coordinates and world coordinates of the detected corner points.

$$\mathrm{H} = [h_1, h_2, h_3] = A[r_1, r_2, t] \quad (10)$$

$$\begin{cases} r_1 = A^{-1}h_1 \\ r_2 = A^{-1}h_2 \end{cases} \quad (11)$$

$$\begin{cases} h_1{}^T A^{-T} A^{-1} h_2 = 0 \\ h_1{}^T A^{-T} A^{-1} h_1 = h_2{}^T A^{-T} A^{-1} h_2 \end{cases} \quad (12)$$

(ii) Camera distortion factor

Ideally, the least square method can be used to calculate the camera distortion coefficient according to formula (13) (14) (15). Ideally (no distortion), the pixel coordinates are (u, v), and the coordinates in the image coordinate system are (x, y). When there is distortion, the real pixel coordinates are $(\bar{u}, \bar{v})$, and the coordinates in the image coordinate system are $(\bar{x}, \bar{y})$.

$$\begin{cases} \bar{x} = x + x(k_1(x^2+y^2) + k_2(x^2+y^2)^2) \\ \bar{y} = y + y(k_1(x^2+y^2) + k_2(x^2+y^2)^2) \end{cases} \quad (13)$$

$$\begin{cases} \bar{u} = u + (u-u_0)(k_1(x^2+y^2) + k_2(x^2+y^2)^2) \\ \bar{v} = v + (v-v_0)(k_1(x^2+y^2) + k_2(x^2+y^2)^2) \end{cases} \quad (14)$$

$$\begin{bmatrix} (u-u_0)(x^2+y^2), (u-u_0)(x^2+y^2)^2 \\ (v-v_0)(x^2+y^2), (v-v_0)(x^2+y^2)^2 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} \bar{u}-u \\ \bar{v}-v \end{bmatrix} \quad (15)$$

4.2. *Calibration results*

By using a standard checkerboard, the camera takes pictures of the imaging area from different angles to obtain a set of calibration images. The image corner information is extracted, and the internal parameters, external parameters and distortion coefficients of the camera are obtained by the mapping relationship between the pixel coordinates of the corner points and the corner coordinates of the world coordinate system.

In this paper, 7×9 black and white checkerboard grids are selected, and the size of each grid is 8mm×8mm. A total of 9 standard checkerboard grid images are collected for calibration. According to formula (10)-(15), the internal parameter matrix A and distortion coefficients dist are obtained as follows:

$$\mathrm{A} = \begin{bmatrix} 26704.79043, & 0, & 1384.94299 \\ 0, & 26754.77535 & 1415.48126 \\ 0, & 0, & 1 \end{bmatrix}$$

$$\mathrm{dist} = \begin{bmatrix} 1.18344 \\ -1034.57967 \\ 0.02802 \\ 0.01921 \\ -6.37039 \end{bmatrix}$$

The maximum calibration error is δ_{max}=0.2125pix, and the average error is δ_{ave} =0.193048pix.

5. Acknowledgement

The authors would like to appreciate for the financial support from Tianjin Municipal Science and Technology Bureau (Grand NO.18YFZCGX00360) and from the National Natural Science Foundation of China (Grand NO.61903274 and NO.41704131), the same appreciate for Tianjin FLY Tech Co., Ltd for offering samples and for Intelligent Edge Computing Joint LAB of TUST for the testing result analyzing.

References

1. Zhou, Q. and B. Zhou, An impending deadlock-free scheduling method in the case of unified automated material handling systems in 300 mm wafer fabrications[J]. Journal of Intelligent Manufacturing. Jan 2018, Vol. 29 Issue 1, pp. 155.
2. Tirkel, I., The efficiency of inspection based on out of control detection in wafer fabrication[J]. Computers & Industrial Engineering. Sept, 2016, Vol. 99, pp. 458.
3. Stefan Bengtsson, "Semiconductor Wafer Bonding - A Review Of Interfacial Properties And Applications[J]" Journal Of Electronic Materials, 1992, vol. 21, no. 8, pp. 841-862.
4. Wenjun, Liu, Thermography techniques for integrated circuits and semiconductor devices[J]. Sensor Review, 2007, Vol. 27, Issue 4, pp. 298-309.
5. Amatya, S, Automated Detection of Branch Shaking Locations for Robotic Cherry Harvesting Using Machine Vision[J]. Robotics, 2017, Vol 6, Issue 4, pp 31.
6. Binch, A. and C.W. Fox, Controlled comparison of machine vision algorithms for Rumex and Urtica detection in grassland[J]. In Computers and Electronics in Agriculture August 2017, 140:123-138.
7. Huang, B, Full Length Article: Research and implementation of machine vision technologies for empty bottle inspection systems[J]. In Engineering Science and Technology, an International Journal February 2018, 21(1):159-169.
8. Eguiraun, H., O. Casquero and I. Martinez, The Shannon Entropy Trend of a Fish System Estimated by a Machine Vision Approach Seems to Reflect the Molar Se:Hg Ratio of Its Feed[J]. Entropy, 2018, Vol 20, Issue 2, pp. 90.
9. Joy, J, Capacitance-voltage profiling of MOS capacitors: A case study of hands-on semiconductor testing for an undergraduate laboratory[J], American Journal of Physics, Oct,2018, Vol. 86 Issue 10, pp 787-796.
10. Wang, K.J, T.C. Hou, Modelling and resolving the joint problem of capacity expansion and allocation with multiple resources and a limited budget in the semiconductor testing industry[J], International Journal of Production Research, 2003, Vol. 41 Issue 14, pp 217-3235.

Wafer defect detection method based on machine vision

Chundong Zhao[1], Xiaoyan Chen[1*], Dongyang Zhang[1], Jianyong Chen[1], Kuifeng Zhu[2], Yanjie Su[2]

[1]College of Electronic Information and Automation, Tianjin University of Science and Technology, Tianjin, China;

[2]Tianjin Fly Tech Co., Ltd, Tianjin, China

*E-mail: * cxywxr@tust.edu.cn*

Www. tust. edu. cn

Abstract:

With the development of integrated electronic circuit manufacturing technology, enterprises have put forward higher requirements for the quality of silicon chips. Aiming at the low efficiency of silicon wafer defect detection, this paper proposes an automatic defect detection method based on machine vision. The voiding algorithm based on flood fill can effectively extract the inner contour information of the wafer profile. A rotation correction algorithm is proposed to correct the wafer yaw angle. The actual wafer was used to verify the performance of the proposed method. The results show that the proposed method is effective in detection accuracy.

Keywords: machine vision, flood fill, defect detection, rotation correction

1. Introduction

Wafer manufacturing is the core process of the semiconductor chip industry. The complexity of semiconductor silicon wafer processing technology has led to more and more sources of defects [10], which puts higher requirements on existing detection technologies. The defect detection is done manually under the microscope, which is affected by visual fatigue and human subjective factors. The detection accuracy cannot be guaranteed, and the artificial vision inspection cannot meet the requirements of large-scale production and manufacturing.

The machine vision detection method can largely overcome the shortcomings of the manual detection method such as low accuracy, poor real-time performance, low efficiency, and high labor intensity, and has been widely studied and applied in industrial testing [1]. Li applies machine vision to the detection of surface defects in integrated circuit wafers, and uses fuzzy logic to analyze different shapes of surface pit defects [2]. Barni based on machine vision system for quality testing of chicken. According to the color information of the chicken image, the mathematical problem is used to extract the features of the potential problem area, and then classified according to the predefined list of quality problems [3]. Hoang introduced a method for detecting surface defects in leather, using OTSU method for defect segmentation [4].

With the development of computer technology and image processing technology, computer vision detection technology has been widely studied and applied in various fields. Aiming at surface defects such as dirt and particles generated during the manufacturing process of silicon wafers, this paper makes research on the technology of silicon wafer surface defect detection based on machine vision. The rotation algorithm based on the cumulative Hough linear transformation can correct the declination of the wafer. The adaptive reconstruction grid method is used to reconstruct the wafer grid, which can effectively filter out the information outside the grain.

2. Image acquisition device

The image acquisition device consists of CMOS industrial camera, lens, ring light source, and light source controller. The structure diagram of the image acquisition device is shown in Fig 1.

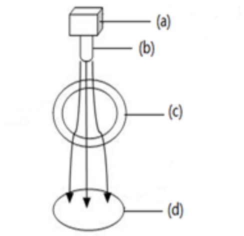


Fig. 1. Structure diagram of the image acquisition device. (a) is the industrial camera. (b) is the lens. (c) is the ring light source. (d) is the silicon wafer.

This device uses a MD-UB1000 CMOS industrial camera. The industrial camera does not need an additional image acquisition card during the image acquisition process, and the camera and the computer can use a professional camera connection cable for data transmission. The camera imager transmits data in parallel at the front end during data transmission, and the bandwidth of the amplifier is very low. The main parameters of the camera include: the pixel size is 1.67um; the effective pixels are 10 million; the resolution ratio is 3664 × 2748; the pixel bit depth is 12 bits. The lens is an MV-JT08 zoom lens. The main parameters include: focal length of 8mm, distortion of less than 1%, closest distance of 0.1m. The ring light source is a FJI-RL150-A00-W ring light source with a diameter of 15cm. The light is radiated vertically from different directions around it. Compared with parallel light, a better quality image can be obtained.

3. Wafer image correction algorithm

There is a deviation in the process of manually placing the wafer, so that the vertical and horizontal intersection lines of the wafer cannot completely match the actual horizontal and vertical directions. In order to accurately locate the grain position coordinates, the tilt angle must be calculated by a certain algorithm, and the wafer is rotated and corrected according to the tilt angle.

Rotation algorithm include graying, thresholding, hole extraction, Hough transform of cumulative probability, and calculation of rotation angle. The flowchart of the algorithm is shown in Fig 2.

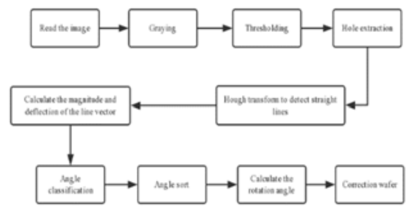


Fig. 2. The flowchart of the algorithm

3.1 Hole filling algorithm

Hough transform requires detection of white straight lines. Hole extraction can turn black grid lines into white and white grains into black. Hole extraction refers to extracting information about the contour of a completely closed contour and filling it with color. The white grid of the image divided by the threshold is used as the outline, and the information of its inner circumference is exacted.

3.1.1 Threshold processing

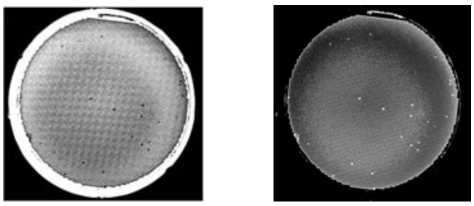
Threshold processing can be regarded as a statistical decision problem, the purpose of which is to minimize the average error introduced in the process of assigning to two or more classes [9,11]. The image has other interference information besides the wafer. Threshold processing of the acquired image can filter out the interference information to obtain the clearest wafer information. Fixed threshold processing is defined as:

$$g(x,y)=\begin{cases} maxval & f(x,y)>T \\ 0 & f(x,y)<T \end{cases} \quad (1)$$

In formula (1), g (x, y) represents the gray value after threshold processing, f (x, y) represents the gray value before threshold processing, and T represents the selected threshold value within the range of pixel gray values. In this paper, the threshold set is T=180.

3.1.2 Flood fill

Flood filling is a method of filling connected areas with a specific color [5-6]. By setting the upper and lower limits of connectable pixels and the connection method, different filling effects are achieved. In order to get the white lines needed for Hough detection, we set the gray value of the fill color to 255. The comparison of the wafer before and after the hole extraction is shown in Fig 3.



(a) (b)

Fig. 3. The comparison of the wafer before and after the hole filling. (a) is threshold image. (b) is hole extraction image

3.2 Hough transform to detect straight lines

Hough transform is a feature extraction technology in image processing [7-8], which is widely used in image analysis and computer vision. In the parameter space, a Hough result is obtained by calculating a local maximum of the accumulated results to obtain a set that conforms to the specific shape. Huff transform uses a transformation between two coordinate systems to map a curve or line with the same shape in one space to a point in another coordinate space to form a peak, thereby transforming the problem of detecting any shape into a statistical peak problem. By performing a Hough transform on the wafer image after the hole filling, the linear distribution of the wafer can be detected.

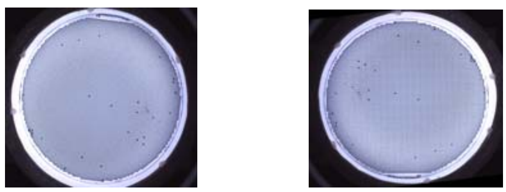
3.3 Angle correction

The coordinates of the two ends of the line segment detected by the Hough transform are subtracted to obtain the size $v(x1-x2, y1-y2)$ of the line segment vector, and the coordinate components v.x and v.y of the vector on the X and Y axes are determined according to the vector size. Let angle be the positive angle of the vector v and the X axis. Then the angle value should satisfy formula (2).

$$\text{angle} = \arctan\left(\frac{v.y}{v.x}\right) \quad (2)$$

The included angle between each line segment and the positive direction of the X axis is not exactly the same, so all the angles need to be further carefully divided. Hierarchical clustering is used to divide a set of data into several categories. After sorting, the angle group containing the largest number of angles is obtained, and the average value of all angles in the array is taken as the final corrected angle.

The affine transformation can realize the rotation correction of the wafer image. Affine transformation refers to the process in which a vector space is transformed into another vector space by completing a linear transformation in geometric space. The affine transformation can maintain the "flatness" and "parallelism" of a two-dimensional image. The comparison before and after correction is shown in Figure 4.



(a) (b)

Fig. 4. The comparison before and after correction. (a) is before correction. (b) is after correction.

4. Wafer grid reconstruction

The number and size of the grains of each wafer are different. Before performing wafer defect detection, different standard templates need to be designed according to the specifications of the grain. The standard template is used as a reference for comparison with the original image during detection. Based on the above problems, this paper proposes an adaptive reconstruction grid method for wafer grid reconstruction. The wafer grid reconstruction algorithm flowchart is shown in Fig 5.

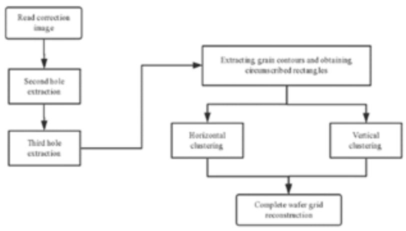


Fig. 5. Wafer grid reconstruction algorithm flowchart

The rule of horizontal clustering is that the Y coordinates at the same point of every two rectangular units are compared, and if the absolute value of the difference between the Y coordinate values is less than 5 pixels, the two rectangles are divided into one class. The rule of vertical clustering is that the X coordinates at the same point of every two rectangular units are compared, and if the absolute value of the difference between the X coordinate values is less than 5 pixels, the two rectangles are divided into one class. Finally, the intersection of the results of horizontal clustering and vertical clustering is taken to complete the grid reconstruction. The effect of wafer grid reconstruction is shown in Fig 6.

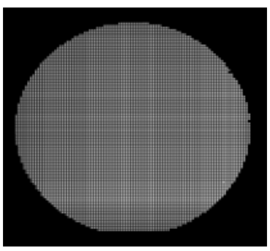


Fig. 6. The effect of wafer grid reconstruction

5. Defect detection

In this paper, a progressive s-shape is used to traverse each grain. The grain defect is composed of a plurality of black pixels in the threshold map. Hole extraction algorithm is performed for each grain to calculate the number of non-zero pixels in the image. The number of pixels in the grain is multiplied by a set threshold, and the number of non-zero pixel points in the grain is more than the number of pixels in the rectangular element, multiplied by the set threshold, the grain is identified as defective.

In this experiment, a total of 1243 grains and 94 defective grains were found in six randomly selected test areas. The total amount of grains and the number of defective grains are different in each test area. The data are shown in Table 1.

Table 1 Statistical results of machine vision detection

Area	Actual defect	Detected defects	Accuracy
1	15	15	100%
2	11	10	90.9%
3	19	18	94.7%
4	20	19	95.0%
5	18	16	88.9%
6	11	11	100%
Total	94	89	94.7%

6. Conclusion

With the rapid development of the semiconductor industry, it is of great significance to apply machine vision technology to detect wafer surface defects. This paper designs an image rotation correction and adaptive grid reconstruction algorithm based on the VS platform and the opencv vision library. The image rotation correction can eliminate the deviation generated in the film placement process and facilitate the positioning of the grain position. Adaptive grid reconstruction algorithm can meet the detection of different specifications of grains. In this paper, the online detection of defective grains is realized, but there are still some shortcomings. Image acquisition is susceptible to interference from the external environment, and the algorithm does not work well for images with uneven light reception.

Acknowledgements

The authors would like to appreciate for the financial support from Tianjin Municipal Science and Technology Bureau (Grant NO.18YFZCGX00360) and from the National Natural Science Foundation of China (Grant NO.61903274 and NO.41704131), the same appreciate for Tianjin FLY Tech Co., Ltd for offering samples and for Intelligent Edge Computing Joint LAB of TUST for the testing result analyzing.

References

1. Tang B, Kong J Y, Wu S Q, Review of surface defect detection based on machine vision. *Journal of Image and Graphics*, 2017, pp 1640-1663.
2. Li H and Lin J C, Using fuzzy logic to detect dimple defects of polished wafer surfaces. *IEEE Transactions on Industry Applications*, 1994, pp 317-323.
3. Barni M, Cappellini V, Mecocci A .Colour-based detection of defects on chicken meat . *Image and Vision Computing*, 1997, pp 549-556.
4. Hoang K, Wen W, Nachimuthu A, Achieving automation in leather surface inspection. *Computers in Industry*, 1997, pp 43-54.
5. Dharmalingham, Vivekanandan, Kumar, Dhananjay, A model based segmentation approach for lung segmentation from chest computer tomography images, *Multimedia Tools* and Applications, 2019.
6. Meem Mahinul Islam, Dhar Pranab Kumar, Khaliluzzaman Md., Shimamura, Tetsuya, Zebra-Crossing Detection and Recognition Based on Flood Fill Operation and Uniform Local Binary Pattern, *2nd International Conference on Electrical, Computer and Communication Engineering, ECCE 2019*, April 1, 2019.
7. Park Hyunhee, Lane detection algorithm based on Hough transform for high-speed self-driving vehicles, *International Journal of Web and Grid Services*, 2019, pp 240-250.
8. Wang Chun Fang, Fast line extraction algorithm based on improved Hough transformation, *Advanced Materials Research,* 2014, pp 3612-3615.
9. Ariawan Angga, Pebrianti Dwi, Ronny, Akbar Yudha Maulana, Margatama Lestari, Bayuaji Luhur, Image Processing-Based Flood Detection, *Lecture Notes in Electrical Engineering,* 2019, pp 371-380.
10. Pan Zhongliang, Chen Ling, Zhang, Guangzhao, IC wafer defect detection using image segmentation based on cultural algorithms, Proceedings of SPIE - *The International Society for Optical Engineering*, 2008.
11. Niu Zuodong, Li Handong, Research and analysis of threshold segmentation algorithms in image processing, *Journal of Physics: Conference Series*, July 12, 2019.

Analysis and Circuit Design of a Novel 4D Chaotic System

Yan Sun, Yongchao Zhang, Jiaqi Chen

Department of Automation, Tianjin University of Science and Technology,
1038 Dagunanlu Road, Hexi District, Tianjin 300222, PR China

E-mail: *sunyantjkj@qq.com*; *www.tust.edu.cn*

Abstract

In this paper, a novel 4D chaotic system is proposed. First, the basic dynamic characteristics of this system are analyzed theoretically. Second, dynamical properties of the system are investigated by phase portraits, Poincaré sections, Lyapunov exponent spectrum and bifurcation diagram. In addition, an analog circuit of the system is designed and simulated by Multisim, a circuit simulation software. The experimental results show that the circuit simulation results are consistent with the numerical simulation results. Therefore, the existence of chaotic attractor in the system is verified from the physical level.

Keywords: Novel 4D chaotic system; Dynamic characteristics; Simulation analysis; Analog circuit design.

1. Introduction

Since 1963, when Lorenz discovered the butterfly effect [1], more and more scholars have studied chaos. In recent decades, with the development of mathematical theory, chaos research has also made great progress with emerging a number of newly built systems. For example, in 1993, Chua et al. designed Chua circuits in which the mathematical model became a very classical chaotic system [2]. And Sprott J et al. proposed Sprott system in the following year [3]. In 1999, Chen system was successfully constructed by Chen et al. [4] In the early 21century, Lü et al. discovered the Lü system [5]. A three-dimensional continuous autonomous chaotic system was constructed by Liu et al. called Liu system [6]. Qi et al. proposed a four-dimensional chaos system, namely Qi system [7]. Except to studying how to construct a system with better dynamic performance, on the basis of existing systems, many scholars also investigate and realize chaotic system through numerical analysis, analog circuit implementation, digital circuit implementation and other methods [8,9,10,11]. In recent years, chaos theory has been applied in many fields, including information studying encryption and secure communication [12,13,14], optoelectronics [15], astrophysics [16], electromagnetic mechanics [17] and economics [18]. Through chaos theory, we can not only further investigate the nonlinear characteristics and their generation mechanism in these fields, but also provide a theoretical basis for the construction of new chaotic systems.

In this paper, a novel 4D chaotic system is proposed. Firstly, the dissipation and equilibrium of the system are theoretically analyzed. Then, simulation analysis is carried out through phase portraits, Poincaré section diagrams, Lyapunov exponent spectrum and bifurcation diagram to analyze the dynamic characteristics of the system in detail. Finally, in order to illustrate the existence of chaotic attractor in the system from the physical level, the circuit design and simulation are carried out with the circuit simulation software Multisim. Moreover, the existence of chaotic attractor in the system is further verified in the experiment.

2. 4D System Model

In this paper, a novel 4D chaotic system is proposed. The system equation of motion is as follows:

$$\begin{cases} \dot{x} = -cx + ay + xz + yw \\ \dot{y} = -ax + bw \\ \dot{z} = -x^2 + w^2 \\ \dot{w} = -xy - by - zw + u \end{cases} \tag{1}$$

In formula (1), x, y, z, w are the system state variables, a, b, c are the system adjustable parameters, u is the external input of the system, and all four numbers are positive.

2.1 Dissipation

The divergence of formula (1) is:

$$\nabla V = \frac{\partial \dot{x}}{\partial x} + \frac{\partial \dot{y}}{\partial y} + \frac{\partial \dot{z}}{\partial z} + \frac{\partial \dot{w}}{\partial w} = -c + z - z = -c \qquad (2)$$

The parameter c is positive, so the divergence of system (1) is negative. This indicates the system (1) is a dissipative system.

2.2 Equilibrium point

Let all the left sides of equation (1) are 0, then equation (1) becomes:

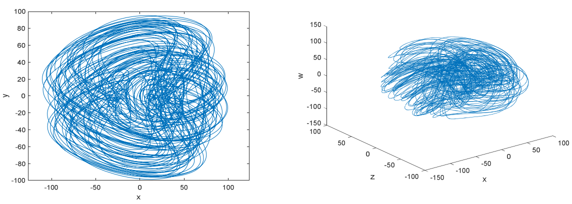
$$\begin{cases} 0 = -cx + ay + xz + yw \\ 0 = -ax + bw \\ 0 = -x^2 + w^2 \\ 0 = -xy - by - zw + u \end{cases} \qquad (3)$$

Through formula (3), calculate the equilibrium point of system (1). It is found that there is no equilibrium point in the system. A chaotic attractor generated by a system without equilibrium point is a hidden chaotic attractor.

2.3 Basic dynamic simulation analysis

2.3.1Phase portraits and Poincaré mapping diagram

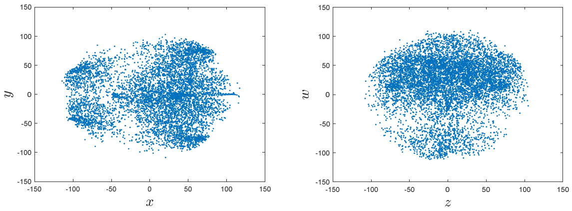
The parameters and external input are a = 50, b = 25, c = 0.47, u = 70, the initial values are (1,1,1,1), the phase portraits of system (1) are shown in Fig.1.



(a) (b)

Fig.1 Phase portraits of system (1) with a = 50, b = 25, c = 0.47, u = 70: (a) $x-y$; (b) $x-z-w$.

As can be seen from Fig.1, chaos is generated in the system with the selected parameters and initial values. In order to further illustrate that the state is chaotic indeed, the Poincaré mapping diagrams under the same parameters and initial values are given. w = 0 and y = 0 are the selected sections, and the results are shown in Fig.2(a) and Fig.2(b).



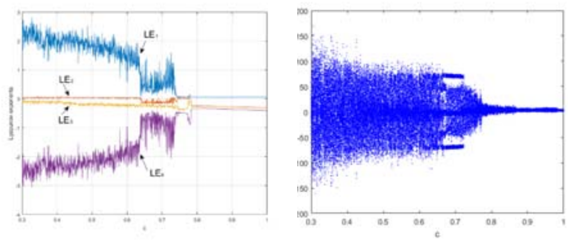
(a) (b)

Fig.2 Poincaré sections of system (1):
(a) w = 0; (b) y = 0.

As can be seen from the Poincaré sections, the map of system (1) under these two sections are many points with dense distribution and fractal structure. This indicates the chaotic behavior occurs in this system.

2.3.2 Lyapunov exponent spectrum and bifurcation diagram

The most common way to determine whether a system is chaotic is to determine the number of positive Lyapunov exponent of a system. For a 4D system, if there are two positive Lyapunov exponents, the system is hyperchaotic. If there is one positive Lyapunov exponent, the system is chaotic. As analyzing this chaotic system, the fixed parameter is a = 50, b = 25, u = 70, and the initial values are (1,1,1,1). The Lyapunov exponent spectrum and bifurcation diagram of system (1) changing with the parameter c in the range (0.3,1) are made. The results are as shown in Fig.3(a) and Fig.3(b) respectively.



(a) (b)

Fig.3 (a): Lyapunov exponent spectrum and (b): bifurcation diagram of system (1).

By combining the Lyapunov exponent spectrum and bifurcation diagram, it can be analyzed that when $c \in (0.3, 0.74)$, system (1) is chaotic; When $c \in (0.74, 0.783)$, system (1) is quasi-periodic. When $c \in (0.783, 1)$, system (1) is periodic. Select c = 0.47, c = 0.7224, c = 0.76 and c = 0.9 of system (1) in these three ranges respectively. Make the corresponding phase portraits, the results are as shown in Fig.4. When c = 0.47, the four Lyapunov exponents are LE_1 = 2.5001, LE_2 = 0, LE_3

= –0.2192, LE_4 = –2.8070, which accord with the feature of chaotic state. When c = 0.7224, the four Lyapunov exponents are LE_1 = 1.2961, LE_2 = 0, LE_3 = –0.2706, LE_4 = –1.7442, which accord with the feature of chaotic state. When c = 0.76, the four Lyapunov exponents are LE_1 = 0, LE_2 = 0, LE_3 = –0.3503, LE_4 = –0.4908, which accord with the feature of quasi-periodic state. When c = 0.9, the four Lyapunov exponents are LE_1 = 0, LE_2 = –0.2664, LE_3 = –0.3201, LE_4 = –0.3752, which accord with the feature of periodic state. These features indicate that the dynamic behavior of the system is very divers.

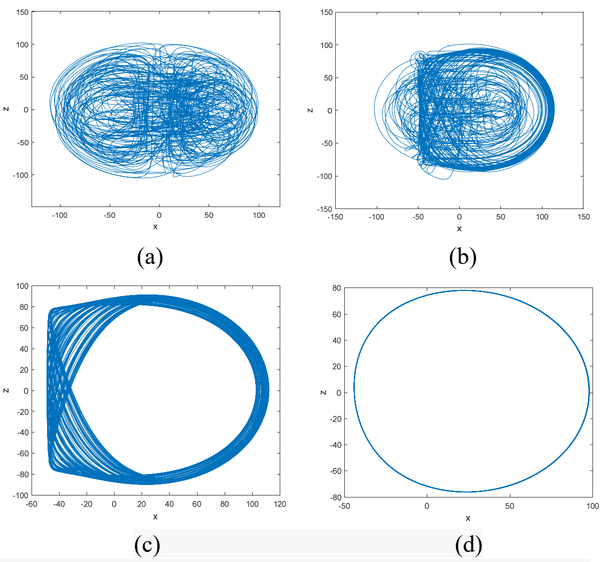


Fig.4 The phase portraits with different values of parameter c: (a) c = 0.47; (b) c = 0.7224; (c) c = 0.76; (d) c = 0.9.

3. Circuit design and implementation

Multisim software is used to design and realize an analog circuit of a system. The selected parameters are a = 50, b = 25, c = 0.47, u = 70, and the initial values are (1,1,1,1). The rated voltage of the power supply in the circuit is $\pm 15V$, so the linear dynamic range of the operational amplifier is $\pm 13.5V$. But the chaotic attractor range of the system is much larger than $\pm 13.5V$, therefore, the spatial proportional compression transformation should be carried out for the system (1). Let:

$$x = 10X, y = 10Y, z = 10Z, w = 10W \qquad (4)$$

Put equation (4) into equation (1) and replace X, Y, Z and W with x, y, z and w again, then:

OS15_Xiaoyan Chen (5)

Now the circuit is designed with the improved method. The circuit is standardized in advance, and the parameters are applied into equation (5). Then, equation (5) becomes:

$$\begin{cases} \dot{x} = -0.47x - 50(-y) - 10(-x)z - 10(-y)w \\ \dot{y} = -50x - 25(-w) \\ \dot{z} = -10x^2 - 10(-w)w \\ \dot{w} = -10xy - 25y - 10zw - (-70) \end{cases} \qquad (6)$$

In order to observe the phase portraits of system (1) in the oscilloscope, time scale transformation is required. Let $t = \tau_0 T, \tau_0 = 100$ and use t to replace T, so:

$$\begin{cases} \dot{x} = -47x - 5000(-y) - 1000(-x)z - 1000(-y)w \\ \dot{y} = -5000x - 2500(-w) \\ \dot{z} = -1000x^2 - 1000(-w)w \\ \dot{w} = -1000xy - 2500y - 1000zw - (-7000) \end{cases} \qquad (7)$$

The circuit model built according to equation (7) is shown in Fig.5. The four parts of the circuit correspond to four first nonlinear differential equations respectively, and the circuit design is realized by using operational amplifiers—LF347N chips, multipliers—AD633 chips, resistors and capacitors [19,20].

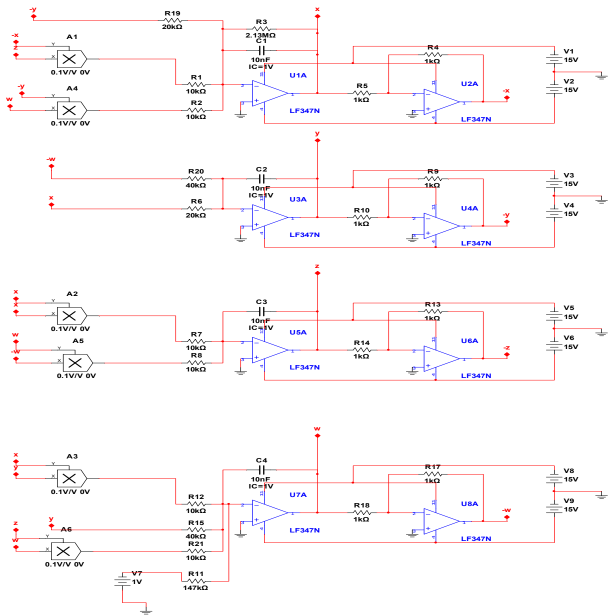


Fig. 5 Circuit diagram of system (1).

The equation of the four states realized by the circuit can be obtained from Fig.5. The result is such as equation (8):

$$\begin{cases} \dot{x} = -\dfrac{1}{R_3C_1}x - \dfrac{1}{R_{13}C_1}(-y) - \dfrac{1}{10R_1C_1}(-x)z \\ \quad -\dfrac{1}{10R_2C_1}(-y)w \\ \dot{y} = -\dfrac{1}{R_6C_2}x - \dfrac{1}{R_{20}C_2}(-w) \\ \dot{z} = -\dfrac{1}{10R_7C_3}x^2 - \dfrac{1}{10R_8C_3}w^2 \\ \dot{w} = -\dfrac{1}{10R_{12}C_4}xy - \dfrac{1}{R_{15}C_4}y - \dfrac{1}{10R_{21}C_4}zw \\ \quad -\dfrac{1}{R_{11}C_4} \end{cases} \tag{8}$$

By equating equation (7) and equation (8), it can be obtained that:

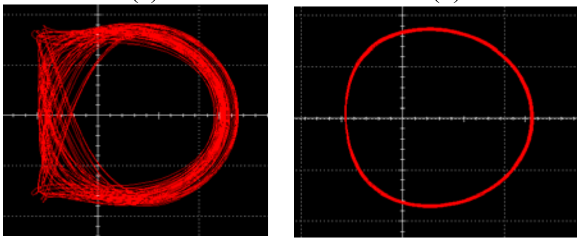
$$\begin{cases} \dfrac{1}{R_3C_1} = 47 \\ \dfrac{1}{R_{13}C_1} = \dfrac{1}{R_6C_2} = 5000 \\ \dfrac{1}{R_{20}C_1} = \dfrac{1}{R_{15}C_2} = 2500 \\ \dfrac{1}{10R_1C_1} = \dfrac{1}{10R_2C_1} = \dfrac{1}{10R_7C_3} = 1000 \\ \dfrac{1}{10R_8C_3} = \dfrac{1}{10R_{12}C_4} = \dfrac{1}{10R_{21}C_4} = 1000 \end{cases} \tag{9}$$

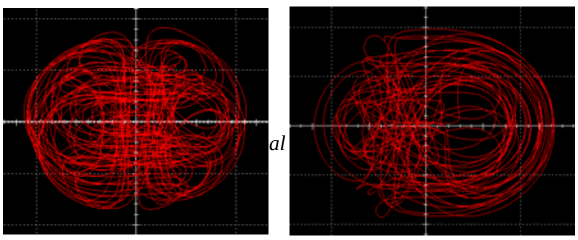
Let $C_1 = C_2 = C_3 = C_4 = 10nF$, then resistance values can be calculated:

$$R_3 = 2.13M\Omega,\ R_{13} = R_6 = 20K\Omega,\ R_{20} = R_{15} = 40K\Omega,$$
$$R_1 = R_2 = R_7 = R_8 = R_{12} = R_{21} = 10K\Omega$$

By the above circuit model design and built, perform circuit simulation experiment with Multisim. When c = 0.47, the phase portrait of system (1) observed on the virtual oscilloscope is shown in Fig.6(a). By the same method, when c = 0.7224, c = 0.76, c = 0.9, the phase portraits of system (1) observed on the virtual oscilloscope are shown in Fig.6(b), Fig.6(c), Fig.6(d) respectively. Compared with Fig.4, it can be seen that the circuit simulation results can consistent with the numerical simulation results.

(a) (b)





(c) (d)

Fig.6 Results of circuit simulation: (a) c = 0.47; (b) c = 0.7224; (c) c = 0.76; (d) c = 0.9.

4. Conclusion

In this paper, an innovative four-dimensional chaotic system is proposed. Theoretical analysis shows that this system is a dissipative system without any equilibrium point. The basic dynamic characteristics of the system are studied by using Lyapunov exponent spectrum and bifurcation diagram and so on. In addition, in order to illustrate the existence of chaotic attractor in the system from the physical level, Multisim software is used to design and implement the analog circuit of the chaotic system. The circuit simulation results are consistent with the numerical simulation results, so the physical realizability of the system is verified. There is no equilibrium point in this system, therefore the chaotic attractor generated by the system is hidden chaotic attractor, which is worthy of further study.

5. References

1. Lorenz EN. Deterministic non-period follow[J]. *Atoms Sci*. 1963, 20(2): 130-141.
2. Chua LO, Wu CW, Huang A, Zhong GQ. A universal circuit for studying and generating chaos[J]. *IEEE Trans Circuits Syst II*,1993,40:732-44.
3. Sprott J. Some simple chaotic flows[J]. *Physical Review E*,1994,50: R647-R650.
4. Chen GR, Ueta T. Yet another chaotic attractor [J]. *International Journal of Bifurcation and Chaos*, 1999, 9(7):1465-1466.
5. Lü Jinhu, Chen Guanrong. A new chaotic attractor coined[J]. *International Journal of Bifurcation and Chaos*, 2002, 12(3):659-661.
6. Liu WB, Chen GR. A new chaotic system and its generation[J]. *International Journal of Bifurcation and Chaos*, 2003, 13(1):261-267.
7. QI GY, Chen GR. On a new hyperchaotic system[J]. *Physics Letters A*, 2008, 372(2):124-136.

8. Li Qingdu. A simple method for finding topological horseshoes[J]. *International Journal of Bifurcation and Chaos*, 2010, 20(2):467-478.
9. Qi Guoyuan. Force Analysis of Qi Chaotic System[J]. *International Journal of Bifurcation and Chaos*, 2016,26(14):1560237.
10. Wang R, Sun H, Wang, J-Z, Wang L, Wang, Y-C. Applications of modularized circuit designs in a new hyper-chaotic system circuit implementation. *Chinese Physics B*, 2015,24(2),020501.
11. Xue Wei, Tan Dongcheng, Zhang Mei, Liu Shilong. A four-wing hyperchaotic system synchronization and its secure video communication based on FPGA[J]. *Journal of ShanDong University*, 2019,49(2):1672-3961.
12. Cang Shijian, Qi Guoyuan, Chen Zengqiang. A four-wing hyper-chaotic attroactor and transient chaos generated from a new 4-d quadratic autonomous system[J]. *Nonlinear Dyn*,2010,59:515-527.
13. Wei Xue, Hui Xiao, Xu Jinkang, Jia Hongyan. A Fractional-order Hyper-chaotic System and its Application in Image Encryption[J], *Journal of Tianjin University of Science & Technology*, 2015, 30(5):67-71.
14. Wang Yaqing, Zhou Shangbo. Image encryption algorithm based on fractional-order Chen chaotic system[J], *Journal of Computer Applications*, 2013, 33(4):1043-1046.
15. Yan Senlin. Chaotic laser parallel series synchronization and its repeater applications in secure communication[J]. *Acta Phys*. 2019,68(17): 170502.
16. Huan Zunxiang, Wu Mingchan. The application of fractal and chaos theory in astrophysics[J]. *Yunnan Observatory*,1996,1.
17. Bai Xiangzhong. The chaos analysis of rectangular plate under distributed load in transverse magnetic field[J]. *Journal of Dynamics and Control*,2006,03.
18. Hadi Jahanshahi, Amin Yousefpour, Zhouchao Wei, Raúl Alcaraz, Stelios Bekiros. A financial hyperchaotic system with coexisting attractors: Dynamic investigation, entropy analysis, control and synchronization [J]. *Chaos Solitons and Fractals*, 2019, 126:66-77.
19. V.T. Pham, S. Jafari, C. Volos, A. Giakoumis, S. Vaidyanathan and T. Kapitaniak. A chaotic system with equilibria located on the rounded square loop and its circuit implementation[J]. *IEEE Transactions on Circuits and Systems*: Express Briefs, 2016,63(9): 878-882.
20. Yang Xiaosong, Li Qingdu, Chen Guanrong. A twin-star hyperchaotic attractor and its circuit implementation[J]. *International Journal of circuit theory and applications*, 2003, 31(6):637-640.

Research on Synchronous Control of a Novel 4D Dissipative Chaotic System

Jiaqi Chen*, Yongchao Zhang, Yan Sun
Department of Automation, Tianjin University of Science and Technology,
1038 Dagunanlu Road, Hexi District, Tianjin 300222, PR China

*E-mail: * xwandcjq2019@126.com*
www.tust.edu.cn

Abstract

A novel four-dimensional dissipative chaotic system is presented in this paper, which with six nonlinear terms, three variable system parameters and one external excitation input. Through theoretical analysis and numerical simulation analysis of the chaotic characteristics of the system, a self-synchronization of the chaotic system is realized by the method of nonlinear synchronous control strategy. Numerical results show that nonlinear synchronous controller is correct and effective.

Keywords: Chaotic system; Numerical simulation analysis; Synchronous control

1. Introduction

Chaos is a general form of motion in nature, it is an irregular motion in a determined system, and its basic feature is the extreme sensitivity to the initial conditions [1]. In 1963, an American meteorologist Edward Lorenz discovered the Lorenz system in his study of the atmospheric convective model [2]. Since then, people have begun to explore and study chaos, and some new systems are proposed by researchers, such as the Chua system [3], the Rössler system [4], the Chen system [5], the Lü system [6], the Liu system [7], and the Qi system [8], etc.

Chaotic synchronization is a kind of generalized chaotic control, which means that the driving system and the response system vibrate in a synchronous manner. The dynamic behavior of the two chaotic systems reaches the same after vibration under different initial conditions. The control and synchronization of chaotic systems have potential application value in physics, chemistry, ecosystem, biological system and so on. Therefore, many scientists have been committed to chaotic synchronization control in recent decades. There are many synchronization control methods have been proposed, such as feedback control [9], adaptive control and synchronization [10], lag control and synchronization [11], pulse synchronization [12], phase synchronization [13] and synovial synchronization [14], etc.

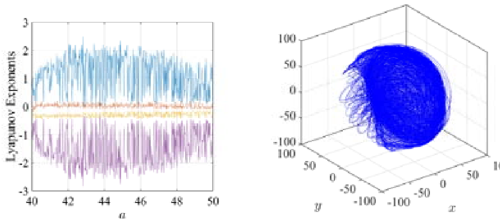
We construct a novel 4D dissipative chaotic system in this paper, which consists of six nonlinear terms, three variable parameters and one fixed value. The dynamic behavior of the system is theoretical analyzed. Based on the Lyapunov stability theory, the feedback synchronization control of the system is carried out by adding four nonlinear feedback controllers. The control results are simulated on MATLAB, and provides a certain reference for the practical application of the system.

2. A Novel 4D Dissipation System Analysis

The novel 4D dissipative chaotic system is proposed, and its equation of motion as follow:

$$\begin{cases} \dot{x} = -cx + ay + xz + yw \\ \dot{y} = -ax + bw \\ \dot{z} = -x^2 + w^2 \\ \dot{w} = -xy - by - zw + d \end{cases} \tag{1}$$

Where x, y, z, w represent the state variables of the system in (1). a, b, c represent the adjustable parameters of the system, d is a constant external input value. The Lyapunov exponent spectrum is made in the range of $a \in [40, 50]$ with the parameter a, as shown in Fig.1(a). When the system parameters are selected as $a = 42.82$, $b = 20$, $c = 0.6$, $d = 50$, the four Lyapunov exponents of the system are $L_1 = 2.502$, $L_2 = 0.0155 \approx 0$, $L_3 = -0.2781$, $L_4 = -2.843$ respectively. At this time, the distribution of the characteristic value is (+, 0, -, -). The system is in a chaotic state, and the chaotic attractor phase portrait is shown in fig. 1(b).



(a) (b)

Fig. 1. (a) Lyapunov exponents spectrum;
(b) phase portrait of 4D system (1)

3. Dynamic Characteristic Analysis

3.1. Initial Value Sensitivity Analysis

The change of chaotic system depends on the initial conditions very much. The initial value sensitivity analysis of the system (1) is carried out by selecting two groups of initial values with slight difference, which are as follows: (1, -2, 3, -4) and (1.001, -2, 3, -4). Under these two initial conditions, the MATLAB numerical simulation is carried out. The corresponding time series of x, y, z, w are shown in Fig. 2 (a), (b), (c), (d) respectively.The red curve represents the time series formed by the initial value (1,-2, 3,-4), and the blue curve represents the initial value (1.0001, -2, 3,-4).

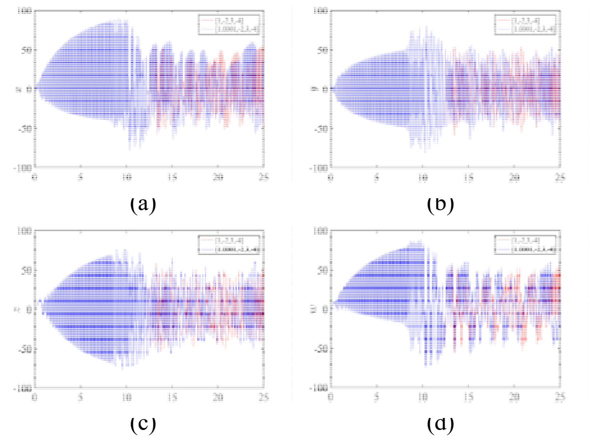


Fig. 2. Time series plot of system (1)

It can be seen from Fig. 2 that the difference between the two initial values is 0.0001 in the x direction when other three directions remain unchanged. The difference of the state sequence diagram between the two initial conditions is relatively small in the initial stage, but with the passage of time, the time series will soon show great differences, and the difference become larger and larger, indicating that the system has better initial value sensitivity.

3.2. Analysis of Dissipation and Symmetry

The divergence of dissipative system (1) is as follows:

$$\nabla V = \frac{\partial \dot{x}}{\partial x} + \frac{\partial \dot{y}}{\partial y} + \frac{\partial \dot{z}}{\partial z} + \frac{\partial \dot{w}}{\partial w} = -c \tag{2}$$

When the parameter $c > 0$, and $c = 0.6$, then $\nabla V = -0.6 < 0$. It is indicating that the system (1) is dissipative. And then, transforming coordination of the system (1), substitute (x, -y, -z, -w), (-x, y, -z, -w), (-x, -y, z, -w), (-x, -y, -z, w) into the mathematical model of the system (1), respectively, It is found that the system equations can be changed, so the system cannot be axisymmetric with x, y, z, and w, respectively.

3.3. Equilibrium Point Analysis

Making the left side of the dissipative system (1) is equal to 0, the formula (3) is as follow:

$$\begin{cases} -cx + ay + xz + yw = 0 \\ -ax + bw = 0 \\ -x^2 + w^2 = 0 \\ -xy - by - zw + d = 0 \end{cases} \tag{3}$$

If $-x^2 - w^2 = 0$ holds, then $x = \pm w$ or $x = w = 0$. Using $x = w = 0$ to verify that when the parameter a, b is any value, $-ax + bw = 0$ is constant. And only when $y = 0$, $-cx + ay + xz + yw = 0$ is true, then $-xy - by - zw + d$ can be simplified to d and $d = 50$, but $-xy - by - zw + d \neq 0$. So, the system has no equilibrium point.

4. Nonlinear Feedback Synchronous Control

Nonlinear feedback synchronization control refers to the design of a nonlinear synchronization controller between the response system and the driving system, so that the response system and the driving system achieve synchronization, finally.

The driving equation is transform from system (1), as shown in system (4):

$$\begin{cases} \dot{x}_1 = -cx_1 + ax_2 + x_1x_3 + x_2x_4 \\ \dot{x}_2 = -ax_1 + bx_4 \\ \dot{x}_3 = -x_1^2 + x_4^2 \\ \dot{x}_4 = -x_1x_2 - bx_2 - x_3x_4 + d \end{cases} \tag{4}$$

Where $a = 42.82$, $b = 20$, $c = 0.6$, $d = 50$. x_1, x_2, x_3, x_4 are the state variables of the driving system.

Establishing the response system as shown in the formula (5):

$$\begin{cases} \dot{y}_1 = -c_1 y_1 + a_1 y_2 + y_1 y_3 + y_2 y_4 + u_1 \\ \dot{y}_2 = -a_1 y_1 + b_1 y_4 + u_2 \\ \dot{y}_3 = -y_1^2 + y_4^2 + u_3 \\ \dot{y}_4 = -y_1 y_2 - b_1 y_2 - y_3 y_4 + d + u_4 \end{cases} \quad (5)$$

Where $a_1 = 32.82$, $b_1 = 10$, $c_1 = 0.1$, $d = 50$, y_1, y_2, y_3, y_4 are the state variables of the response system. $U = (u_1, u_2, u_3, u_4)^T$ is the synchronization controller.

The synchronous error signal of the response system (5) and the driving system (4) is as follow:

$$\begin{cases} e_1 = y_1 - x_1 \\ e_2 = y_2 - x_2 \\ e_3 = y_3 - x_3 \\ e_4 = y_4 - x_4 \end{cases} \quad (6)$$

The error system is as follow:

$$\begin{cases} \dot{e}_1 = (y_3 - c)e_1 + (a_1 + y_4)e_2 + x_1 e_3 \\ \quad + x_2 e_4 + k y_1 - g x_2 + u_1 \\ \dot{e}_2 = -a e_1 + b_1 e_4 + g y_1 - h x_4 + u_2 \\ \dot{e}_3 = -(x_1 + y_1)e_1 + (y_4 + x_4)e_4 + u_3 \\ \dot{e}_4 = -y e_2 - x_2 e_1 - b_1 e_2 - y_3 e_4 - x_4 e_3 \\ \quad + h x_2 + u_4 \end{cases} \quad (7)$$

Where $g = a - a_1 = 10$, $h = b - b_1 = 10$, $k = c - c_1 = 0.5$. Selecting the nonlinear feedback controller is as follow:

$$\begin{cases} u_1 = -y_3 e_1 - (a_1 + y_4)e_2 - x_1 e_3 - x_2 e_4 \\ \quad - k y_1 + g x_2 \\ u_2 = a e_1 - b_1 e_4 - g y_1 + h x_4 - m e_2 \\ u_3 = (x_1 + y_1)e_1 - (y_4 + x_4)e_4 - n e_3 \\ u_4 = y e_2 + x_2 e_1 + b_1 e_2 + y_3 e_4 + x_4 e_3 \\ \quad - h x_2 - f e_4 \end{cases} \quad (8)$$

Where $a = 42.82$, $b = 20$, $c = 0.6$, $m = 5$, $n = 4$, $f = 3$. Letting the formula (8) is substituted into the formula (7), the error system is as follow:

$$\begin{cases} \dot{e}_1 = -c e_1 = -0.6 e_1 \\ \dot{e}_2 = -m e_2 = -5 e_2 \\ \dot{e}_3 = -n e_3 = -4 e_3 \\ \dot{e}_4 = -f e_4 = -3 e_4 \end{cases} \quad (9)$$

The synchronous error system represented by the formula (9) is a linear time-invariant system with an equilibrium point, $S_0 = (0, 0, 0, 0)$, and the Lyapunov function is selected:

$$V(e) = 0.5e_1^2 + 0.5e_2^2 + 0.5e_3^2 + 0.5e_4^2 \quad (10)$$

Differentiating the formula (10) and substituting into the formula (9):

$$\begin{aligned} \dot{V}(e) &= e_1 \dot{e}_1 + e_2 \dot{e}_2 + e_3 \dot{e}_3 + e_4 \dot{e}_4 \\ &= -0.6e_1^2 - 5e_2^2 - 4e_3^2 - 3e_4^2 \end{aligned} \quad (11)$$

It can be seen from the formula (11) that $V(e) = -0.6e_1^2 - 5e_2^2 - 4e_3^2 - 3e_4^2 \leq 0$, the equal sign is established only when $e_1 = e_2 = e_3 = e_4 = 0$, and the visible $\dot{V}(e)$ is negative. It can be known from the Lyapunov stability theory that the error system equation (9) is asymptotica-lly stable, so $e_1 \to 0$, $e_2 \to 0$, $e_3 \to 0$, $e_4 \to 0$. Thus, the designed synchroniza-tion controller should be able to synchronize the response system (5) with the driving system (4).

The driving system and response system of synchronous circuit are built by Simulink, and the error system is obtained by subtract the output of driving system from the response system. By inputting specific values, select the different initial conditions $[x_1(0), x_2(0), x_3(0), x_4(0)] = [20, -15, 16, -10]$ and $[y_1(0), y_2(0), y_3(0), y_4(0)] = [-10, 15, -15, 20]$. The synchronization error curve, synchronization phase diagram and synchroniza-tion sequence diagram are obtained by simulation.

The convergence curve of the synchronization error of the system is shown in Fig. 3, and it can be seen from the diagram that the controller can make the system synchronize quickly.

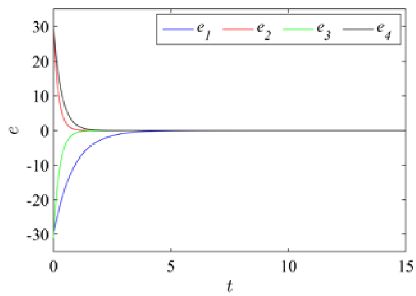


Fig. 3. Evolution of synchronization errors $e_1(t)$, $e_2(t)$, $e_3(t)$, $e_4(t)$

Fig. 4 is the synchronous phase diagram of the system. We can see that the four phase trajectories of $x_1 - y_1$, $x_2 - y_2$, $x_3 - y_3$ and $x_4 - y_4$ are 45° straight lines across the origin through the phase diagram, indicating that the system can achieve the synchronization state.

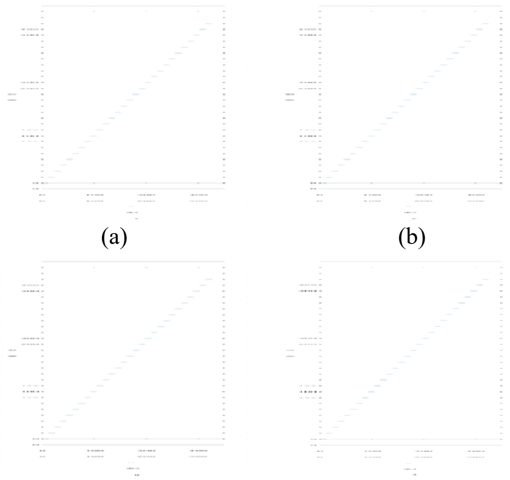


Fig. 4. Synchronization phase portrait of system (4) and system (5)

Fig. 5 is the time response diagram of the response system (4) and the driving system (5). By selecting the appropriate controller parameters, the response system and the driving system have their own motion trajectory, but with the increase of time, the motion tracks begin to coincide, which the system achieves self-synchronization quickly in a short time.

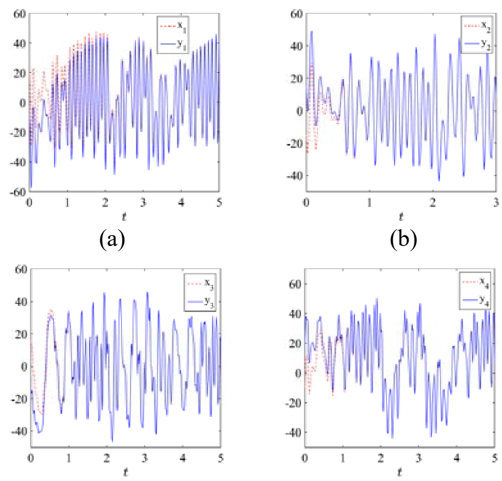


Fig. 5. Evolution of variables (a) $x_1(t)$, $y_1(t)$; (b) $x_2(t)$, $y_2(t)$; (c) $x_3(t)$, $y_3(t)$; (d) $x_4(t)$, $y_4(t)$

5. Conclusion

In this paper, the basic dynamic characteristics and numerical simulation analysis of a novel proposed 4D chaotic system are carried out, and its chaotic characteristics, initial value sensitivity and dissipation are verified. A nonlinear four-term feedback controller is designed to control the chaotic system synchronously. By setting the appropriate parameters, the system can achieve the synchronization state quickly when the initial conditions are quite different. It provides a reference for the further study of the system.

References

[1] G. Wang, H. He, A new Rössler hyper-chaotic system and realization with systematic circuit parameter design, *Chinese Phys. B*, 2008, 17 (11) : pp.4014-4021.

[2] E.N. Lorenz, Deterministic non-periodic flow, *Atmos. Sci.*, 1963, 20 (2): pp.130-141.

[3] T. Matsumoto, L.O. Chua, M. Komuro, The double scroll, *IEEE Trans. Circuit Syst.*, 1985, 32 (8): pp.797–818.

[4] O.E.Rössler, An equation for continuous chaos, *Phys. Lett. A*, 1976, 57(5): pp.397-398.

[5] G. Chen, T.Ueta, Yet another chaotic attractor, *Int. J. Bifurc. Chaos*, 1999, 9 (07): pp.1465-1466.

[6] J. Lü, G. Chen, A New Chaotic Attractor Coined, *Int. J. Bifurc. Chaos*, 2002, 12(03): pp.659-661.

[7] C. Liu, T. Liu, L. Liu, K. Liu, A new chaotic attractor, *Chaos, Solitons and Fractals*, 2004, 60(02): pp.1031-1038.

[8] G. Qi, G. Chen, S. Du, Z. Chen, Z. Yuan, Analysis of a new chaotic system, *Physica A* 2005, 353: pp.295-308.

[9] Z Li, J N Yu, X L An, J K Peng. Nonlinear feedback Synchronization Control of a New Chaotic System, *Journal of Wuhan University of Technology (Transportation Science & Engineering)*, 2009, 33 (4): pp.811-814.

[10] X Y Wang, Y Q Wang. Adaptive control for synchronization of a four-dimensional chaotic system via a single variable, *Nonlinear Dynamics*, 2011, 65 (3): pp.311-316.

[11] M J Wang, X Y Wang. A secure communication scheme based on parameter identification of first order time-delay chaotic system, *Journal of Physics*, 2009, 58 (3): pp.1467-1472. (in Chinese)

[12] T D Ma, W B Jiang, J Fu. Impulsive synchronization of fractional order hyperchaotic system based on comparison system, *Journal of Physics*, 2012 (9): pp.39-44. (in Chinese)

[13] J Meng, X Y Wang. Phase synchronization of chaotic systems based on nonlinear observers, *Journal of Physics*, 2007, 56 (9): pp.5142-5148. (in Chinese)

[14] G Pan, J Wei. Design of an adaptive sliding mode controller for synchronization of fractional- order chaotic system, *Journal of Physics*, 2015, (4): pp.45-51. (in Chinese)

EEG classification based on common spatial pattern and LDA

Lei Wang [1]*, Zixuan Li[2]
([1]College of Electronic Information and Automation, Tianjin University of Science and Technology, China
[2]International Business College, Dongbei University of Finance & Economics)
**wanglei404402644@outlook.com ; tjlzx@qq.com*

Abstract

In this paper, common spatial pattern (CSP) and linear discriminant analysis (LDA) are used to extract and classify motor imagery (MI) EEG signals, respectively. That is, according to the physiological phenomena of event related desynchronization (ERD) and event related synchronization (ERS) of MI-EEG signal, the optimal eigenvector representing MI-EEG signal was selected, and its characteristics was further extracted by using common spatial pattern (CSP). Then, the MI-EEG signals was classified by adopting linear discriminant analysis (LDA). At last, according to the data Ⅲ provided by Graz University, we studied the recognition accuracy of left and right hand, and the result is 80%. It shows that LDA classifier can be applied to classify the eigenvector extracted from CSP, and has a good recognition accuracy.

Keywords: EEG; motor imagery (MI); ERD/ERS; common spatial pattern (CSP); linear discriminant analysis (LDA).

1. Introduction

With the rapid development of Computer technology and the continuous deepening of Brain science research, people begin to try to establish a new direct communication and control pathway between the Brain and the external world, which is independent of peripheral nerves and muscles, and it is called as brain-computer Interface (BCI)[1].One of the important uses of this new mode of communication and control is to provide people with normal thinking and dyskinesia a way to communicate and control with the external environment [2]. The research field of BCI involves many subjects, so many complex problems need to be solved. Among them, feature extraction and classification are one of the keys. Feature extraction involves how to extract a small amount of useful information from control signals, and classification refers to how this information is distinguished to identify different brain states [3].

Motor imagery (MI) is one of the important control methods in brain computer interface technology. This method refers to that people can imagine the movement of limbs through the brain, collect, record and calculate the EEG signals, and translate the thinking activities of the brain into the instructions that can be executed by the computer, so as to realize the function of brain and external intelligent equipment interaction and control [4]. Compared with SSVEP and P300, motor imagination does not need any device to provide stimulation, and is not limited by different people's sensitivity to stimulation. [5-6].

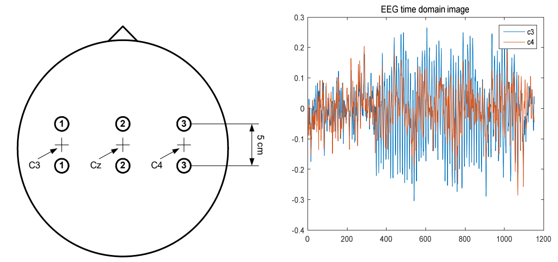
In motor imagery EEG signals, the main methods of feature extraction are Fourier transform (FFT), approximate entropy (ApEn), common spatial pattern (CSP), and so on. The main classification methods are linear discriminant analysis (LDA), support vector machine (SVM), and so on. [7-11]. This study attempts to apply CSP algorithm to extract the feature of two-dimensional motor imagery EEG signals, and combine it with the linear discriminant analysis (LDA) method to classify and recognize them. Compared with the previous SVM classifier method, the method in this study improves the classification accuracy of EEG signals and speeds up the signal processing speed.

2. Principles of EEG signals in motor imagery

2.1 The principle of the ERD / ERS characteristics

Human EEG signal frequency range is 0.5~50Hz, including delta wave (0.5~3Hz), theta wave (4~7Hz), alpha wave (8~13Hz), beta wave (14~30Hz)， gamma wave (31~50Hz), and mu rhythm (8~13Hz) and beta rhythm are closely related to motor imagery EEG signal [12]. Studies have shown that when people are doing unilateral limb movement or exercising imagination, the rhythm amplitude of mu/beta in the ipsilateral motor sensory area of the brain will increase, generating event-related synchronization（ERS）, while the rhythm amplitude of mu/beta in the contralateral side will decrease, generating event-related desynchronization(ERD)[13].

According to this physiological phenomenon, this paper will process data set Graz2003iii MI-EEG from Austria. This data set completely records the EEG signals of C3, CZ and C4 bipolar EEG channels. The electrode position is shown in Figure 1, and the EEG signals of C3 and C4 channels in one experiment are shown in Figure 2.



（a）Electrode location map (b) EEG signal map

Fig.1 Electrode position and signal diagram

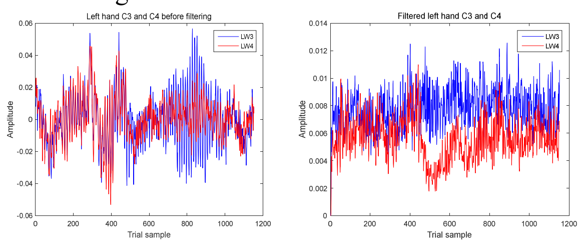
2.2 Verification of ERS / ERD characteristics

Suppose $x_{(i,\ j)}$ is the data of each experiment, where i represents the number of experiments and j represents a certain data point of experiments, the average energy can be expressed as,

$$W = \frac{1}{N}\sum_{i=1}^{N} x^2_{(i,\ j)} \qquad (1)$$

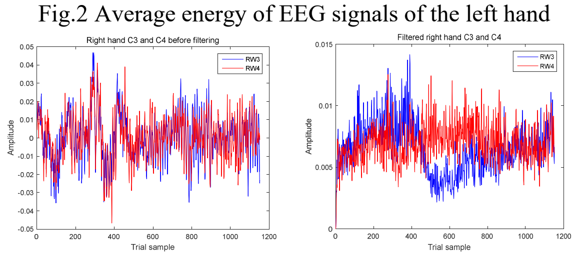
Here, N=70, i=1,2,3,⋯,140, j=1,2,3, ⋯,1152.

In order to better observe the changes of power of left and right hand, let LW_3、 LW_4, and RW_3、 RW_4 respectively represent the average instantaneous power of C3 and C4 EEG signals when the left and right hands are imagined to move for 70 times. The changes before and after filtering of LW_3、 LW_4 are shown in figure 2. The changes before and after filtering of RW_3、 RW_4 are shown in figure 3.



(a)Before filtering (b) After filtering

Fig.2 Average energy of EEG signals of the left hand



(a) Before filtering (b) After filtering

Fig.3 Average energy of EEG signals of the right hand

It can be clearly seen from Figures 3 and 4 that in the filtered motor imaging EEG signal, an obvious ERD / ERS phenomenon occurs between 3s and 9s. We find that the signal of C4 channel appears obvious ERD phenomenon, which shows the subject are imagining left hand movement. Similarly, the signal of C3 channel appears obvious ERD phenomenon, which shows the subjects are imagining right hand movement. Therefore, we select the signal between 3s and 9s as experiment signals, and uses CSP method for feature extraction and LDA classifier for classification

3. Experimental

3.1 Experimental steps

The experimental steps are as follows:

(1) Based on diagonalization of the matrix of the pre-processing training set, create an optimal spatial filter, which maximizes the difference between the variance values of the two kinds of signals and obtains the most distinguishable eigenvector.

(2) Train LDA classifier by using the eigenvectors of EEG signals of training set.

(3) Extract characteristic of the pre-processing test set by using CSP.

(4) Classify the feature vectors of EEG signals in the test set. And the flow of the method is shown in Figure 4.

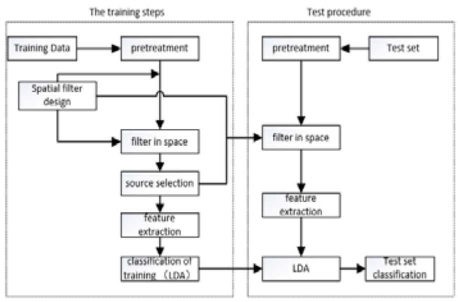


Fig.4 Method flow chart

3.2 CSP feature extraction based on ERD/ERS characteristics

3.2.1 Find the mixed spatial covariance matrix of two kinds of data

Suppose there are Z channels on the device to record EEG signals, and each channel collects N discrete signals, so the collected data can be regarded as a matrix E_i with Z × N, where i represents a certain experiment. And the standard space covariance matrix can be written as,

$$C_i = \frac{E_i E_i^T}{trace(E_i E_i^T)} \qquad (2)$$

Where E_i^T the transpose of isE_i, $trace(E_i E_i^T)$is the sum of the main diagonal of the matrix$E_i E_i^T$. So the average covariance matrix of these two types of movements imaginary can be written as,

$$C_l = \frac{1}{l}\sum_{i=1}^{l} C_i^l,\ \ C_r = \frac{1}{r}\sum_{i=1}^{r} C_i^r \qquad (3)$$

Where l and r represent the number of samples of left hand and right hand in the training set, respectively

3.2.2 Find out the whitening eigenvalue matrix P

According to matrix theory, combining C_l and C_r, transformation is carried out,

$$C = C_l + C_r = B\lambda B^T \quad (4)$$

Where B is the eigenvector matrix of combinatorial space covariance matrix C, B^T is the transpose of B, λ is a diagonal matrix composed of non-zero eigenvalues, and is arranged in descending order according to the elements on the diagonal. And then can obtain albinism matrix,

$$P = \sqrt{\lambda^{-1}}B^T \quad (5)$$

3.2.3 Tectonic spatial filtering

According to (4), it can be obtained,

$$S_l = PC_lP^T, S_r = PC_rP^T \quad (6)$$

$$S = S_l + S_r = PCP^T = I \quad (7)$$

Which proved that C_l and C_r have common eigenvector. Therefore, it can be assumed that,

$$S_l = D\lambda_1 D^T, S_r = D\lambda_2 D^T, \lambda_1 + \lambda_2 = I \quad (8)$$

Where the common feature vector of S_l and S_r is D, and I is the identity matrix. In order to maximize the difference between the two types of tasks in spatial arrangement, the corresponding m maximum eigenvalues and the m minimum eigenvalues in λ_1 are selected to construct D_l and D_r, and then the spatial filters for the two types of tasks are obtained as follows,

$$W_l = D_l^T P, W_r = D_r^T P, \ W = [W_l; W_r] \quad (9)$$

3.2.4 Feature extraction

Then the EEG data E_i collected in each experiment were filtered by the common space mode to get Z_i, that is $Z_i = WE_i$. For the transformed signal Z_i (i=1,2,⋯,2n).

$$f_i = \lg\left[\frac{var(Z_i)}{\sum_{j=1}^{2n} var(Z_i)}\right] \quad (10)$$

Where $var(.)$ represents variance, f_i represents features. The logarithm operation is to make the eigenvectors f_i closer to the normal distribution.

3.3 Principle of linear discriminant classification algorithm

Linear discriminant analysis (LDA), also known as Fisher linear discrimination (FLD), is a classic algorithm of pattern recognition [14]. The basic idea is to project the labeled data points into the space with lower dimension by projection, so that the projected points can be distinguished by category, and then ensure that the samples have the maximum distance between classes and the minimum distance within classes in the new subspace [15].

The dispersion matrix in a sample class can be defined as,

$$S_W = \sum_{i=1}^{C} P_i S_i \quad (11)$$

S_i can be defined as:

$$S_i = \sum_{x \in C_i}(x - m_i)(x - m_i)^T, \text{(i=1, 2… C)} \quad (12)$$

Among them, P_i is a prior probability, and m_i is the mean value of C_i. The inter class dispersion matrix of samples can be defined as: $S_b = \sum_{i=1}^{C} P_i (m_i - m)(m_i - m)^T$. m is the mean of all samples. It can be deduced from the above formula that if it is a non-anisotropic matrix, the projection effect can be obtained. Generally, the greater the dispersion is, the more obvious the projection effect is. The smaller the intra class dispersion is, the better the projection effect is [16-17]

The Fisher criterion function can be defined as follows

$$J(W_{opt}) = \frac{W^T S_b W}{W^T S_W W} \quad (13)$$

W is an n-dimensional column vector. Fisher's linear discriminant analysis is to select the vector W that maximizes $J(W_{opt})$ as the projection direction. Its physical meaning is that the projected samples have the maximum inter class dispersion and the minimum intra class dispersion [18]. The dividing line of Fisher classification is shown in Figure 5.

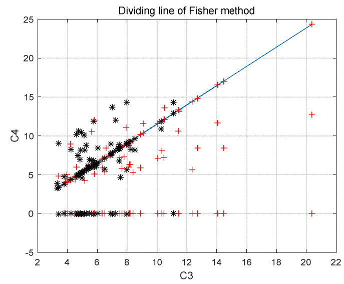


Fig.5 Dividing line of Fisher classifier method

3.4 Result analysis

The final classification result is shown in Figure 6. After calculation, the recognition rate of EEG signals of left and right hand is 80%. The results of each method are compared as shown in Figure 7, it can be seen that the recognition rate is significantly improved, compared with the previous signal processing methods. (These methods are not described in detail and they use the same data set).

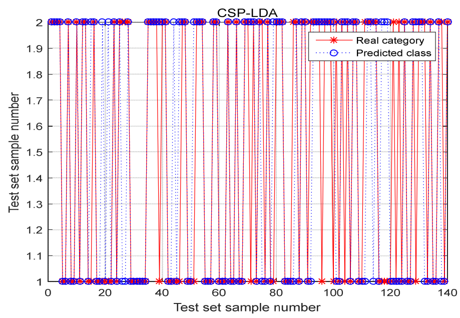


Fig.6 Classification results of EEG signals

Method	ApEn-SVM	CSP-SVM	FFT-SVM	CSP-LDA
Recognition rate:	41.06%	47.85%	75%	80%

Fig.7 Recognition rate of several methods

4. Conclusion

Through the data analysis provided by Graz University, the feasibility of CSP-LDA method is verified. Compared with CSP-SVM method, it has a great improvement in the final classification results. Compared with ApEn-SVM and FFT-SVM method, it also has good classification results.

5. References

1. CHENG Ming, REN Yu-peng, GAO Xiao-rong, et al. Key Technique on Controlling Rehabilitation Robot by Electroencephalographic [J].Robot Technique and Application (S1004-6437), 2003, 10 (4):45-48.
2. Blankertz B, Muller K R, Curio G, et al. The BCI competition2003: progress and perspectives in detection and discrimination of EEG single trials [J]. IEEE Trans on Biomedical Eng. (S0018-9294), 2004, 51(6): 1044-1051.
3. Yang Banghua, Yan Guozheng, Ding Guoqing, Yu Lianzhi. Research on Key Technologies of brain computer interface [J]. Beijing Biomedical Engineering, 2005, 24 (4): 308:310
4. Bhattacharyya S, Sengupta A, Chakraborti T, et al. Automatic feature selection of motor imagery EEG signals using differential evolution and learning automata.[J]. Medical & Biological Engineering & Computing, 2014, 52(2):131.
5. Neumann N, Kubler A. Training locked-in patients: A challenge for the use of brain-computer interfaces [J]. IEEE Trans onNeural Systems and Rehabilitation Eng (S1534-4320), 2003,11(2): 169-172.
6. Hua Xiaomei, Lin Jiarui, Guan Jin'an. Research progress of brain computer interface [J]. Journal of Biomedical Engineering, 2004, 27 (2): 94-111
7. Quadrianto N, Guan Cun Tai, Dat TH, et al. Sub-band Common Spatial Pattern(SBCSP)for Brain-Computer Interface[A] In: 2007 3rd International IEEE EMBS Conference on Neural Engineering[C]. Piscataway, NJ, USA:IEEE, 2007.219- 225.
8. Peters, BO, Pfurtscheller G., Flyvbjerg H. Automatic differentiation or multichannel EEG signals[J]. Transactions on Biomedical Engineering 2001,48(1):111-116.
9. Lei Songze. Application of fast LDA Algorithm Based on K-means clustering in speech recognition [J]. Science and Technology Bulletin 2013
10. Lei Yun, Gong Kuihua. Medical image feature extraction based on linear moment and wavelet transform [J]. Science and technology information, 2010 (03)
11. Molina G.G..BCI adaptation using incremental-SVM learning[A].In: Proceedings of 3rd International IEEE EMBS Conference on Neural Engineering[C]. Piscataway, NJ, USA: IEEE,2007. 337- 341
12. Huang Sijuan, Wu Xiaoming. Extraction of EEG characteristics based on mu / beta rhythm imagination [J]. China tissue engineering research and clinical rehabilitation, 2010, 14 (43): 8061-8064.
13. Birbaumer N, Breaking the silence: Brain-computerinter-interfaces(BCI) for communication and motor control [J]. Psy -2006, 43 (6): 517 – 532.
14. Zhou Zhihua. Machine learning [M]. Tsinghua University Press, 2016.
15. L.F.Nicolas-Alonso and J. Gomez-Gil. Brain computer interfaces,a review.Sensor, 2012(12):1211-1279.
16. Thanh N, Khosravi A, Creighton D, et al. EEG signal classification for BCI applications by wacelets and interval type-2 fuzzy logic systems[J]. Expert Systems With Applications,2015, 42(9): 615-624.
17. Pfurtscheller G, Brunner C, Schlogl A, et al. Mu rhythm synchronization and EEG single-trial classification of different motor imagery tasks[J]. Neuro Image, 2006, 31:153-159.
18. Shi Yuan. Research on EEG data analysis based on Fisher discriminant method [J]. Microcomputer information, 2010, 7 (26): 226-228

Dynamic characteristics analysis of the Shimizu–Morioka chaotic system

Wenxin Shi, Hongyan Jia

Department of Automation, Tianjin University of Science and Technology,
1038 Dagunanlu Road, Hexi District, Tianjin 300222, PR China

E-mail: w.x.shi@qq.com

Abstract

In the paper, we investigate the Shimizu–Morioka chaotic system based on Numerical simulations and FPGA implementation. Firstly, the stability of three equilibrium points of the system is analyzed by using the Routh-Hurwitz criterion. Then the coexistence characteristics of the system are studied by using digital analysis methods such as the phase portrait diagram, Lyapunov exponents diagram and bifurcation diagram. Finally, the digital circuit of the system is realized by FPGA, and the feasibility of the system is illustrated.

Key words: the Shimizu–Morioka chaotic system; coexistence characteristics; FPGA

1. Introduction

Complex behavior of the Shimizu-Morioka model has been found by computer simulation in Ref [1]. Subsequently, it was found in Ref [2,3] that there were two types of Lorenz-like attractors in the Shimizu-Morioka model, one is orientable and the other is non-orientable. In recent decades, the Shimizu-Morioka system has gained interest due to its rich dynamical behavior and potential physical applications [4, 5-7,8,9]. Although some dynamical characteristics of the Shimizu-Morioka system have been shown in the literature [5], some other dynamic characteristics of the Shimizu-Morioka system remain to be studied, such as Multi-stability. Multi-stability means multiple coexistence of attractors in a nonlinear dynamic system and the attractors will lead to different behaviors for the same parameter values [10]. In recent years, the multi-stability of attractors has become a very hot research topic and coexistence of attractors has been found in several some systems, such as, a dynamical system, an electronic circuit, a laser system, a neural system and so on [11-19]. In this paper, we will study the coexistence characteristics of the Shimizu-Morioka system, give the causes of coexistence, and analyze other dynamic characteristics in the system.

Since Altera corporation launched DSP builder in 2002, it has solved the problem of designing anti-lock algorithm and its analog signal processing [20]. DSP builder be used to design the digital circuit of continuous chaotic system on FPGA platform. On the basis of numerical analysis, the FPGA implementation of the Shimizu–Morioka chaotic system is introduced, which provides a feasible method for engineering applications.

2. Stability analysis of the Shimizu–Morioka chaotic system

The Shimizu–Morioka system is described as:

$$\begin{cases} \dot{x} = y \\ \dot{y} = x - by - xz \\ \dot{z} = -az + x^2 \end{cases} \tag{1}$$

Where *x, y* and *z* are status variables, and *a, b* are system parameters. The system (1) is simple three-dimensional autonomous system, in which there is only one nonlinear cross term and one square term. The divergence of the system (1) is $\nabla f = -(a + b) < 0$ $(a, b > 0)$. Therefore, the system (1) is a dissipative system. This means that as time approaches infinity, the system (1) will eventually form an attractor.

2.1 Equilibrium point of the system (1)

By setting the right-hand side of the system (1) to zero, we can obtain that the system (1) has three equilibrium points:

$$\begin{cases} s_0(0,0,0) \\ s_1(\sqrt{a},0,1) \\ s_2(-\sqrt{a},0,1) \end{cases} \tag{2}$$

It shows that the three equilibrium points of the system (1) are only related to the system parameter *a*. The Jacobian matrix of the system (1) is

$$J_C = \begin{bmatrix} 0 & 1 & 0 \\ 1-z & -b & -x \\ 2x & 0 & -a \end{bmatrix} \tag{3}$$

When the equilibrium point is $s_0(0, 0, 0)$, the eigenvalues of the system (1) are $\lambda_1=-a$ and $\lambda_{2,3}=(-b\pm\sqrt{b^2+4})/2$, which indicates that the equilibrium point s_0 is the stable focus-point. When the equilibrium points are $s_1(\sqrt{a}, 0, 1)$ and $s_2(-\sqrt{a}, 0, 1)$, the characteristic equation of the system (1) is

$$f(\lambda) = \lambda^3 + (a+b)\lambda^2 + ab\lambda + 2a \tag{4}$$

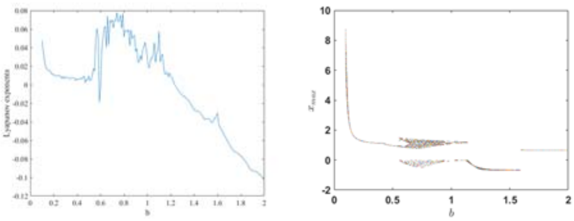
According to the classic Routh–Hurwitz stability criterion, we can get that the equilibrium points s_1 and s_2 are stable when $b + a > b/2$ (a, $b>0$). Otherwise, both equilibrium points are unstable.

2.2 Symmetry of the system (1)

Like Lorenz chaotic system, the Shimizu-Morioka chaotic system is unchanged under the coordinate transformation (x, y, z) → (–x, –y, z). It means that the system (1) is symmetric with respect to z-axis.

3. The dynamic characteristics of the Shimizu–Morioka chaotic system

Now, the system parameter *a* is fixed to 0.45, the dynamic behaviors of the system (1) are discussed by changing system parameter *b*. When the initial value is set to (1, 1, 1), Lyapunov exponent diagram and bifurcation diagram of the system (1) are shown in Fig. 1(a) and Fig. 1(b).

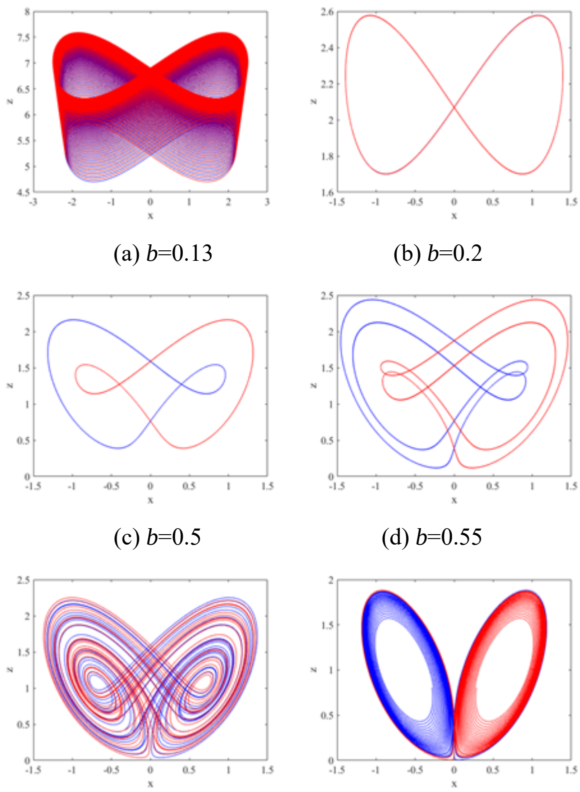


(a) Largest Lyapunov exponent diagram (b) Bifurcation diagram

Fig. 1. Lyapunov exponent diagram and Bifurcation diagram of system (1)

Fig. 1 (a) shows the maximum Lyapunov exponent. According to Fig. 1 (a), (b), it can be seen that when $b\in[0.1, 0.15)\cup(0.56, 1.2]$, the system (1) is in a chaotic state, when $b\in(0.15, 0.55)$, the system (1) presents a periodic state, and when $b\in(1.26, 2]$, the system (1) shows point attractor.

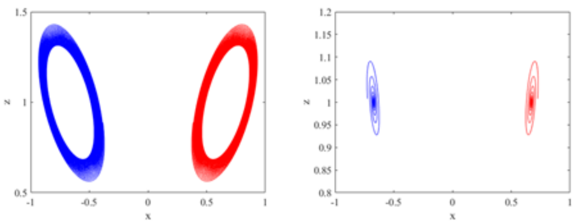
In addition, it is mentioned in the section 2 that symmetry will lead to attractors coexistence of the system (1). As shown in Fig. 2, when initial values of (1, 1, 1) and (-1, -1, 1) are selected, coexistence of different states occurs in the system (1) as the system parameter *b* changes.



(a) *b*=0.13 (b) *b*=0.2

(c) *b*=0.5 (d) *b*=0.55

(e) *b*=0.75 (f) *b*=1.15



(g) *b*=1.2 (h) *b*=1.5

Fig. 2. Coexisting phase diagram with variation of the system parameter *b*

As can be seen from Fig. 2, with the change of the system parameter *b*, the system goes from the chaotic attractor to the periodic attractor, then from the periodic attractor to the chaotic attractor, and finally from the chaotic attractor to point attractor. Fig. 2 (a), (e), (f) and (g) respectively show the coexistence of different chaotic attractors. Fig. 2 (b), (c) and (d) respectively show the coexistence of different periodic attractors. Fig. 2 (h) show the coexistence of different periodic attractors. However, it is worth noting that when the system parameters *b* is 0.75,1.15,1.2 and 1.5, the system (1) changes from a double-wing chaotic attractor to a single-wing chaotic attractor to point attractor.

4. FPGA implementation of the Shimizu–Morioka chaotic system

4.1 Discrete algorithm analysis

First, we take the Laplace transform of system (1), and get:

$$\begin{cases} L\{x(t)\} = \dfrac{1}{S} L\{y\} \\ L\{y(t)\} = \dfrac{1}{S} L\{x - by - xz\} \\ L\{z(t)\} = \dfrac{1}{S} L\{- az + x^2\} \end{cases} \tag{5}$$

Where *x*, *y* and *z* are functions of *t* and H(S)=1/S is continuous integer order operator. Then, the impulse response invariant method is used for discretization. The discrete transfer function is

$$H(Z) = \frac{1}{f} ? \frac{Z}{Z - 1} \tag{6}$$

Where, *f* is sample frequency. According to discrete transfer function (6), the model of integer order integral operator is established by using DSP Builder Library on Matlab/Simunlink platform, as shown in the Fig. 3.

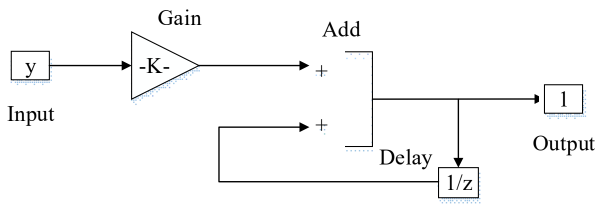


Fig. 3. H(S)=1/S of discrete structure diagram

In this way, the digital integrator can be embedded into the chaotic structure of the whole digital calculation, so as to realize the digital circuit model of the system (1).

4.2 FPGA implementation

In this paper, the software used is Quartus15.1 developed by Altera company and DSP builder. One of the main characteristics of the platform is that it can directly use Matlab/Simunlink to call each unit library in DSP builder library, which is convenient for circuit module design and FPGA hardware implementation [20,21].

Fig. 5 (a), (b) are a periodic attractor and a chaotic attractor observed on the oscilloscope in the experiment when the system parameters *a*=0.45, *b*=0.5, 0.75 and the initial value is (1, 1, 1). Compared Fig. 5 (a), (b) with Fig. 2 (c), (e), it can be seen that the experimental results are in good agreement with the numerical analysis results. Thus, the feasibility of the system (1) is illustrated

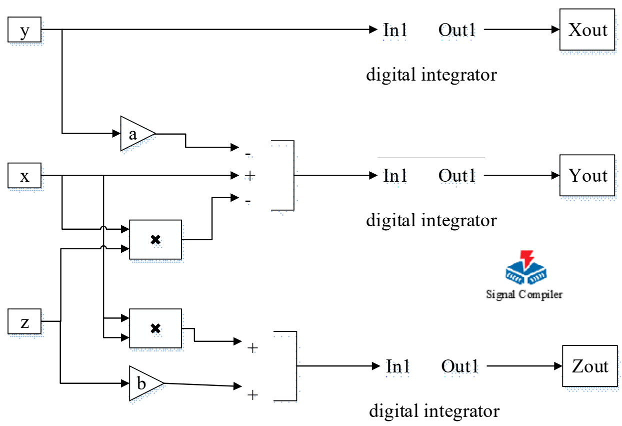
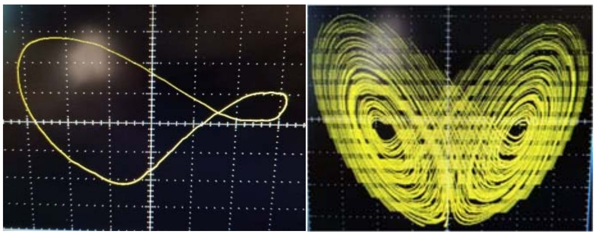


Fig. 4. FPGA implementation of the system (1)



(a) b=0.5 (b) b=0.75

Fig. 5. The results of FPGA implementation x-z phase plane

5. Conclusion

In this paper, firstly, the Shimizu–Morioka chaotic system with three equilibrium points is introduced and its stability is analyzed. Secondly, the dynamic characteristics of the system are discussed by means of bifurcation diagram, Lyapunov exponent diagram and phase diagram. Attractors coexistence of the system (1) and the reasons for coexistence are discussed. Finally, the digital circuit model of this system is implemented on the FPGA hardware platform, and the experimental results are consistent with the numerical analysis results, which further explains the feasibility of this system and provides a method for the realization of this system.

References

[1] T. Shimizu and N. Morioka, Phys. Lett. A76 (1980) 201.

[2] A.L. Sil'nikov, Sel. Math. Sov. 10 (1991) 105 (originally published in: Methods of Qualitative Theory of Differential Equations, Gorky St. Univ. (1986), p. 180).

[3] A.L. Sil'nikov, Methods of Qualitative Theory and Theory of Bifurcations, Gorky St. Univ. (1989), p. 130.

[4] Kokubu, H., Roussarie, R.: Existence of a singularly degenerate heteroclinic cycle in the Lorenz system and its dynamical consequences: Part I. J. Dyn. Differ. Equ. 16(2), 513–557 (2004)

[5] Shilnikov, A.L.: On bifurcations of the Lorenz attractor in the Shimizu-Morioka model. Physica D 62, 338–346 (1993)

[6] Shimizu, T., Morioka, N.: On the bifurcation of a symmet riclimit cycle to an asymmetric one in a simple model.Phys. Lett. A 76(3,4), 201–204 (1980)

[7] Tigan, G., Turaev, D.: Analytical search for homoclinic bifurcations in the Shimizu–Morioka model. Physica D 240(12), 985–989 (2011)

[8] Vladimirov, A.G., Volkov, D.Y.: Low-intensity chaotic operations of a laser with a saturable absorber. Opt. Commun.100, 351–360 (1993)

[9] Yu, S., Tang, W.K.S., Lü, J., Chen, G.: Generation of n × m-wing Lorenz-like attractors from a modified Shimizu–Morioka model. IEEE Trans. Circuits Syst. 55(11), 1168–1172 (2008)

[10] C. Li, J.C. Sprott, Finding coexisting attractors using amplitude control, Nonlinear Dynam. 78 (2014) 2059–2064.

[11] C. Li, J.C. Sprott, Multistability in the Lorenz system: A broken butterfly, Int. J. Bifur. Chaos 24 (2014) 1450131, http://dx.doi.org/10.1142/S0218127414501314.

[12] A.N. Pisarchik, U. Feudel, Control of multistability, Phys. Rep. 540 (2014) 167–218.

[13] C. Li, W. Hu, J.C. Sprott, X. Wang, Multistability in symmetric chaotic systems, Eur. Phys. J. Spec. Top. 224 (2015) 1493–1506.

[14] B.C. Bao, H. Bao, N. Wang, M. Chen, Q. Xu, Hidden extreme multistability in memristive hyperchaotic system, Chaos Solitons Fractals 94 (2017) 102–111.

[15] F. Yuan, G. Wang, X. Wang, Extreme multistability in a memristor-based multiscroll hyper-chaotic system, Chaos 26 (2016) 73107.

[16] X. Zhang, Z. Li, D. Chang, Dynamics, circuit implementation and synchronization of a new three-dimensional fractional-order chaotic system, AEU - Int. J. Electron. Commun. 82 (2017) 435–445.

[17] Chen, M., Xu, Q., Lin, Y. & Bao, B. [2017] "Multistability induced by two symmetric stable node-foci in modified canonical Chua's circuit," Nonlin. Dyn. 87, 789–802.

[18] Leonov, G. A. & Kuznetsov, N. V. [2013] "Hidden attractors in dynamical systems: From hidden oscillations in Hilbert–Kolmogorov, Aizerman, and Kalman problems to hidden chaotic attractor in Chua circuits," Int. J. Bifurcation and Chaos 23, 133002–133007.

[19] Pisarchik, A. N. & Feudel, U. [2014] "Control of multistability," Phys. Rep. 540, 167–218.

[20] WANG Guang-yi, BAO Xu-lei, WANG Zhong-lin. Design and FPGA implementation of a new hyperchaotic system[J]. Chinese physics B, 2008, 17(10): 3596-3602.

[21] LV Jin-hu, YU Xing-huo, CHEN Guan-rong, et al. Characterizing the synchronizability of small-world dynamical networks[J]. IEEE transactions on circuits and systemsI, 2004, 51(4):787-796.

Research on the Motion Track of High-speed Objects

Qianqian Zhang[1,2*], Fengzhi Dai [1,3], Jichao Zhao [1], Haokang Wen [1], Hongbo Hao [1]

[1] Tianjin University of Science and Technology, China;

[2] Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China;
3 Advanced Structural Integrity International Joint Research Centre, Tianjin University of Science and Technology, 300222, China

*E-mail: *577972810@qq.com*
www.tust.edu.cn

Abstract

Measuring the location coordinates of high-speed moving objects is of great significance in the strike accuracy test. Currently, there are a variety of measurement methods, such as the light screen target measurement and acoustic target measurement. The integrated control and verification workstation includes the calculation system and wireless transmission module, which is mainly responsible for the final modeling and analysis of the data collected by the system to obtain the calculation results and display them. "Acquisition station 1" and "acquisition station 2" are composed of acquisition computer, battery, DCDC module (secondary power supply), camera, trigger module and other parts. The main function of acquisition station is to analyze and process the data collected by the camera and upload them to the integrated control algorithm workstation. The purpose of this design is to develop a set of high speed object trajectory landing point prediction system, and realize the recording, processing, wireless uploading, human-computer interaction and other functions of high speed linear motion target video collected by 4 sets of high speed cameras set up at two points.

Keywords: motion tracking, high-speed moving objects, algorithm workstation, location coordinates

1. Introduction

In this paper, four cameras are arranged in a rectangular way. Among them, four cameras are placed on the ground in a rectangular form. The long side of the rectangle is about 60 meters, and the short side is about 20 meters. The two cameras form a collection station, which consists of a comprehensive control checking calculation workstation, named as collection station 1 and collection station 2. "Acquisition station 1" and "acquisition station 2" are composed of acquisition computer, battery, DCDC module (secondary power supply), camera, trigger module and other parts. The main function of acquisition station is to analyze and process the data collected by the camera and upload them to the integrated control algorithm workstation. The integrated control and verification workstation includes the calculation system and wireless transmission module, which is mainly responsible for the final modeling and analysis of the data collected by the system to obtain the calculation results and display them. It is a challenge for any optical method to measure object surfaces with a large range of reflectivity variations[1].

2. Communication unit

According to the analysis of the application site, the indoor calculation system is connected with the acquisition station, the communication distance is about 2km, and the communication frequency band of 2.4g is assumed. Therefore, 2.4g wireless networking technology is selected to network the calculation system and two collection stations, and control instructions are sent and data collection is reported within the LAN.

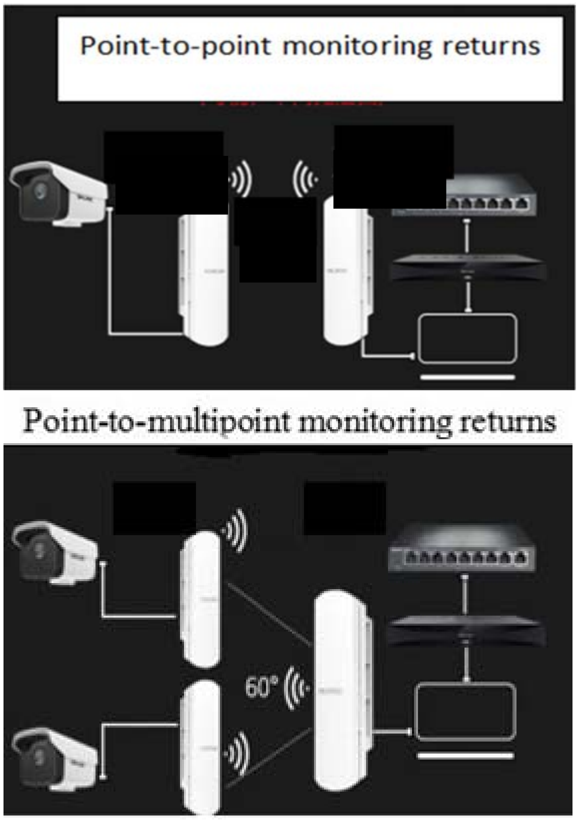


Fig.1. collection stations

Different selection of point-to-point 2.4/5g modules can achieve 2-15km visual distance directional communication. Due to the long transmission distance, the module adopts directional antenna technology, which greatly improves the communication distance, but the communication Angle is limited, generally ranging from 30° to 60° . The line covers the direction of the included Angle inside. Therefore, for the application of 2.4/5g technology, it is necessary to:

- Ensure the communication distance to see.
- The communication distance is less than 15km.
- The antenna coverage Angle can cover two sets of acquisition stations at the same time.

Compared with other schemes, the biggest advantage of adopting 2.4/5g scheme is that the whole system can be networked to form an independent LAN, which brings major benefits to the system.

The selection of system telecommunication technology has been considered as follows: Low-frequency Lora communication, 2.4/5g networking, 4G and other technologies. The accuracy of the measured gradient is low since the captured fringe pattern shows low signal to noise ratio[2].

(1) 4G networking technology has been proved to have the lowest environmental requirements and can realize communication between any distance and any obstacle. However, it is more troublesome to set up the system, which requires network card and regular renewal. And the use of third party server, long-term consideration service relationship is not reliable, delay is not controllable, stability is not controllable.

(2) Low frequency Lora communication (433), communication distance of more than 2km (meet the requirements), obstacle diffraction performance can meet the requirements of 1 building obstacle diffraction (meet the requirements).

(3) 2.4/5g network, communication distance of more than 2km (meet the requirements), obstacle diffraction performance can meet the requirements of obstacles of a building (meet the requirements).

In conclusion, both low-frequency Lora and 2.4/5g networking schemes can meet the needs of this case, but compared with 2.4g /5G technology, more mature Ethernet networking technology is adopted[2], and the communication protocol is more mature and stable. The communication between the acquisition computer and the camera in the system is also selected as Ethernet interface. Conclusion: unified technology (2.4/5g Ethernet technology) in the system will be more stable and more flexible.

3. Software

3.1 Software component

Each system needs three software, and the corresponding software is deployed in the collection station 1, collection station 2 and the comprehensive control algorithm workstation respectively. Among them, the software of acquisition station 1 and collection station 2 is mainly responsible for the camera control and data collection, and the original camera data is packaged and parsed into target information message, which is reported to the calculation system.

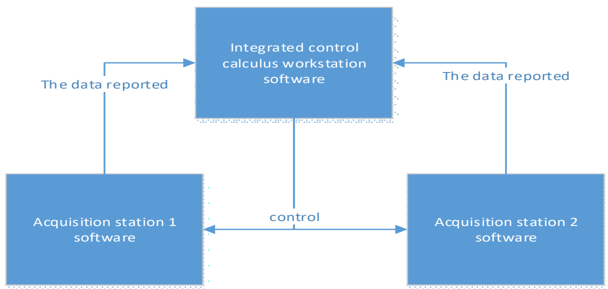


Fig.2. Software component

The specific composition of the acquisition station software is shown in Fig.3.

Fig.3. Acquisition station software composition

The working flow charts of the two acquisition stations are basically similar. After receiving the start instruction of the integrated control algorithm workstation, the two cameras are controlled for image acquisition at the same time, and the original grayscale collected by the cameras is converted into the target position information for information analysis and reporting. Motion blur usually invalids the vision-based measurement algorithms designated for static objects[4].

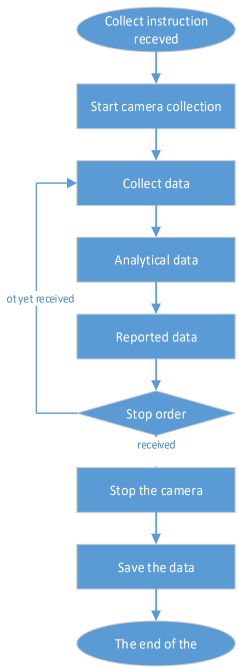


Fig.4. Acquisition station software flow

3.2 Integrated control calculus workstation software

The software of integrated control calculus workstation is shown in Fig.5.

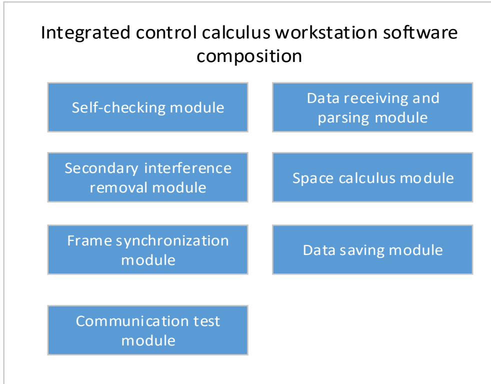


Fig.5. Integrated control calculus workstation software composition

The operator gives the test instructions in the integrated control algorithm workstation, which sends the instructions to acquisition station 1 and acquisition station 2, and then begins to receive the data returned from the two stations. The returned data are removed twice to eliminate the interference, and then the settlement code is carried out to obtain the ballistic characteristics and the coordinates of the falling point. Until the operator completes the test process.

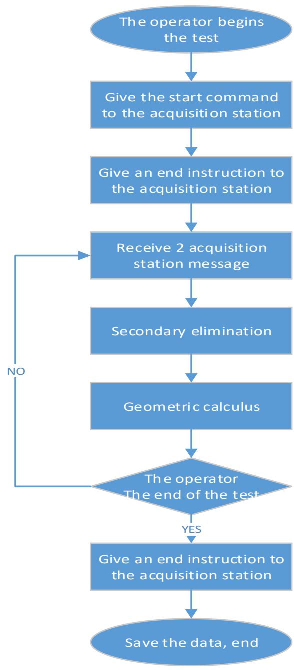


Fig.6. Integrated control calculus workstation software flow

3.3 Interference to rule out (Synchronization strategies)

Original frame information using relative timestamp, each frame of the original data after the acquisition is immediately on relative timestamp, namely frame number, camera first frame after each began experimenting with the timestamp of 0 x00 00 00 00 (data capacity of 0 ~ 4294967295 frames) accumulate a per frame, frame frequency calculated at 9.4 k/s material/s 10000, data capacity conversion for time:

429496 seconds = 7158 minutes = 119 hours material 4.9 days, enough to test the largest number. 4 cameras in the system start to collect at the same time, so the data with the same frame number is the collection information at the same time, so the system synchronizes the frame alignment.

In order to evaluate these errors, experimental tests on a CCD/CMOS ToF camera sensor have been done [5].

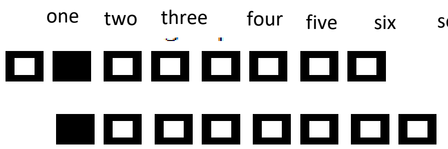


Fig.7 Experimental tests

4. Coplanar camera simultaneous trigger (acquisition computer processing)

The field of view of two linear array cameras connected by the same acquisition computer is coplanar and partially coincident. The effective target must pass through the overlap area of two field of view, and the effective target must trigger two opposite cameras at the same time[6]. Therefore, each triggered target can find 1 pair of data with the same timestamp in the system data, which are from opposite 2 cameras. No matching data could be triggered by the distracter at A and B. Therefore, the data that did not find a match was considered to be triggered by mistake. Interference removal is performed on the acquisition computer. It is a cost-effective alternative to the fundus camera[6].

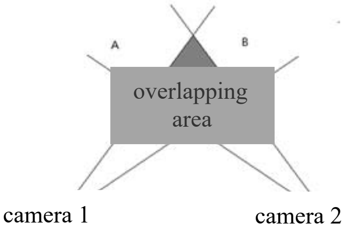


Fig.8. Coplanar camera

Target matching (integrated control algorithm workstation processing): According to the test method and equipment arrangement, the correct target must first trigger the two coplanar cameras connected by "computer 2", and then trigger the two cameras connected by "computer 1". And the process has a time difference t time difference.

"Computer 1" and "computer 2" are 60 meters apart, and the target speed is about 700 m/s.

T time difference = 60m ÷700m/s ≈ 0.00857s

Set the camera frame frequency at 9.4k/s.

The time stamp difference of the same target through 2 sets of acquisition computer is n frames:

N = t time difference * frame frequency = 0.00857 * 9.4k = 80.558 frames

As a rule, the data with too large velocity error is regarded as interfering object. However, taking into account the system error and the velocity error of the target object, this value will also change, so this value is taken as a configurable parameter and finally determined before the test, Comparative experiments show that, the proposed tracking method can successfully cope with the fast motion[3].

5. Geometric modeling (algorithm)

In Fig.9, let the camera height be a, the distance between the cameras AD=b, AB=c, and Get points in Q3(x,y,z):

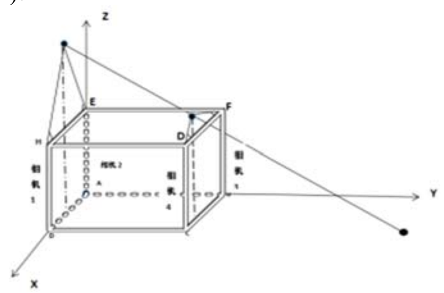


Fig.9 Geometric modeling

$$\frac{|Q_1H|}{\sin\beta} = \frac{|HZ|}{\sin(a+\beta)} = \frac{|Q_1E|}{\sin a} \quad (\triangle Q_1HE \text{ Sine theorem})$$

$$|Q_1H| = \frac{b\sin\beta}{\sin(a+\beta)}, \quad |Q_1H| = \frac{b\sin\alpha}{\sin(\alpha+\beta)}$$

$$\cos\beta = \frac{|OE|}{|Q_1E|}, \quad |OE| = \frac{b\sin\alpha\cos\beta}{\sin(\alpha+\beta)}$$

$$Q_1\left(0, \frac{b\sin\alpha\cos\beta}{\sin(\alpha+\beta)}, a + \frac{b\sin\alpha\sin\beta}{\sin(\alpha+\beta)}\right)$$

$$\frac{|Q_2F|}{\sin a_1} = \frac{|Q_2G|}{\sin\beta_1} = \frac{b}{\sin(\alpha_1+\beta_1)}$$

$$|Q_2F| = \frac{b\sin\alpha_1}{\sin(\alpha_1+\beta_1)} \qquad |Q_2G| = \frac{b\sin\beta_1}{\sin(\alpha_1+\beta_1)}$$

$$|Q_2O_2| = \frac{b\sin\alpha_1\sin\beta_1}{\sin(\alpha_1+\beta_1)}$$

$$|FO_2| = \frac{b\sin\alpha_1 + \cos\beta_1}{\sin(\alpha_1+\beta_1)}$$

$$Q_2\left(c, \frac{b\sin\alpha_1\cos\beta_1}{\sin(\alpha_1+\beta_1)}, a + \frac{b\sin\alpha_1\sin\beta_1}{\sin(\alpha_1+\beta_1)}\right)$$

$$\frac{y-\frac{b\sin a_1\cos\beta_1}{\sin(a_1+\beta_1)}}{y-\frac{b\sin\alpha\cos\beta}{\sin(\alpha+\beta)}} = \frac{x-c}{x} \qquad (1)$$

$$\frac{c}{x} = \frac{\frac{b\sin\alpha_1\cos\beta_1}{\sin(\alpha_1+\beta_1)} - \frac{b\sin\alpha\cos\beta}{\sin(\alpha+\beta)}}{y-\frac{b\sin\alpha\cos\beta}{\sin(\alpha+\beta)}} \qquad (2)$$

$$H_3m = \frac{z}{\tan\theta}$$

$$|H_2H_3| = \sqrt{\left(y-\frac{b\sin\alpha_1\cos\beta_1}{\sin(\alpha_1+\beta_1)}\right)^2 + (x-c)^2}$$

$$|mH_2| = |H_2H_3| - |\theta_3 m|$$

$$= \sqrt{\left(y-\frac{b\sin\alpha_1\cos\beta_1}{\sin(\alpha_1+\beta_1)}\right)^2 + (x-c)^2} - \frac{Z}{\tan\theta}$$

$$\cos\theta_1 = \frac{\frac{|H_3n|}{2}}{\tan\theta}$$

$$= \frac{|H_3k|}{\sqrt{\left(y-\frac{b\sin a_1\cos\beta_1}{\sin(\alpha_1+\beta_1)}\right)^2 + (x-c)^2}}$$

$$|H_3n| = \frac{(x-c)\frac{z}{\tan\theta}}{\sqrt{\left(y-\frac{b\sin\alpha_1\cos\beta_1}{\sin[\alpha_,+\beta_1)}\right)^2+(x-c)^2}}, \quad \frac{|H_3m|}{|mH|} = \frac{|H_3N|}{|nK|}$$

$$\frac{\frac{z}{\tan\theta}}{\sqrt{\left(y-\frac{b\sin\alpha_1\cos\beta_1}{\sin(\alpha_1+\beta_1)}\right)^2+(x-c)^2}-\frac{z}{\tan\theta}} = \frac{\frac{(x-c)\frac{z}{\tan\theta}}{\sqrt{\left(y-\frac{b\sin\alpha_1\cos\beta_1}{\sin(\alpha_1+\beta_1)}\right)^2+(x-c)}}}{x-c-\frac{(x-c)\frac{z}{\tan\theta}}{\sqrt{\left(y-\frac{b\sin\alpha_1\cos\beta_1}{\sin(\alpha_1+\beta_1)}\right)^2+(x-c)^2}}} \qquad (3)$$

Solve three equations (1), (2) and (3) of three variables x, y, and z to find the coordinates of Q_3（x,y,z）.

References

1. Wang Jianhua, Yang Yanxi, High-speed three-dimensional measurement technique for object surface with a large range of reflectivity variations, *Applied optics*, 2018(30): pp. 397.
2. Yuxiang Wu,Huimin Yue,Yong Liu, High-precision measurement of low reflectivity specular object based on phase measuring deflectometry, *Opto-Electronic Engineering*, 2017(08): pp. 772-780.
3. Peng-Cheng Zhang, De Xu, Tracking and Guiding Multiple Laser Beams for Beam and Target Alignment, *International Journal of Automation and Computing*, 2015, 12(06): 600-610.
4. Hance Zhou, Mingjun Chen, Liyan Zhang, Nan Ye, Cong Tao, Measuring shape and motion of a high-speed object with designed features from motion blurred images, *Measuremen*t, 2019, pp.145.
5. Filiberto Chiabrando, Roberto Chiabrando, Dario Piatti et al., Sensors for 3D Imaging: Metric Evaluation and Calibration of a CCD/CMOS Time-of-Flight Camera, *Sensors*, 2009, 9(12): pp. 339-405.
6. Biju Raju, N S D Raju, John Davis Akkara et al., Do it yourself smartphone fundus camera – DIYretCAM, *Indian Journal of Ophthalmology*, 2016, 64(9): pp. 129-131.

Research on Surface Defect Detection of Aluminum based on Image Processing

Xuemin Liu [1], Ce Bian [2*], Di Yin [3,4], Yuxuan Zhu [3,4], Yasheng Yuan [3,4], Hao Wang[3]

[1] *China petroleum engineering & construction corporation,*

[2] *Tianjin Tianke Intelligent and Manufacture Technology CO., LTD, China*

[3] *Tianjin University of Science and Technology, China*

[4] *Advanced Structural Integrity International Joint Research Centre, Tianjin University of Science and Technology, Tianjin, 300222, China*

*E-mail: * 969016280@qq.com*

Abstract

Aluminum material is relatively smooth. Aluminum surface engender scratches and bruises easily when collide with other metal. Surface defect detection of aluminum products is particularly important. It is very convenient to use machine vision method for defect detection. Defect contour extraction is an important part of machine vision for defect detection. The surface of aluminum metal is very reflective and shallow scratches are easily mistaken for defects. There are many kinds of filtering, such as the mean filtering, gauss filtering, median filtering and directed filtering. With the help of filtering, dynamic threshold can achieve a good effect. The severe scratch defect and the slight scratch can be clearly separated from the surface of the aluminum product.

Key words: defect detection, Aluminum metal, machine vision, Industrial Light source, Visual inspection

1. Introduction

In the traditional machining industry, the quality of the finished workpiece will be different due to various factors. Due to the improvement of industrial requirements, the quality requirements of customers are more and more stringent. At present, the quality of the products is judged by human eye in the piston processing factory made of aluminum alloy. This method not only has great labor intensity, but also has a high rate of missing detection in the finished workpiece. A simple method of image contrast can be used in a few cases. But there is no standard device that can detect complex defects such as scratches. The quality inspection requirements and difficulties of aluminum alloy pneumatic piston are studied in this paper.

The study of aluminum piston scratch detection is made in this paper. And the results showed that the workload of the staff was reduced by 3 / 1 and the miss rate was reduced to less than 1 / 1000.

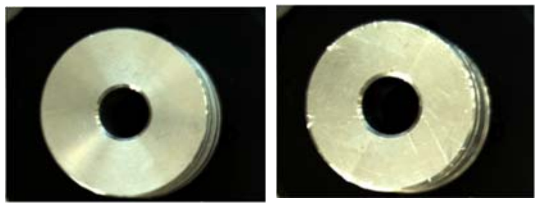


Fig.1 Normal product and Abnormal product

2. Defect detection method

The first is to build the experimental platform. A 2048 x 1536 pixel color industrial camera is mounted on a retractable camera mounting bracket. The camera captures at 10 FPS. The light source is an integral sphere blue light source. The distance between the camera and the workpiece is 300 mm. The sample taken is shown in the following figure.

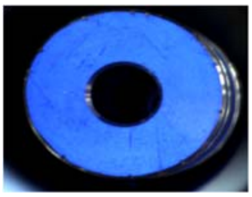


Fig.2 Blue Light

The camera is connected to the computer through the mini USB signal line for image acquisition. The entire research platform is shown in the figure below.

Fig.3 Experimental Platform

The mode and intensity of illumination have been determined and the inner wall of the spherical light source is hemispherical. It has the effect of converging light. The light comes from a 360-degree circular light source at the bottom and bounces off evenly. It equalizes the brightness of the entire image. It is suitable for metal surface defect detection.

Image processing is the key process of the whole system, which includes image acquisition, image pre-processing and defect extraction. The read image is collected by industrial camera. The image is then written using Halcon software. Image pre-processing is also the key step to determine whether the defect can be accurately extracted. For this reason, we have done the following processing.

The time domain image is transformed into the frequency domain image by Fourier affine transformation. Generate a bandpass filter with sinusoidal shape and set the distance of the filter's maximum from the DC term is 0.4.Use rft mode to set location of the DC term in the frequency domain.

Set the filter size as with original image size.

Convolve original image with the bandpass filter in the frequency domain, then we can get an image without high-frequency noise in frequency domain.

Fig.4 Image in frequency domain.

In order to give people a sense of what's going on, we need to compute the real-valued with fast Fourier transform for the image in frequency domain. After threshold setting we can get an image in time domain.

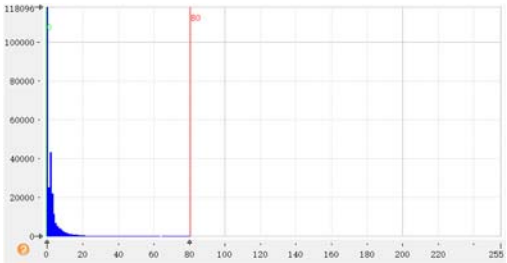


Fig.5 Gray histogram

Segment the image using global threshold to get an image with isolated each defection. Compute connected components of all region for next step. As show in the image below.

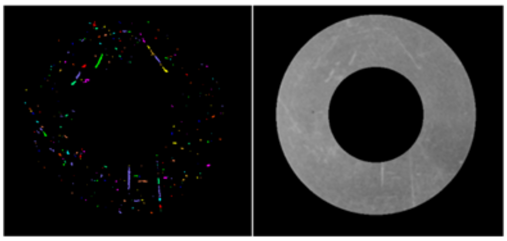


Fig.6 Influencing factor

The surface of aluminum alloy has the characteristics of gray level approaching to the defection. We're going to do the first selection of the defect area. At this point, the approximate defect region can be obtained without the effect of the material surface.

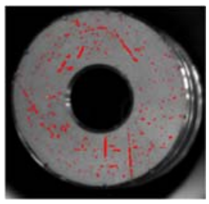


Fig.7 Defect display

In order to ensure that the image defect area is not missing, we dilate a region with a circular structuring element. The expanded region is extracted from the source image, It cut down the running time of the program and reduces the amount of data to process objects. In order to avoid separation of lines due to overlapping scratches, we need to extract complete lines from the image. We use the method of line segment fitting after extracting the region skeleton to connect the continuous defects. Select the length of the line segment to determine defect of scratches. The extracted lines are returned in Lines as sub pixel precise contours.

Based on the customer requirements we can get criteria for defection which size is 0.5 Square millimeter.

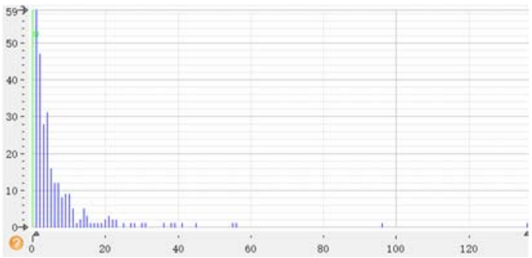


Fig.8 Area histogram

We can select the size of the area to get real defect region. And obtain a roughly selected defect position in the original image.

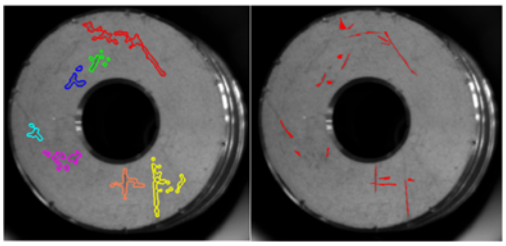


Fig.9 algorithm research

In this project, the pixel value of the scratches in the image can be placed in the lower level by changing the illumination. So that the degree of the scratches can be increased when the image is collected. Bandpass filter with sinusoidal shape enhancement of the high-frequency signal is better than most. So that we use 255 to subtract the Pixel value of the source image to achieve the image Pixel inversion effect.

The inverted image is transformed by Fourier affine transformation. Convert the image to the frequency domain. The next step is the same as before. we make the

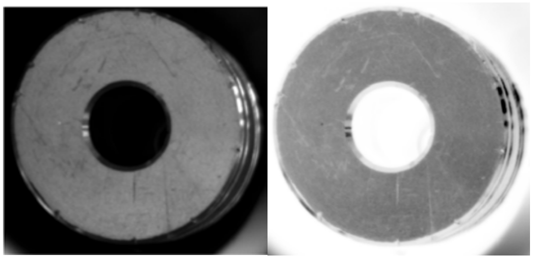


Fig.10 original image and inverted image

two images be convolved in the frequency domain. The pixels of the affine transformed image are multiplied by the corresponding pixels of the filter image. Compute the real-valued in time domain by fast Fourier transform of an image.

3. Testing and conclusion

3.1 The miss detection rate of detection has been greatly improved

Compare to the mode of manual inspect scratch defects, application of machine vision in scratch defect recognition has many advantages. The recognition ability can be improved a lot by adjusting parameters. The miss rate can be reduced to less than one in ten thousand. With the increase of working time, the miss detection rate of manual detection will increase. Unlike manual inspect, the vision detection system studied in this paper runs stably. The most important miss rate in detection attains one in a million. The results are as follows.

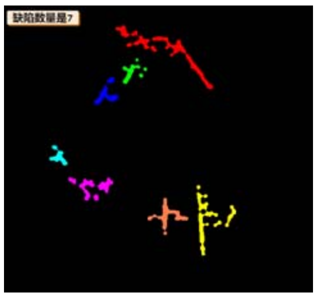


Fig11. Decision result

3.2 The detection time reduced

Compared to manual inspection of scratch defects machine vision well reduce the detection time. Using machine vision to detect surface scratch of Aluminum alloy material defects, the detection time is reduced to about 0.1 S. For different defects, the manual detection time may be vary greatly. And there is double-checking situation. The algorithm of this paper is stable in the detection period. The experimental results show that the detection time of the single piece is reduced to 0.127518 seconds. It can't do that manually。As shown in the image below.

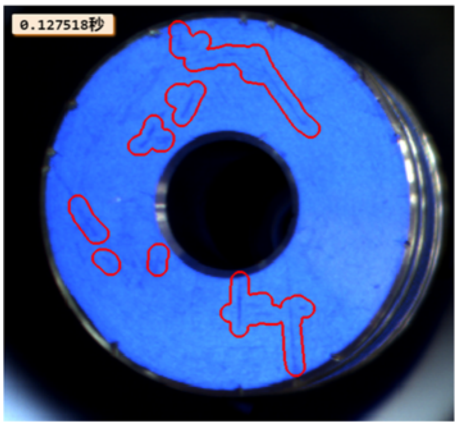


Fig12. Detection time

4. Conclusion

The quality inspection requirements of the workpiece and the problems existing in the current quality inspection are known by us. In this paper, machine visual inspection method is used to detect the surface scratch defect of Aluminum Alloy Cylinder Piston. The goal of reducing the Labor Force by 1 / 3 and the rate of missed inspection by less than 1 / 10000 was achieved. In the image processing process, the algorithm is optimized by studying the defect gray value. The algorithm is more accurate for defect detection. This paper also has some shortcomings. Depth of field has a huge impact on the algorithm. A surface with depth of field less than 0.05 will improve the detection rate.

Acknowledgements

The research is partly supported by the Project of Tianjin Enterprise Science and Technology Commissioner to Tianjin Tianke Intelligent and Manufacture Technology Co., Ltd (19JCTPJC53700). It is also supported by the Industry-University Cooperation and Education Project (201802286009) from Ministry of Education, China.

References

1. Lin su, Research on the key technology of fast detection of high reflection spherical defects based on vision. *Tianjin University,* 2013.
2. Yukui Wang, Hong Jia, Yongqing Chen, Wei Hou, Magnetic ring end surface defect detection based on machine vision. *Journal of Engineering*, 2010, 5: pp.63-68.
3. Hangxu Jin, Research on HALCON's vision measurement system for semi-axial parts. *North University of China*, 2018.
4. Yudan Wang, Research on surface defect detection based on machine vision. *University of Chinese Academy of Sciences*, 2018.
5. Chen Li, Research on the detection technology of weak defects in different attribute surfaces based on machine vision. *Zhejiang University,* 2018.
6. Haiyang Hou, Research and development of magnetic ring size and surface defect detection system based on machine vision. *Zhejiang University of Science and Technology*, 2016.
7. Huinan Wang, Cylindrical roller surface defect detection system based on machine vision. *Henan University of Science and Technology*, 2012.
8. Quan Li, Automatic detection and classification of surface defects. *Zhejiang University,* 2015.

Research on the Control of Multi Position Production Line based on PLC

Yong Hou [1]*, Hao Wang [1], Runhua Mao [2], Xuemin Liu [3]
[1] *Tianjin University of Science and Technology, China,*
[2] *Tianjin Tianke Intelligent and Manufacture Technology CO, LTD, China,*
[3] *China petroleum engineering & construction corporation, China*
*E-mail: * houyong@tust.edu.cn*
www.tust.edu.cn

Abstract

In view of the disadvantages and limitations of the traditional manual production line with the cylinder processing system as the main body, this paper proposes an intelligent production line control system based on PLC. The rearrangement of production line and automatic planning of production process are realized. The application of the system in practical production shows that the designed system can achieve 8 hours of unmanned automatic operation. It greatly improves production efficiency and product quality、reduces labor intensity and production cost ,and it makes the traditional production line intelligent and automatic.

Keywords: Cylinder, PLC, automatic control, production line, automatic production

1. Introduction

Production lines for processing products can be seen everywhere, but the traditional manufacturing production lines cannot achieve online control and real-time operation, so it is difficult to ensure operation accurately、fast and effectively of production lines. Based on the above discussion, the focus of this design is to add PLC intelligent equipment to the traditional production line equipment.

The object of this design is the automatic machining control system of the parts with the cylinder as the main body. A set of multi-station production line control system based on PLC and industrial computer is designed, which is applied to automatic machining of CU series cylinder cylinder parts with different cylinder diameters and strokes.20 series cylinder parts with multiple cylinder diameter and multiple stroke are completed on the same production line, and two workpieces can be clamped and unloaded at the same time and run in multi-station mode through the design of the program. At the same time, the main parameters of the automatic control system can be timely viewed and modified to achieve real-time controllable requirements.

2. The hardware structure design

This design of the control system is based on PLC as the control core of the system, with the industrial computer as the upper computer. First, Solidworks software platform is used to draw a schematic diagram and design a reasonable layout. Finally, sensors, solenoid valves, cylinders and electric cylinders are used as control execution components of the production line to realize automatic processing of cylinder cylinder parts.

Delta PLCDVP28SV11T was chosen as the controller in the design. The 5-position 3-pass SMC solenoid valve is selected as the pneumatic control valve. Figure 1 is the overall layout of the production line designed.

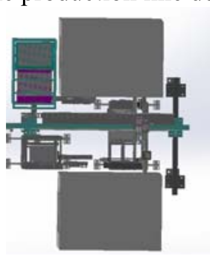


Fig.1 General layout of production line

- ***Selection of Servo System***

Motor type selection should be made according to the motion characteristics of the control object and the characteristics of the load, and power selection is also the primary one. Formula (2-1) can be applied to calculate the motor power to obtain the power of the servo motor to be selected.

$$P=F*V/100 \quad (1)$$

Considering that the motion of the mechanical arm of the automatic processing system, the cylinder body needs to be fast, stable and high-precision, the servo motor is chosen as the driving device of the mechanical arm. The control system uses delta's ASDA-A servo drive series. Next, draw the mechanical diagram required by this design through solidworks, as shown in figure 2.

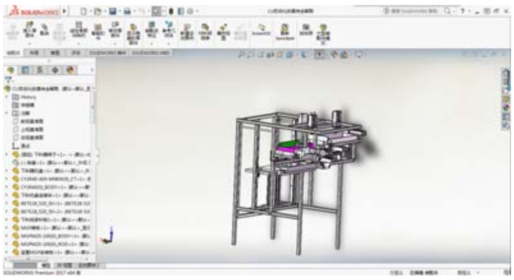


Fig.2. Mechanical drawings drawn with solidworks

3. Overall operation design

The design of the overall operation of this paper is combined with the process of technology、the overall layout and mechanism design for the overall operation of the system control. The overall operation of the design is #1 and #2 machining center as the center, mechanical arm as the transfer mechanism; Feeder, feeder and turnover mechanism are auxiliary mechanisms, which are logically matched with each other to complete automatic processing of cylinder body parts.

3.1. System Circuit Module Design

The power distribution cabinet of this design chooses 220V as the main power supply, and is equipped with the main power control switch to control the power supply on and off of the whole circuit; In order to ensure the safety of equipment and personnel, the main switch is equipped with an overcurrent circuit breaker; Each servo driver is connected to the power supply via a separate switch. The connection mode of communication between PLC and the touch screen RS485 is S/S and c0-c4 as the common end of the input and output sensor. Since the signal line derived from the previous servo IN1 is connected with the PLC terminal, a power circuit must be formed between IN1 and the common end of PLC to make the signal effective. The terminals here and CN1 share DC24V and 0V derived from CN1.

3.2 Operational Program Design

This design uses ISPSoft software to design the automatic start program、servo control program、 alarm program,、origin regression program and upper computer monitoring program. When operators operate the production line, users can choose automatic mode or manual mode according to their own needs. When the operator needs to run the automatic operation, it shall first carry out the origin regression to the device. When the origin regression is not completed, the automatic operation cannot be started, and the automatic operation can be started after the completion of the origin regression.

4. Assembly line design

The span length of the arm is 2,500mm, and the beam is made of aluminum standard parts. Flange joint is adopted between the aluminum profile of the beam and the support frame, which is connected with bolts to facilitate the transportation and installation of the equipment. The bottom of the frame and the ground are fixed with expansion bolts, so that the whole mechanical arm can have good stability and rigidity. Figure 3 is the design drawing of the robot arm.

4.1. Design of Feeding Mechanism

The function of feeding mechanism is automatic feeding of parts. The feeder is composed of silo, separating station and grasping station. The bin position of the feeder takes up most of the space of the feeder. The bin capacity design is calculated according to the processing time of a workpiece, and the single feeding can meet the requirement of 8 hours of operation. The feeder adopts cylinder and electric cylinder as actuators. The sensor is used to collect information, and the automatic feeding of parts is completed by logic control. Feeding mechanism is shown in figure 4.

Fig.3 The design of the manipulator

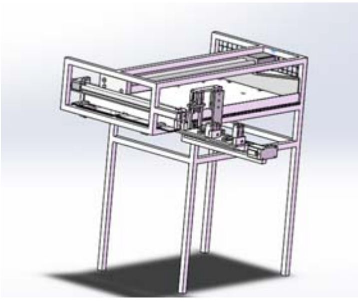


Fig.4. Design drawing of feeding mechanism

4.2. Design of Flip Mechanism

The turnover mechanism consists of the turnover station, the cleaning station and the preparation station. After the completion of the first working procedure, the parts shall be turned over and the surface chips shall be cleaned to prepare for the second working procedure. Two guides are installed at the bottom of the tilting mechanism, whose function is to allow the tilting mechanism to move left and right as a whole. The flip mechanism is shown in figure 5.

4.3. Design of Blanking Mechanism

The part of the feeder is composed of the feeding place, the workpiece vertical place, the spacer place, the tray and the basket. The function of the feeder is to automatically stack the finished parts and then add spacers between the rows in the process of automatic stacking. It plays an isolation and protection role for the processed workpiece and prevents the processed parts from colliding with each other in the handling and subsequent processes. The feeder is shown in figure 6.

Fig.5. Design of the flip mechanism

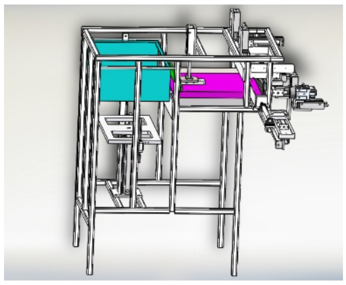


Fig.6. Structural drawing of feeder

5. Conclusion

100 data statistics were made for 20 models of different cylinder diameters and different strokes processed by the production line. The results show that the production line has reached the operation requirements of the enterprise, and the production line has been put into actual production. The whole process of automatic operation of eight hours, the design can be multi - cylinder diameter and multi - stroke processing and multi - station loading and unloading. During the operation of mechanical arm, the material taking and feeding work can be completed within 10 seconds, which shortens the time compared with manual material unloading and improves the processing efficiency. The field operation diagram is shown in figure 7.

Fig.7. Field diagram

Acknowledgements

The research is partly supported by the Project of Tianjin Enterprise Science and Technology Commissioner to Tianjin Tianke Intelligent and Manufacture Technology Co., Ltd (19JCTPJC53700). It is also supported by the Industry-University Cooperation and Education Project (201802286009) from Ministry of Education, China.

References

1. Roberts A. W, *Advances in the design of mechanical conveyors*. Bulk Solids Handing，1994, 14(2): pp.255-281.
2. SolidWorks 2004 (Chinese version) basic and application tutorial, *Beijing: Beijing institute of technology press*, 2004: pp.10-39.
3. Zhihu Liu, Fei Wang, Youmin Wang, Ming li, Design of manipulator control system for packaging bag based on PLC, *Hefei: Anhui University of Engineering,* 2015.
4. Runhua Mao, *Study on automatic machining system of cylinder body based on PLC*, Tianjin University of Science and Technology, 2018.
5. Wang Guohe, Li Weiguang, Trajectory plan-Ning and simulation for valve electric upset-Tingforming robot, *Machinery Design and Manufacture*, 2012, 0-0152-03.
6. Clavel, R, A Fast Robot with Parallel Geometry, *Industrial Robots*, 1988, 60(6): pp.91-100.
7. *Japan Robot Association (JAPA)*, Production and shipments of Manipulators and Robots by Application Areas (based on Value) (JAPA Member Companies), 2012.
8. Yong Hou, Runhua Mao, Xianzhang He, Zhang Xiaoyu, li Jintong, Design of CNC multi-station automatic loading and unloading mechanical arm, *Manufacturing automation*.

Research on Acoustic Source Localization Based on Acoustic Holography

Xiuqing Wang, Xiaoyun Jia , Jimin Zhao*
College of Electronic Information and Automation,
Tianjin University of Science and Technology, China
E-mail: lwxqly@163.com,1094622439@qq.com, zhaojiminemail@tust.edu.cn
www.tust.edu.cn

Abstract

According to statistically optimal cylindrical near-field acoustic holography (SOCNAH), Simulation analysis of single and multiple sound sources is established. In this paper, the influence of noise on multiple sound source signal is also studied. In addition, the experimental scheme of data acquisition is designed. Simulation and experimental results show that the method based on SOCNAH can realize the sound source location of pipeline and determine the damaged state of pipeline.

Keywords: Statistically optimal cylindrical near-field acoustic holography, single and multiple sound sources, Data acquisition, the sound source location

1. Introduction

Pipeline system is widely used in industrial sectors such as electric power, petroleum, chemical industry, natural gas and urban water supply.[1]At present, pipeline transportation is an important means of transportation. In order to solve the transportation problem of energy such as oil, natural gas, across the country have laid a lot of long-distance pipeline. Oil and gas pipelines are subject to external forces and corrosion due to its own material aging. During the period of work, it will produce a certain degree of damage and cause leaks occur due to internal and external pressure difference. Meanwhile, Pipeline damage will injure national economic interests and even endanger life safety.[2] Therefore, the detection of pipeline damage is very important.

Near-field acoustic holography is a technique for locating and identifying sound sources on the theoretical basis of acoustic radiation. It is not only used for locating and identifying noise sources, but also predicted the radiation characteristics of sound sources in the sound field.[3] Compared with near-field acoustical holography, statistical optimal acoustical holography algorithm avoids the spatial Fourier transform complicated calculation and reduces calculation error. It also has the advantages of high precision and computing speed. In addition, the algorithm, which can effectively solve the problem of large size, is suitable for the cylinder and flat. So the algorithm in the noise of the cylinder identification and location have a certain advantage.

In the space domain of holographic cylinder, SOCNAH, which can acquire the sound field of reconstructed surface by linear superposition of complex sound pressure, does not need to use spatial Fourier transform to achieve. It overcome the window from the theory source and winding error.[4-6] In this paper, based on the statistical optimal cylindrical near-field acoustic holography method, the sound source reconstruction simulation is carried out to obtain the best simulation parameters, acquire the cylinder sound source location and identification method, establish the conformal surface pipeline experimental model, and then judge the pipeline damage state.

2. Based on the theory of SOCANH

The Helmholtz equation can be obtained by separating variables in cylindrical coordinate system

$$p(r,\theta,z)=\frac{1}{2\pi}\sum_{-\infty}^{+\infty}e^{in\theta}\int_{-\infty}^{+\infty}p_n\left(a,k_z\right)\times e^{ik_z z}\frac{H_n^{(1)}\left(k_r r\right)}{H_n^{(1)}\left(k_z a\right)}dk_z \quad (1)$$

After the superposition of the unit cylindrical wave determined by the wave number vector *K=(n, kz)* in the spatial frequency domain, the sound pressure on the cylinder $r = a$ can be expressed as.

$$\Phi_k(a,\theta,z)=e^{in\theta}e^{ik_z z} \quad (2)$$

The radiated sound pressure from $\Phi_k(a,\theta,z)$ to cylinder $r=r_h$ and reconstruction surface $r=r_s$ can be expressed as.

$$\Phi_K(r_h,\theta,z)=e^{in\theta}e^{ik_z z}\frac{H_n^{(1)}(k_r r_h)}{H_n^{(1)}(k_r a)} \quad (3)$$

$$\Phi_K(r_s,\theta,z)=e^{in\theta}e^{ik_z z}\frac{H_n^{(1)}(k_r r_s)}{H_n^{(1)}(k_r a)} \quad (4)$$

In Eq(2), (3) and (4), $e^{in\theta}$ and $e^{ik_z z}$ represent cylindrical wave functions.[6]

Fig.2.1 represents the position relationship among *r=a*, *r=rs* and *r=rh*. *r=a* represents the minimum cylinder radius surrounded by sound source, *r=rh* represents the radius of the holographic surface, *r=rs* represents the radius of the reconstructed surface

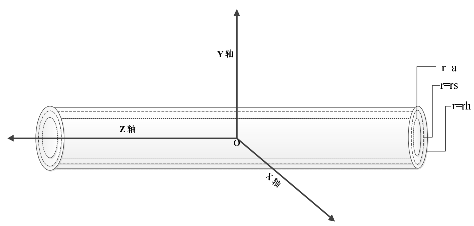


Fig. 2.1 the position relationship among holographic plane and reconstructed plane and minimum cylinder

If integral calculation is regarded as sum calculation, the sound pressure value can be obtained on the reconstructed surface:

$$p(r_s,\theta,z)=\frac{1}{2\pi}\sum_{m=1}^{M}p(a,K_m)\Phi_{K_m}(r_s,\theta,z) \quad (5)$$

According to the superposition theory of sound field, the unit cylindrical wave whose vector is *km* at any point *rs=(rs ,θ, z)* can be superimposed by the unit cylindrical wave whose vector is *km* at point *rH=(rH, θn, zn)*, [6] which can be expressed as:

$$\Phi_{K_m}(r_s)=\sum_{n=1}^{N}C_n(r_s)\Phi_{K_m}(r_{H_n})\ \ m=1,2,.....,M \quad (6)$$

Therefore, in space, the acoustic waves of the unit cylinder are superimposed on the holographic surface. By reconstructing the acoustic waves of the unit cylinder, the cylindrical waves at a point can be obtained.

Substitute Eq(5) into Eq(6) and get:

$$p(r_s)=\frac{1}{2\pi}\sum_{m=1}^{M}p(a,K_m)\sum_{n=1}^{N}C_n(r_s)\Phi_{K_m}(r_{H_n})=\sum_{n=1}^{N}C_n(r_s)p(r_{H_n}) \quad (7)$$

If
$$b=\begin{bmatrix}\Phi_{K_1}(r_s)\\ \cdots\\ \cdots\\ \Phi_{K_M}(r_s)\end{bmatrix} \quad (8)$$

$$c=\begin{bmatrix}c_1(r_s)\\ \cdots\\ \cdots\\ c_N(r_s)\end{bmatrix} \quad (9)$$

$$A=\begin{bmatrix}\Phi_{K_1}(r_{H_1}) & \cdots & \cdots & \Phi_{K_1}(r_{H_1})\\ \cdots & \cdots & \cdots & \cdots\\ \cdots & \cdots & \cdots & \cdots\\ \Phi_{K_1}(r_{H_1}) & \cdots & \cdots & \Phi_{K_1}(r_{H_1})\end{bmatrix} \quad (10)$$

The above linear system of equations is transformed into a matrix, and the following equation is obtained:

$$b = Ac(r_s) \quad (11)$$

The regularization solution of Eq(11) can be obtained:

$$c(r_s)=(A^{+}A+\varphi^2 I)^{-1}A^{+}b \quad (12)$$

A^{+} ------- the conjugate transpose of the matrix A;

φ -------regularization parameter;

I ---------the unit diagonal matrix;

$(A^{+}A+\varphi^2 I)^{-1}$ is the inverse of $(A^{+}A+\varphi^2 I)$

Substitute Eq(12) into Eq(7) to obtain the sound pressure value on the reconstructed surface

$$p(r_s)\approx\sum_{n=1}^{N}C_n(r_s)p(r_{H_n})=p^T c(r_s)=p^T(A^{+}A+\varphi^2 I)^{-1}A^{+}b \quad (13)$$

Eq(13) is the reconstruction formula based on the statistical optimal algorithm.

3. The experimental model

The acoustic emission signal of pipeline damage is a erupting wave and can be used as a point sound source. Most of the pipelines are the state of cylinder. Therefore, during the reconstruction of sound field, the holographic surface of cylinder is selected to establish the experimental model of conformal surface (cylinder-cylinder). The space rectangular coordinate system

(r,θ,z) is established of the inner center of the pipe as the origin. If the position of AE source is (x,y,z), let's take $x = r\sin\theta$, $y = r\cos\theta$, the coordinate position can be converted to (z,θ). The coordinates of sound source points are shown in fig.3.1.

Fig.3.1 experimental model diagram

As shown in fig.3.1, in the sound source point $M_{(z,\theta)}$, the value of z is the transverse distance between point M and coordinate $x-y$, and the value of θ is the Angle between the projection of point M on the y axis and the two lines on the x axis. This research establish an experimental model based on this coordinate.

4. Simulation experiment and analysis

4.1. *Simple sound source simulation*

According to the characteristics of the damaged sound source, the overclocking frequency signal is usually selected for analysis when detecting whether there is a crack in the pipe. In this experiment, the sound source frequency is selected as 60kHz, 150kHz and 200kHz respectively, the point sound source radius is selected as 0.06m, the vibration velocity is selected as 2.5 $m.s^{-1}$, the propagation speed in the pipe is selected as C=5200 $m.s^{-1}$. The main parameters affect the simulation results include: reconstruction source frequency range, reconstruction radius range, the number of holographic measurement points, etc. According to the actual situation, simulation parameters are selected to obtain the best simulation results.

Firstly, the sound source frequency is selected as 60kHz, the point sound source radius *r0*=0.06m, the reconstructed cylinder radius *rs*=0.07m, the holographic cylinder radius *rh*=0.075m, the holographic cylinder length *L*=10m, and the measurement point *M**N* =8*8. The simulation results are shown in fig.4.1.1. The sound source is located near (-3.5m,-128°).

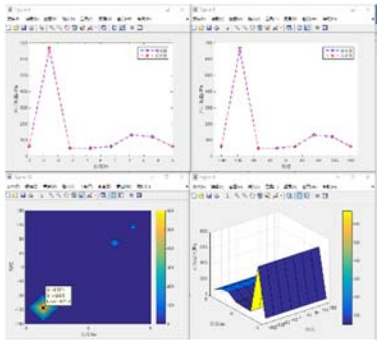


Fig.4.1.1*f*=60kHz,*r0*=0.06m,*rs*=0.07m,*rh*=0.075m,*L*=10m,*M*=8,*N*=8

Other parameters remained unchanged, and the distance was adjusted to d=0.03m between the holographic plane and the reconstructed plane. As shown in fig.4.1.2, the sound source can still be identified, and the position of the sound source is located near (-3.5m,-128°).

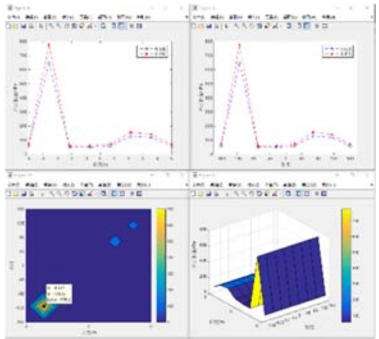


Fig.4.1.2*f*=60kHz,*r0*=0.06m,*d*=0.03m ,*L*=10m,*M*=8,*N*=8

Other parameters unchanged, change the sound source frequency *f*=150 kHz and the distance was adjusted to *d* =0.13m between the holographic surface and the reconstructed surface. As shown in fig.4.1.3, the sound source could still be identified and the position of the sound source is located near (-3.5m,-128°).

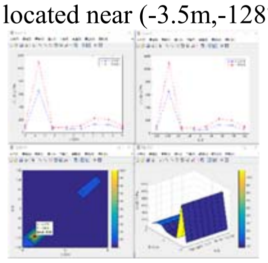


Fig.4.1.3*f*=150kHz,*r0*=0.06m,*d*=0.13m ,*L*=10m,*M*=8,*N*=8

Other parameters remain unchanged, change the sound source frequency *f*=200 kHz; the distance was adjusted to *d*=0.43m between the holographic surface and the reconstructed surface; change the measurement points *M**N*=11*11 of the holographic surface. As shown in fig.4.1.4, the sound source can still be identified, and the position of the sound source is located near (-4m,-144°).

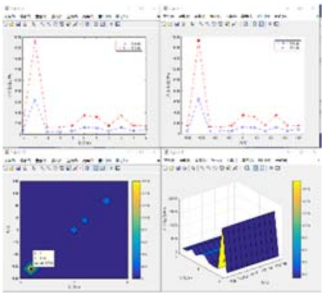


Fig.4.1.4*f*=200kHz, *r0*=0.06m,*d*=0.43m ,*L*=10m,*M*=11,*N*=11

Based on the above simulation, a large number of simple sound source simulation is done in the later stage, and a large number of simulation results were compared and sorted out, as shown in table 4-1.

Table 4-1 Reconstruction simulation positioning results

Grouping	*f*	*r0*	*d*	*L*	*M*N*	results
1	60	0.06	0.001	10	8*8	excellent
2	60	0.06	0.005	10	8*8	excellent
3	60	0.06	0.03	10	8*8	better
4	60	0.06	0.43	10	8*8	bad
5	60	0.06	0.005	10	11*11	better
6	150	0.06	0.001	10	8*8	excellent
7	150	0.06	0.005	10	8*8	better
8	150	0.06	0.03	10	8*8	good
9	150	0.06	0.43	10	8*8	bad
10	150	0.06	0.005	10	11*11	better
11	200	0.06	0.001	10	8*8	excellent
12	200	0.06	0.005	10	8*8	better
13	200	0.06	0.43	10	8*8	bad
14	200	0.06	0.005	10	11*11	better
15	200	0.06	0.43	10	11*11	bad

According to the above table 4-1, as the distance between the holographic plane and the reconstructed plane increases, the reconstruction error will increase and the reconstruction accuracy will decrease. The higher the value of sound source frequency will lead to the lower the reconstruction accuracy. However, the smaller the value of frequency will easy to locate the sound source. As the distance between the holographic plane and the reconstructed plane increases and the number of holographic measuring points increases, the reconstruction accuracy decreases gradually. The difficulty of actual operation will increase accordingly. On the premise that the reconstruction accuracy is not affected, the measurement accuracy is reduced as far as possible considering the actual measurement problem.

4.2. *Multiple sound source simulation*

According to actual situation of oil and gas pipelines, generally geographic environment is noisy. Multiple frequency multiple sound source simulation is carried out on the basis of simple sound source. The gauss white noise is introduces into multiple sound source of *f1*=60KHZ, *f2*=120KHZ, *f3*=150KHZ, the selection of measurement point *M*N*=8*8, length *L*=2m, radius of point source *r0*=0.06m, rebuild radius *rs*=0.065m, holographic radius *rh*=0.066m. The simulation results are shown in fig.4.2.1

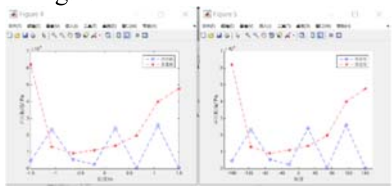


Fig.4.2.1 reconstructed sound pressure diagram without filtering

Based on the above fig.4.2.1, without filtering, noise is introduced into the simulation of multiple sound source which is seriously affected by noise.

In order to eliminate the noise influence problem and realize the identification and location of multiple sound sources. Window function is selected to filter the noise. Its function expression are shown in Eq(14) and Eq(15):

$$W(n,k_z)=W(n/r_H)W(k_z) \tag{14}$$

$$W(y)=\begin{cases}1-\frac{1}{2}e^{[(|y|/k_c-1)/a]} & |y|\le k_c\\ \frac{1}{2}e^{[(1-|y|/k_c-1)/a]} & |y|>k_c\end{cases} \tag{15}$$

kc------cut-off wave number;

a------the steepness coefficient of the window function

The selection of cut-off wave number can be calculated according to the formula *kc=o.6π/ls*, *ls* represents the interval distance of the sampling point.[7] According to the best possible result, the specific value of cut-off wave number should be adjusted. The steepness factor is also adjusted for the best possible result. After several simulation experiments, it is found that the best filtering effect can be achieved when the cut-off wave number is *kc*=18 and the steepness coefficient of the window function is *a*=15. Therefore, *kc*=18 and *a*=15 are taken for filtering.

Selecting the same parameters as above: three sound source frequency *f1*=60 kHz, *f2*=120 kHz, *f3* =150 kHz, point sound source radius *r0*=0.06m, reconstruction surface radius rs=0.065m, holographic surface radius rh=0.066m, length L=2m, measurement point M*N=8*8. Gaussian white noise is introduce into it and filter it. The simulation results are shown in fig.4.2.2:

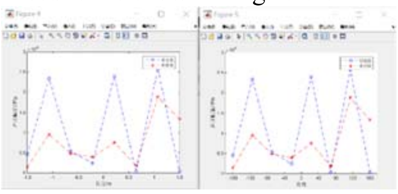


Fig.4.2.2 reconstructed sound pressure diagram after filtering

According to the above fig.4.2.2, the position of the sound source is located near the positions of (-1m, -124°), (0.3m, 36°) and (1.2m, 135°) respectively. By comparing fig.4.2.1 and fig.4.2.2, it can be concluded that the influence of noise is improved and the sound source can be identified by adding the window function.

5. Experimental scheme for data acquisition

After several times of simulation, different parameters were simulated and debugged. According to the simulation results, the feasible experimental parameters were obtained. Considering the actual parameters of the pipeline, reasonable experimental parameters were selected to establish scheme of data acquisition.
The data acquisition system are shown in fig.5.1:

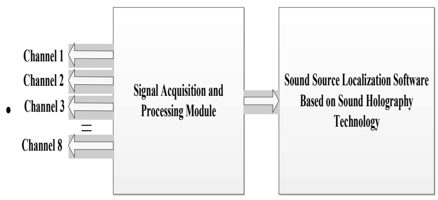


Fig.5.1 The data acquisition system

In this paper, active ultrasonic signal acquisition method is adopted to obtain acoustic emission signal, that is, multi-channel piezoelectric sensor PK15I is used to directly obtain acoustic emission signal. As shown in fig.5.1, the multi-channel piezoelectric sensor is connected with the signal acquisition processing module through multiple channels, and lead breaking experiment is done near the sensor. The acoustic emission signal is analyzed by the signal acquisition and processing module, and then the acoustic emission signal data is imported into the acoustic source positioning software to obtain the positioning result and determine the state of the pipeline.

6. Conclusion

In this paper, based on SOCNAH, a cylindrical - cylindrical acoustic holographic experimental model was established. Based on the model parameters of the actual pipeline, the values of experimental parameters are constantly changed such as frequency range, radius range, and number of holographic measurement points to carry out simulation experiment analysis of simple sound source and multiple sound source. Through a large amount of data simulation, the best sound source location results was acquired. According to the actual situation, the actual experiment parameters was chosen. Acoustic emission signal is obtained by data acquisition system. Based on acoustical holography technology, after the acoustic emission data of dealing with and analysis, the data is import into positioning software to determine sound location result, which eventually proved effective SOCNAH algorithm can accurately locate the sound source location. It can prevent major accidents in advance to ensure the safety of pipeline.

References

1. Penglin Zhang, Rujiao Wang, Yicheng Jiang, Mei Li. Pipeline leakage detection based on acoustic emission technology. *Journal of gansu science*, 2017, 29(03): pp. 86-91.
2. Liying Sun, Yibo Li, Zhigang Qu, Shijiu Jin, Yan Zhou. Study on leakage detection of acoustic emission pipeline by EMD signal analysis method. *Vibration and impact*, 2007, (10): pp.161-164+195.
3. Zhenshan Wang. Localization and identification of underwater motion noise source based on statistical optimal near field acoustic holography. *Harbin University of Engineering*, 2010.
4. Weibin Li, Jian Chen, Fei Yu, Xinzhao Chen. Statistical optimal cylindrical near-field acoustic holography. *Chinese journal of mechanical engineering*, 2005, 04: pp.123-127.
5. Steiner R, Hald j. near-field acoustical holography without the errors and limitations caused by the use of spatial DFT. *International Journal of Acoustics and Vibration*, 2001, 6(2): pp.83-89.
6. Xiuqing Wang, Haiyan Wang, Chunxia Zhang. Study on statistical optimal cylindrical near-field acoustic holography recognition of acoustic emission sources. *Acoustics technology*, 2013, 32(03): pp.175-180.
7. Yinchao Wang. Detection of plant disease stress based on near-field acoustic holography. *Tianjin University of science and technology*, 2013.

Research on Anti-interference Method of Electrical Fast Transient Pulse Group for Fire Alarm of Substation

Guo Wangyong *, Ju Zhenfu, Chen Guang, Zhu Bo

Beijing NARI Yihe Environmental Protection Technology Co., Ltd, China

*E-mail: * dx2233@126.com*

Abstract

Aiming at the problem of electrical fast transient (EFT) in substation, the mechanism of EFT generation and electromagnetic compatibility is analyzed. In this paper, the anti-interference method of fire alarm is designed for complex EFT in substation. The test results show that the design method can effectively suppress the influence of EFT on fire alarm in substation, improve the immunity level of EFT of fire alarm in substation, and meet the requirements of level 4 in GB / t17626.4-2006 standard.

Keywords: substation; electrical fast transient pulse group; anti-interference

1. Introduction

Power station is an essential component of the power system. In order to ensure the normal operation of the power station, a large number of power components and electronic products are running in the power station. These power components and electronic products will produce a variety of complex electromagnetic interference, which will not only affect the normal operation of the electronic equipment in the power station, but also cause serious damage to the electronic products themselves, endangering the safe operation of the power system, of which the most influential is the electrical fast transient pulse group.

As the central control core of alarm information, substation intelligent fire alarm not only requires the system to have accurate and stable control performance, but also requires the product to have a strong ability of anti electromagnetic interference, anti shock and anti harsh environment, so as to ensure that the fire alarm information is correct and reliable in any case and ensure the safety of the equipment. Most of the intelligent fire alarms in substations are designed by single-chip microcomputer technology. Electromagnetic interference is one of the main reasons for the failure or performance degradation of the intelligent fire alarm in substations 错误!未找到引用源。.

Based on the mechanism of EFT generation and disturbance, this paper studies the anti EFT method of the fire alarm, and puts forward the design method for the anti EFT of the power input of the fire alarm. The test results show that the proposed method can effectively improve the immunity level of EFT of fire alarm in substation.

2. Mechanism analysis of EFT generation and disturbance

Ground fault or switching inductive load will produce electrical fast transient pulse group disturbance to electronic system. The characteristic of the disturbance is that the disturbance signal is not a single pulse, but a series of pulses. On the one hand, because the pulse group can produce the accumulation effect at the input of the circuit, the amplitude of the disturbance level may eventually exceed the noise tolerance of the circuit. On the other hand, the period of the pulse group is short, and the interval time of each pulse wave is short. When the first pulse wave has not disappeared, the second pulse follows closely. For the input capacitance of the circuit, it starts to charge when the discharge is not completed, so it is easy to reach a higher voltage, which has a great impact on the normal operation of the circuit.

In power system, the operation of circuit breaker, relay and high voltage disconnector will cause a series of very fast transient over-voltage process, and the formation mechanism of EFT / B is closely related to this process. When these switches are closed or opened, breakdown will occur between the contacts, and compared with the process of

extinction when the switch is closed, the breaking of the switch will lead to repeated breakdown between the contacts. The amplitude of EFT / b voltage has a great relationship with the operation of the switch. The closing operation of the switch causes a decreasing voltage, while the breaking operation of the switch causes a growing process, as shown in Figure 1. The electrical fast transient pulse group caused by switching off operation is a kind of very fast transient over-voltage, which belongs to the wideband electromagnetic disturbance signal. The rise time of a single pulse is 5ns, the duration of a single pulse is 50ns, the pulse repetition frequency is 5KHz or 100kHz, which can reach dozens of MHz, the duration of the pulse group is 15ms, and the period of the pulse group is 300ms.

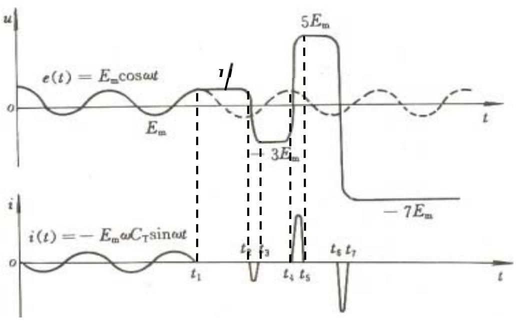


Fig.1 Change trend of voltage between breaking contacts of switch

3. Coupling analysis of electrical fast transient pulse group

The fire alarm controller realizes the functions of power input, signal acquisition and communication through the secondary cable connected to it, but the electrical fast transient pulse group will also be connected to the fire alarm controller through the secondary cable. The secondary cables in power system are all cables with metal shielding layer, which can effectively reduce the coupling of external electromagnetic interference to cable core and improve the EMC level of substation. The grounding of cable shielding layer shall adopt two-point grounding to shield the interference of external high-frequency electromagnetic field to cable core. When conducting the electrical fast transient pulse group test, when the electrical fast transient pulse group current appears on the conductor in the circuit under test, because the current is transient, it will cause the magnetic field around the circuit 1 to change, so that the other conductor circuit 2 in the changing magnetic field on the circuit induces the electromotive force, and the signal of the conductor circuit 1 is coupled into the conductor electricity Circuit 2, causing interference to wire circuit 2.

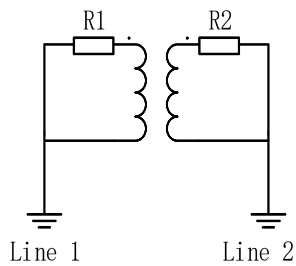


Fig.2 Equivalent circuit of inductive coupling

4. EFT protection method

EFT sensitive components and electrical fast transient pulse group can control EFT disturbance in a certain range through certain methods and measures, so that it is not enough to cause greater harm to electronic products. Based on the mechanism of EFT, two methods of EFT protection are proposed in this paper

In the design of power supply, the main measure to restrain the interference of electrical fast transient pulse group is filter. In the actual design of filter, the common mode choke and differential mode capacitor are used to form the filter. The common mode choke is made of ferrite beads and copper wires. The anti-interference ability of ferrite beads is related to its own characteristics and geometric size. For the same kind of core beads, the longer the length, the smaller the inner diameter and the larger the outer diameter, the better the suppression effect. Therefore, the impedance of ferrite beads increases with the increase of frequency. In the low frequency, the impedance is mainly composed of inductance component, and the impedance is very small. In the high frequency, the impedance is mainly determined by the resistance component, and the inductance component is still very small. Therefore, ferrite beads can effectively suppress the high frequency interference of EFT, and convert its energy into heat to dissipate.

In addition to taking pulse absorption and filtering measures for the interference direct transmission channel, in order to prevent EFT interference from radiating to other ports through space, so as to invade the sensitive equipment, the interference port and other ports should be divided in space, and appropriate shielding measures should be taken for other ports to suppress common mode interference.

5. Experimental analysis of EFT

Problem based on substation fire alarm According to GB / t17626.4, the test level of EFT is three, that is, the peak voltage is 2KV, the pulse rise time is 5ns, the duration is 50ns, and the repetition frequency is 5KHz. In the actual test, in order to fully understand the ability of the fire alarm of the substation to resist EFT interference current disturbance, the test is started from 0.5kV, 1kV and 2KV, and the specific test results are shown in Table 1. It is found that the alarm works normally when the voltage is below 1KV. Under the test level of 1kV and 2KV, the fire alarm in the substation works abnormally: display disorderly code, data rolling, etc.

Table 1 Test results of three levels of experimental voltage

Test level /kV	Result
0.5	normal
1	abnormal
2	abnormal

For the common mode interference current of EFT on the power line, the common mode inductance or common mode choke is usually used to suppress it. There are two common mode inductance coils in the common mode choke, which are respectively connected with L line and N line of the power line. The winding direction and turns of the two coils are the same, because the common mode current has the same phase characteristics, according to the electromagnetic field theory and Maxwell equations, the magnetic field generated by the two coils overlaps each other, the inductance of the coil increases correspondingly, showing high impedance, and the common mode current is consumed in the form of heat energy.

Table 2 Capacitance value

Capacitance	Capacitance value	Capacitance	Capacitance value
C_1	105	C_4	105
C_2	104	C_5	104
C_3	105	C_6	105

Because the common mode choke and capacitor are used together to form a low-pass filter, which can have a good suppression effect on the common mode EMI interference signal, the EFT filter structure shown in Figure 3 is designed in combination with the previous rectification experience. LX1, lx2 and LX3 are respectively wound by ferrite magnetic rings with outer diameter of 35mm, inner diameter of 25mm and height of 12mm. The coil is made of two copper wires and wound with ferrite magnetic rings for 25, 20 and 15 turns. The capacitance values are shown in Table 2.

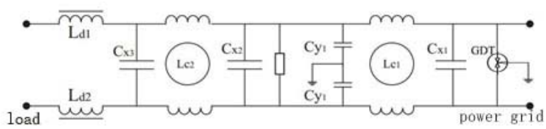
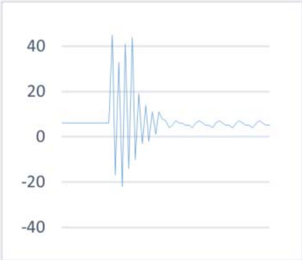
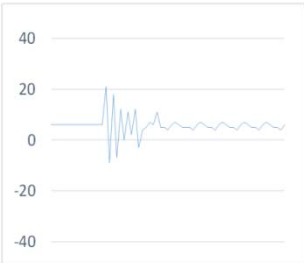


Figure 3 EFT filter structure

The power port is the main way for the electrical fast transient pulse group coupling into the fire alarm controller. The fire alarm controller uses a small volume, light weight, high efficiency switching power supply to provide the DC power it needs. The switch power supply of the fire alarm controller generally adopts AC220V power supply, with output of 5V, ± 15V and ± 24V DC voltage. At present, the fire alarm controller mainly adopts centralized power supply and distributed power supply. In order to effectively suppress the interference of the electrical fast transient pulse group to the power port of the microprocessor-based protection device, in addition to the improved distributed power supply, the microprocessor-based protection device can also introduce conventional overvoltage suppression and filtering measures to suppress the interference of the electrical fast transient pulse group to the power port.

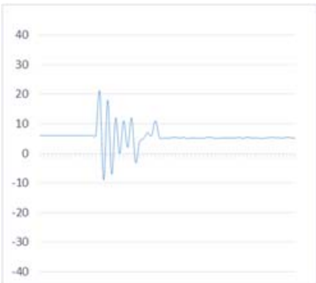


(a)



(b)

(c)



(d)

(a)、(c): No action taken;

(b)、(d): Steps taken

Fig.4 Waveform of voltage receiving fast transient pulse group interference without and with suppression measures

The purpose of shielding is to limit the induction or radiation of electromagnetic wave from one area of circuit or equipment to other areas. After the interference source is shielded, it can not radiate and diffuse the interference pulse; after the sensitive body is shielded, it can prevent from being affected by the external electromagnetic field.

EFT interference signal will affect other sensitive circuits through radiation. In order to reduce the radiation of the interference source and enhance the immunity of the sensitive body, the interference source and the sensitive body on the circuit board can be shielded. Combined with the above two measures, after the test, the substation fire alarm can work normally, and the display screen is normal.

6. Conclusion

Aiming at the problem of immunity of EFT produced by substation fire alarm, this paper analyzes the generation and disturbance mechanism of EFT, establishes the coupling model of EFT, puts forward two measures to restrain the disturbance of EFT, and carries out the test verification. The results show that the method proposed in this paper can effectively improve the level of EFT immunity of substation fire alarm, and meet the requirements of GB / t17626.4-2006 standard.

References

1. WangYufeng, Zou Jiyan, LiaoMinfu, Study of EFT/B generated in primary circuit and it scounter measure. *Electric Power Automation Equipment*, 2007, 27(9):22-27. (in Chinese)
2. XuL, WangS, ZhaiG, etal., Study on EMS of DCSS R under interference of EFT burst, *Proceedings of the 2011 6th IEEE Conference on Industrial Electronics and Applications*, Beijing, 2011: pp.1938-1942.
3. ZhouZ, Quan xingJ., Analysis of the effectiveness of ferrite core for improving EFT immunity, *Proceedings of the 2002 3rd International Symposium on Electro magnetic Compatibility*, 2002, pp.718-721.
4. Wang Rao, Li Kui, Guo Zhitao, etc., Research on the resistance of intelligent leakage circuit breaker to fast transient pulse group interference. *Power automation equipment*, 2012, 32(4): pp.129-133.
5. General Administration of quality supervision, inspection and Quarantine of the people's Republic of China. GB/T17626. 4-2006 electromagnetic compatibility test and measurement technology: Electrical fast transient pulse group immunity test. Beijing: China Standards Press, 2007.